Finite-Time Last-Iterate Convergence for Multi-Agent Learning in Games

Tianyi Lin [*1] **Zhengyuan Zhou** [*2] **Panayotis Mertikopoulos** [3] **Michael I. Jordan** [1]

Abstract

We consider multi-agent learning via online gradient descent (OGD) in a class of games called λ-cocoercive games, a broad class of games that admits many Nash equilibria and that properly includes strongly monotone games. We characterize the finite-time last-iterate convergence rate for joint OGD learning on λ-cocoercive games; further, building on this result, we develop a fully adaptive OGD learning algorithm that does not require any knowledge of the problem parameter (e.g., the cocoercive constant λ) and show, via a novel double-stopping-time technique, that this adaptive algorithm achieves the same finite-time last-iterate convergence rate as its non-adaptive counterpart. Subsequently, we extend OGD learning to the noisy gradient feedback case and establish last-iterate convergence results—first qualitative almost sure convergence, then quantitative finite-time convergence rates—all under non-decreasing step-sizes. These results fill in several gaps in the existing multi-agent online learning literature, where three aspects—finite-time convergence rates, non-decreasing step-sizes, and fully adaptive algorithms—have not been previously explored.

1. Introduction

In its most basic incarnation, online learning (Blum, 1998; Shalev-Shwartz et al., 2012; Hazan, 2016) can be described as a feedback loop of the following form:

1. The agent interfaces with the environment by choosing an *action* $a_t \in \mathcal{A} \subset \mathbf{R}^d$ (e.g., bidding in an auction, selecting a route in a traffic network).
2. The environment then yields a reward function $r_t(\cdot)$, and the agent obtains the reward $r_t(a_t)$ and receives some *feedback* (e.g., reward function $r_t(\cdot)$, gradient $\nabla r_t(a_t)$, or reward $r_t(a_t)$), and the process repeats.

As the reward functions $r_t(\cdot)$ are allowed to change from round to round, the standard metric that quantifies the performance of an online learning algorithm is that of regret (Blum & Mansour, 2007): at time T, the regret is the difference between $\max_{a \in A} \sum_{t=1}^{T} u_t(a)$, the total reward achieved by the best fixed action in hindsight, and $\sum_{t=1}^{T} u_t(a_t)$, the total reward achieved by the algorithm. In the extensive online learning literature (Zinkevich, 2003; Kalai & Vempala, 2005; Shalev-Shwartz & Singer, 2007; Arora et al., 2012; Shalev-Shwartz et al., 2012; Hazan, 2016), perhaps the simplest algorithm that achieves a minimax-optimal regret guarantee is Zinkevich's online gradient descent (OGD) algorithm, where the agent simply takes a gradient step (at the current action) to form the next action, performing a projection if necessary. Due to its simplicity and success in practice, it is one of the most widely-used algorithms in online learning theory and applications (Zinkevich, 2003; Hazan et al., 2007; Quanrud & Khashabi, 2015).

Increasingly, however, the most salient applications of online learning, particularly those in which the reward functions are changing over time, arise in the setting of multi-agent online learning. In this setting, each agent is making online decisions in an environment that consists of other agents who are simultaneously making online decisions and whose actions impact the rewards of other agents. Thus, each agent's reward is determined by an (unknown) game, and each agent's reward function, when viewed solely as a function of its own action, also changes, despite the fact that the underlying game mechanism is fixed. It is a stringent test of the universality of the OGD regret bounds that they apply in some form in the multi-agent setting. In particular, we have the following fundamental question in game-theoretic learning (Cesa-Bianchi & Lugosi, 2006; Shoham & Leyton-Brown, 2008; Viossat & Zapechelnyuk, 2013; Bloembergen et al., 2015; Monnot & Piliouras, 2017):

Can OGD learning, and more broadly no-regret learning, lead to Nash equilibria?

As an example, if all users of a computer network individually follow some no-regret learning algorithm (e.g., OGD)

[*]Equal contribution [1]University of California, Berkeley [2]New York University [3]Univ. Grenoble Alpes, CNRS, Inria, LIG, 38000 Grenoble. Correspondence to: Tianyi Lin <darren_lin@berkeley.edu>.

to learn the best route for their traffic requests, would the system eventually converge to a stable traffic distribution, or would it devolve to perpetual congestion as users ping-pong between different routes (like commuters changing lanes in a traffic jam)? Note that whether the process converges at all pertains to the stability of the joint learning procedure, while whether it converges to Nash equilibria pertains to the game-theoretic interpretation of the stable points: if the learning procedure converges to a non-Nash point, then each agent can do better by not following that learning procedure.

1.1. Related Work

Despite the seeming simplicity of this fundamental question, the existing literature has only provided partial and qualitative answers. This is in part due to the strong convergence mode implied by the question: much of the existing literature focuses on time-average convergence (i.e., convergence of the time average of the joint action), rather than last-iterate convergence (i.e., convergence of the joint action). However, not only is last-iterate convergence theoretically stronger and more appealing, it is also the only type of convergence that actually describes the system's evolution. This point was only recently expounded rigorously, by Mertikopoulos et al. (2018b), who show that even though follow-the-regularized-leader (another no-regret learning algorithm) converges to a Nash equilibrium in linear zero-sum games in the sense of time averages, actual joint actions orbit Nash equilibria in perpetuity. Motivated by this consideration, a growing literature (Krichene et al., 2015; Lam et al., 2016; Palaiopanos et al., 2017; Zhou et al., 2017; Mertikopoulos et al., 2017; Zhou et al., 2018; Mertikopoulos & Zhou, 2019) has aimed at obtaining last-iterate convergence results. However, all of these last-iterate convergence results are qualitative. In particular, except in strongly monotone games, where Zhou et al. (2020) very recently established a $O(\frac{1}{T})$ last-iterate convergence rate for OGD learning with noisy feedback[1] in strongly monotone games, there are no quantitative, finite-time last-iterate convergence rates available.[2]

Additionally, an important element in multi-agent online learning is that the horizon of play is typically unknown. As a result, no-regret learning algorithms need to be employed with a decreasing learning rate (e.g., because of a doubling trick or as a result of an explicit $\mathcal{O}(1/t^{\alpha})$ step-size tuning). In particular, in order to achieve last-iterate convergence-to-Nash results, existing work rests crucially on using decreasing step sizes (often converging to zero no slower than a particular rate) in their algorithm designs. This, however, embodies the following general tenet:

New information is utilized with decreasing weight.

From a rationality viewpoint, this tenet is counter-intuitive and it flies at the face of established economic wisdom. Instead of *discounting* past information, players end up indirectly *reinforcing* it by assigning negligible weight to recent observations relative to those in the distant past. This negative recency bias is unjustifiable from economic principles, and it does not accord with a plausible model of human/consumer behavior. These considerations raise another important open question, one that, if answered, can bridge the gap between online learning and rational microeconomic foundations:

Is no-regret learning without discounting recent information compatible with Nash equilibria?

1.2. Our Contributions

Reflecting on those two gaps simultaneously, we are thus led to the following research question, one that aims to close two open questions at once: **Can we obtain finite-time last-iterate convergence rate using only nondecreasing step sizes?**

Our goal here is to make initial but significant progress in answering this question, and our contributions are threefold. First, we introduce a class of games that we call *cocoercive*; these contain all strictly monotone games as a special case. We show that if each player adopts OGD, then the joint action sequence converges in a last-iterate sense to the set of Nash equilibria at a rate of $o(\frac{1}{T})$. The convergence speed more specifically refers to how fast the gradient norm squared converges to zero: note that in cocoercive games, gradient norms converge to zero if and only if the iterate converges to the set of Nash equilibria. To the best of our knowledge, this is the first rate that provides finite-time last-iterate convergence beyond the strong monotonicity setting.

Second, we study in depth the stochastic gradient feedback case, where each player adopts OGD in λ-cocoercive games, with gradient corrupted by zero-mean, martingale-difference noise, whose variance is proportional to current gradient norm squared (this is the relative random noise model due to Polyak (1987)). In this more challenging setting, we first establish that the joint action sequence converges in a last-iterate sense to Nash equilibria almost surely under a constant step size. The previous best such qualitative convergence is due to Mertikopoulos & Zhou (2019), who shows that such almost sure convergence is

[1] The perfect gradient feedback case has a last-iterate convergence of $O(\rho^T)$, for some $0 < \rho < 1$ when the game is further Lipchitz. This follows from a classical result in variational inequalities (Facchinei & Pang, 2007).

[2] Except in convex potential games. In that case, the problem of converging to Nash equilbria reduces to a convex optimization problem, where standard techniques apply.

guaranteed in a variationally stable game. Despite the fact that variationally stable games contain cocoercive games as a subclass, our result is not covered by theirs because Mertikopoulos & Zhou (2019) assume a compact action set, where we consider unconstrained action sets—a more challenging scenario since the action iterates can be unbounded *a priori*. Our result is also unique in that constant step size is sufficient to achieve last-iterate almost sure convergence, while Mertikopoulos & Zhou (2019) require decreasing step sizes (that is, square-summable, but not summable). Note that the relative random noise model is necessary for obtaining such a constant-step-size result: in an absolute random noise model (where the second moment of the noise is bounded by a constant), the gradient descent iterate forms an ergodic and irreducible Markov chain, which induces an invariant measure that is supported on the entire action set, thereby making it impossible to obtain any convergence-to-Nash result. We then proceed a step further and characterize the finite-time convergence rate. We establish two rates here: first, the expected time-average convergence rate is $O(\frac{1}{T})$; second the expected last-iterate convergence rate is $O(a(T))$, where $a(T)$ depends on how fast the relative noise proportional constants decrease to zero. As a simple example, if those constants decrease to zero at an $O(\frac{1}{\sqrt{t}})$ rate, then the last-iterate convergence rate is $O(\frac{1}{\sqrt{T}})$. For completeness, we also present (in the appendix) a parallel set of results—last-iterate almost sure convergence, time-average convergence rate and last-iterate convergence rate—for the absolute random noise model.

Third, and even more surprisingly, we provide—to the best of our knowledge—the first adaptive gradient descent algorithm that has last-iterate convergence guarantees for games. In particular, the online gradient descent algorithms mentioned above—both in the deterministic and stochastic gradient case—require the cocoercive constant λ to be known beforehand. Thus, this calls for adaptive variants that do not require such knowledge. In the deterministic setting, we design an adaptive gradient descent algorithm that operates without needing to know λ and adaptively chooses its step size based on past gradients. We then show that the same $o(\frac{1}{T})$ last-iterate convergence rate can be achieved with a nondecreasing step size. Previously, the closest existing result is that of Bach & Levy (2019), who provided an adaptive algorithm on variational inequality with time-average convergence guarantees. Providing adaptive algorithms for last-iterate convergence is much more challenging. Moreover, Bach & Levy (2019) require the knowledge of the diameter of the domain set (which they assume to be compact) in their adaptive algorithm, whereas we operate in unbounded domains. Our analysis relies on a novel double-stopping-time analysis, where the first stopping time characterizes the first time until gradient norm starts to decrease monotonically, and the subsequent second stopping time characterizes the first time the underlying pseudo-contraction mapping starts to converge rapidly. Finally, we also provide an adaptive algorithm in the stochastic gradient feedback setting and establish the same finite-time last-iterate convergence guarantee.

2. Problem Setup

In this section, we present the definitions of a game with continuous action sets, which serves as a stage game and provides a reward function for each player in an online learning process. The key notion defined here is λ-*cocercivity*, which is weaker than λ-strong monotonicity and covers a much wider range of games in real applications.

2.1. Basic Definition and Notation

Throughout this paper, we focus on games played by a finite set of *players*, $i \in \mathcal{N} = \{1, 2, \ldots, N\}$. During the learning process, each player selects an *action* $\mathbf{x}_i$ from a convex subset $\mathcal{X}_i$ of a finite-dimensional space $\mathcal{V}_i$ with the induced norm $\|\cdot\|$. The reward is determined by the profile $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ of all players' actions.

Definition 2.1 *A continuous game is a tuple $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$, where $\mathcal{N}$ is the set of N players $\{1, 2, \ldots, N\}$, $\mathcal{X}_i$ is a convex set of some finite-dimensional vector space $\mathbb{R}^{d_i}$ representing the action space of player i, and $u_i : \mathcal{X} \to \mathbb{R}$ is the i-th player's payoff function.*

Regarding the players' payoff functions, we make the following assumptions: (i) for each $i \in \mathcal{N}$, the function $u_i(\mathbf{x})$ is continuous in $\mathbf{x}$; (ii) for each $i \in \mathcal{N}$, the function u_i is continuously differentiable in $\mathbf{x}_i$ and the partial gradient $\nabla_{\mathbf{x}_i} u_i(\mathbf{x})$ is Lipschitz continuous in $\mathbf{x}$. The notation $\mathbf{x}_{-i}$ denotes the joint action of all players but player i. Consequently, the joint action $\mathbf{x}$ will frequently be written as $(x_i; \mathbf{x}_{-i})$. Two important quantities are specified as follows:

Definition 2.2 $\mathbf{v}(\mathbf{x})$ *is the profile of the players' individual payoff gradients; i.e., $\mathbf{v}(\mathbf{x}) = (v_1(\mathbf{x}), v_2(\mathbf{x}), \ldots, v_N(\mathbf{x}))$, where $v_i(\mathbf{x}) \triangleq \nabla_{x_i} u_i(\mathbf{x})$.*

Definition 2.3 $\mathbf{x}^* \in \mathcal{X}$ *is called a (pure-strategy) Nash equilibrium of a game $\mathcal{G}$ if for each player $i \in \mathcal{N}$, it holds true that $u_i(x_i^*, \mathbf{x}_{-i}^*) \geq u_i(x_i, \mathbf{x}_{-i}^*)$ for each $x_i \in \mathcal{X}_i$.*

As an important special case, we consider *individually (pseudo-)concave* payoff functions that have the property that, for each $i \in \mathcal{N}$, the function $u_i(x_i; \mathbf{x}_{-i})$ is (pseudo-)concave in x_i for all $\mathbf{x}_{-i} \in \prod_{j \neq i} \mathcal{X}_j$. For this case, the game itself is called (pseudo-)concave and $\mathbf{x}^*$ is the Nash equilibrium if and only if it satisfies the first-order condition: $(\mathbf{x} - \mathbf{x}^*)^\top \mathbf{v}(\mathbf{x}^*) \leq 0$ for all $\mathbf{x} \in \mathcal{X}$.

Proposition 2.1 *If $\mathbf{x}^* \in \mathcal{X}$ is a Nash equilibrium, then $(\mathbf{x} - \mathbf{x}^*)^\top \mathbf{v}(\mathbf{x}^*) \leq 0$ for all $\mathbf{x} \in \mathcal{X}$. The converse also holds true if the game is (pseudo-)concave.*

Proposition 2.1 shows that Nash equilibrium of concave games are precisely the solutions of the variational inequality $(\mathbf{x} - \mathbf{x}^*)^\top \mathbf{v}(\mathbf{x}^*) \leq 0$ for all $\mathbf{x} \in \mathcal{X}$. Using this variational characterization, the equilibrium existence follows from standard results while Rosen (1965) provided the following sufficient condition for equilibrium uniqueness:

Theorem 2.2 *Assume that $\mathcal{G}$ satisfies the payoff monotonicity condition: $(\mathbf{x}' - \mathbf{x})^\top (\mathbf{v}(\mathbf{x}') - \mathbf{v}(\mathbf{x})) \leq 0$ for all $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$, with equality if and only if $\mathbf{x} = \mathbf{x}'$. Then, $\mathcal{G}$ admits a unique Nash equilibrium.*

Games satisfying the payoff monotonicity condition are called (strictly) monotone and they enjoy properties similar to those of (strictly) concave functions. In particular, when $\mathbf{v} = \nabla f$ for some sufficiently smooth function f, this condition is equivalent to f being (strictly) concave. The notion *strictly* refers to the *only if* requirement in the payoff monotonicity condition. Analogously, we can define the notion of strongly monotone games:

Definition 2.4 *A continuous game $\mathcal{G}$ is called λ-strongly monotone if the payoff strongly monotone condition holds: $(\mathbf{x}' - \mathbf{x})^\top (\mathbf{v}(\mathbf{x}') - \mathbf{v}(\mathbf{x})) \leq -\lambda \|\mathbf{x}' - \mathbf{x}\|^2$ for all $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$.*

For strongly monotone games, the last iterate $\mathbf{x}_t$ induced by the OGD learning provably converges to a Nash equilibrium asymptotically regardless of imperfect feedback. The appealing feature is that the finite-time rate can be derived in terms of $\|\mathbf{x}_t - \mathbf{x}^*\|$ where $\mathbf{x}^*$ is an unique Nash equilibrium. On the other hand, $\mathbf{x}_t$ possibly converges to a limit cycle or repeatedly hits the boundary in the setting of general monotone games (Mertikopoulos et al., 2018a; Daskalakis et al., 2018), while the average iterate, $\sum_{j=1}^t \mathbf{x}_j / t$, provably converges with the rate expressed in terms of the equilibrium gap function.

Recently, Mertikopoulos & Zhou (2019) have analyzed online mirror descent (OMD) learning for so-called variational stable games and proved last-iterate convergence to Nash equilibria regardless of imperfect feedback. This result is quite surprising since the notion of variational stability is even weaker than that of strict monotonicity, meaning that strong monotonicity is not necessary for the last-iterate convergence of the OGD-based learning. However, the rate is unknown even in terms of the equilibrium gap function.

2.2. λ-cocercive Games

We consider the following notion of cocercivity which is weaker than the notion of strong monotonicity.

Definition 2.5 *A continuous game $\mathcal{G}$ is called λ-cocoercive if the payoff cocoercive condition holds: $(\mathbf{x}' - \mathbf{x})^\top (\mathbf{v}(\mathbf{x}') - \mathbf{v}(\mathbf{x})) \leq -\lambda \|\mathbf{v}(\mathbf{x}') - \mathbf{v}(\mathbf{x})\|^2$ for all $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$.*

Two comments are useful in understanding λ-cocercive games. First, the Nash equilibrium is not necessarily unique but shares the same individual payoff gradient; i.e., $\mathbf{v}(\mathbf{x}) = \mathbf{v}(\mathbf{x}')$ for any $\mathbf{x}, \mathbf{x}' \in \mathcal{X}^*$. Second, λ-cocercivity is weaker than the notion of strictly monotonicity. Indeed, for a λ-cocercive game, $(\mathbf{x}' - \mathbf{x})^\top (\mathbf{v}(\mathbf{x}') - \mathbf{v}(\mathbf{x})) = 0$ only implies that $\mathbf{v}(\mathbf{x}') = \mathbf{v}(\mathbf{x})$ and $\mathbf{x}' = \mathbf{x}$ does not necessarily hold true.

In this paper, we focus on the unconstrained setting in which $\mathcal{X}_i = \mathbb{R}^{n_i}$. All the Nash equilibria $\mathbf{x}^* \in \mathcal{X}^*$ satisfy $\mathbf{v}(\mathbf{x}^*) = 0$ in this setting so the notion of *optimality gap function* for a candidate solution $\mathbf{x} \in \mathcal{X}$ is natural: $\epsilon(\mathbf{x}) = \|\mathbf{v}(\mathbf{x})\|^2$. However, it is more difficult to analyze the convergence property of online algorithms, especially when the feedback information is noisy, since the iterates are not necessarily assumed to be bounded.

2.3. Learning via Online Gradient Descent

We describe the *online gradient descent* (OGD) algorithm in our game-theoretic setting. Intuitively, the main idea is as follows: At each stage of process, every player $i \in \mathcal{N}$ forms an estimate $\hat{v}_i$ of the individual gradient of their payoff function at the current action profile, possibly subject to noise and uncertainty. Subsequently, they choose an action x_i for the next stage using the current action and feedback $\hat{v}_i$, and continue playing.

Formally, starting with some arbitrarily (and possibly uninformed) iterate $\mathbf{x}_0 \in \mathbb{R}^n$ at $t = 0$, the scheme can be described via the recursion

$$x_{i,t+1} \; = \; x_{i,t} + \eta_{t+1} \hat{v}_{i,t+1}, \tag{1}$$

where $t \geq 0$ denotes the stage of process, and $\hat{v}_{i,t}$ is an estimate of the individual payoff gradient $v_i(\mathbf{x}_t)$ of player i at stage t. The learning rate $\eta_t > 0$ is a nonincreasing sequence which can be of the form c/t^p for some $p \in [0, 1]$.

Feedback and uncertainty: we assume that each player $i \in \mathcal{N}$ has access to a "black box" feedback mechanism—an *oracle*—which returns an estimate of their payoff gradients at their current action profile. This information can be imperfect for a multitude of reasons; see Section 3.1 of Mertikopoulos & Zhou (2019). We thus consider the following noisy feedback model:

$$\hat{v}_{i,t+1} \; = \; v_i(\mathbf{x}_t) + \xi_{i,t+1}, \tag{2}$$

where the noise process $\xi_t = (\xi_{i,t})_{i \in \mathcal{N}}$ is an L^2-bounded martingale difference adapted to the history $(\mathcal{F}_t)_{t \geq 1}$ of $\mathbf{x}_t$ (i.e., ξ_t is $\mathcal{F}_t$-measurable but ξ_{t+1} isn't).

We focus on two types of random noise proposed by Polyak (1987). The first type is called *relative random noise*:

$$\mathbb{E}[\xi_{t+1} \mid \mathcal{F}_t] = 0, \quad \mathbb{E}[\|\xi_{t+1}\|^2 \mid \mathcal{F}_t] \leq \tau_t \|\mathbf{v}(\mathbf{x}_t)\|^2. \quad (3)$$

and the second type is called *absolute random noise*:

$$\mathbb{E}[\xi_{t+1} \mid \mathcal{F}_t] = 0, \quad \mathbb{E}[\|\xi_{t+1}\|^2 \mid \mathcal{F}_t] \leq \sigma_t^2, \quad (4)$$

These conditions are mild (note that an i.i.d. condition is not imposed) and allows for a broad range of error processes. For the relative random noise, the variance decreases as it approaches a Nash equilibrium which admits better convergence rates for learning algorithms.

3. Convergence under Perfect Feedback

In this section, we analyze the convergence property of OGD learning under perfect feedback. In particular, we show that the finite-time last-iterate convergence rate is $o(1/T)$ regardless of fully adaptive learning rates. To our knowledge, the proof techniques for analyzing adaptive OGD learning is new and may be of independent interest.

3.1. OGD Learning

We first provide a lemma which shows that $\|\mathbf{v}(\mathbf{x}_t)\|^2$ is nonnegative, nonincreasing and summable.

Lemma 3.1 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and with an nonempty set of Nash equilibria $\mathcal{X}^*$. Under the condition that $\eta_t = \eta \in (0, \lambda]$, the OGD iterate $\mathbf{x}_t$ satisfies that $\|\mathbf{v}(\mathbf{x}_{t+1})\| \leq \|\mathbf{v}(\mathbf{x}_t)\|$ for all $t \geq 0$, and*

$$\begin{aligned} \|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_0)\| &\leq \|\mathbf{x}_0 - \Pi_{\mathcal{X}^*}(\mathbf{x}_0)\|, \\ \sum_{t=0}^{+\infty} \|\mathbf{v}(\mathbf{x}_t)\|^2 &\leq \frac{\|\mathbf{x}_0 - \Pi_{\mathcal{X}^*}(\mathbf{x}_0)\|^2}{\eta\lambda}. \end{aligned}$$

Remark 3.2 *For OGD learning with perfect feedback and constant step size, the update formula in Eq. (1) implies that $\|\mathbf{v}(\mathbf{x}_t)\|^2 = \|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2/\eta^2$. This implies that $\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2$ also serves as the candidate for an optimality gap function. Such a quantity is called the iterative gap and is frequently used to construct stopping criteria in practice.*

We are now ready to present our main results on the last-iterate convergence rate of OGD learning.

Theorem 3.3 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and with an nonempty set of Nash equilibria $\mathcal{X}^*$. Under the condition that $\eta_t = \eta \in (0, \lambda]$, the OGD iterate $\mathbf{x}_t$ satisfies $\epsilon(\mathbf{x}_T) = o(1/T)$.*

Algorithm 1 Adaptive Online Gradient Descent

```
Initialization: x_0 ∈ R^n, η_1 = 1/β_1 for some β_1 > 0 and
  tuning parameter r > 1.
for t = 0, 1, 2, ... do
  for i = 1, 2, ..., N do
    x_i^{t+1} ← x_i + η_{t+1} v(x_t).
    if ||v(x_{t+1})|| > ||v(x_t)|| then
      β_{t+2} ← r β_{t+1}.
    else
      β_{t+2} ← β_{t+1}.
    end if
    η_{t+2} ← 1/sqrt(β_{t+2} + Σ_{j=0}^{t} ||v(x_j)||^2).
  end for
end for
```

Proof. Lemma 3.1 implies that $\{\|\mathbf{v}(\mathbf{x}_t)\|^2\|\}_{t \geq 0}$ is nonnegative, nonincreasing and $\sum_{t=0}^{+\infty} \|\mathbf{v}(\mathbf{x}_t)\|^2 < +\infty$. Therefore,

$$T\|\mathbf{v}(\mathbf{x}_{2T-1})\|^2 \leq \sum_{t=T}^{2T-1} \|\mathbf{v}(\mathbf{x}_t)\|^2 \ \rightarrow \ 0 \text{ as } T \rightarrow +\infty.$$

which implies that $\|\mathbf{v}(\mathbf{x}_T)\|^2 = o(1/T)$. By the definition of $\epsilon(\mathbf{x})$, we conclude the desired result. □

3.2. Adaptive OGD Learning

Our main result in this subsection is the last-iterate convergence rate of Algorithm 1. Here the algorithm requires no prior knowledge of λ but still achieves the rate of $o(1/T)$.

Theorem 3.4 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and with an nonempty set of Nash equilibria $\mathcal{X}^*$. The adaptive OGD iterate $\mathbf{x}_t$ satisfies that $\epsilon(\mathbf{x}_T) = o(1/T)$.*

Proof. Since the step-size sequence $\{\eta_t\}_{t \geq 1}$ is nonincreasing, we define the first iconic time in our analysis as follows,

$$t^* \ = \ \max\{t \geq 0 \mid \eta_{t+1} > \lambda\}.$$

In what follows, we establish a last-iterate convergence rate for two cases: $t^* = +\infty$ (**Case I**) and $t^* < +\infty$ (**Case II**).

Case I. First, we have $1/\lambda^2 - \beta_0 \geq 0$ since $\eta_0 > \lambda$. Noting that $\beta_{t+2} \leftarrow r\beta_{t+1}$ with $r > 1$ is updated when $\|\mathbf{v}(\mathbf{x}_{t+1})\| > \|\mathbf{v}(\mathbf{x}_t)\|$, there exists $T_0 > 0$ such that $\|\mathbf{v}(\mathbf{x}_{t+1})\| \leq \|\mathbf{v}(\mathbf{x}_t)\|$ for all $t \geq T_0$. If not, then $\beta_t \rightarrow +\infty$ as $t \rightarrow +\infty$ and $\eta_t \rightarrow 0$. However, $t^* = +\infty$ implies that $\eta_{t+1} \geq \lambda$ for all $t \geq 1$. This leads to a contradiction. Furthermore, it holds true for all $t \geq 0$ that

$$\sum_{j=0}^{t} \|\mathbf{v}(\mathbf{x}_j)\|^2 \ \leq \ \frac{1}{\lambda^2} - \beta_{t+2} \ \leq \ \frac{1}{\lambda^2} - \beta_0 \ < \ +\infty.$$

By starting the sequence at a later index T_0, we have $\sum_{t \geq T_0} \|\mathbf{v}(\mathbf{x}_t)\|^2 < +\infty$ and $\|\mathbf{v}(\mathbf{x}_{t+1})\| \leq \|\mathbf{v}(\mathbf{x}_t)\|$ for all $t \geq T_0$. Using the same argument as in Theorem 3.3, the adaptive OGD iterate $\mathbf{x}_t$ satisfies $\epsilon(\mathbf{x}_T) = o(1/T)$.

Case II. First, we claim that $\|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\| \leq D$ where $D = \max_{1 \leq t \leq t^*} \|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\|$. Indeed, it suffices to show that $\|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\| \leq \|\mathbf{x}_{t^*} - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\|$ holds for $t > t^*$. By the definition of t^*, we have $\eta_{t+1} \leq \lambda$ for all $t > t^*$. The desired inequality follows from Lemma 3.1.

Using the update formula in Eq. (1), we have

$$\begin{aligned}(\mathbf{x}_{t+1} - \mathbf{x}^*)^\top \mathbf{v}(\mathbf{x}_t) &= \frac{1}{2\eta_{t+1}} \left(\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2 \right.\\ &\quad \left. + \|\mathbf{x}^* - \mathbf{x}_{t+1}\|^2 - \|\mathbf{x}^* - \mathbf{x}_t\|^2\right).\end{aligned}$$

Using the update formula for OGD learning, we have

$$\begin{aligned}\lambda\|\mathbf{v}(\mathbf{x}_t)\|^2 &\leq \frac{1}{\eta_{t+1}} \left(\|\mathbf{x}^* - \mathbf{x}_t\|^2 - \|\mathbf{x}^* - \mathbf{x}_{t+1}\|^2\right) \\ &\quad + \left(\frac{1}{\lambda} - \frac{1}{\eta_{t+1}}\right) \|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2. \qquad (5)\end{aligned}$$

Recalling that the step-size sequence $\{\eta_t\}_{t\geq 1}$ is nonincreasing and $\|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\| \leq D$, we let $\mathbf{x}^* = \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})$ in Eq. (5) and obtain that

$$\sum_{t=0}^T \lambda\|\mathbf{v}(\mathbf{x}_t)\|^2 \leq \frac{D^2}{\eta_{T+1}} + \sum_{t=0}^T \left(\frac{1}{\lambda} - \frac{1}{\eta_{t+1}}\right)\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2.$$

To proceed, we define the second iconic time as

$$t_1^* = \max\left\{t \geq 0 \mid \eta_{t+1} > \frac{\lambda}{2D^2+1}\right\} > t^*.$$

If $t_1^* = +\infty$, it is straightforward to show that the adaptive OGD iterate $\mathbf{x}_t$ satisfies $\epsilon(\mathbf{x}_T) = o(1/T)$ using the same argument as in **Case I**.

Next, we consider $t_1^* < +\infty$. We recall that $\eta_{t+1} \leq \lambda$ for all $t > t_1^*$ which implies that $1/\lambda - 1/\eta_{t+1} \leq 0$. Since $\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2 = \eta_{t+1}^2\|\mathbf{v}(\mathbf{x}_t)\|^2$ (cf. Eq. (1)) and assuming that T is sufficiently large without loss of generality, we have

$$\sum_{t=0}^T \lambda\|\mathbf{v}(\mathbf{x}_t)\|^2 \leq \frac{D^2}{\eta_{T+1}} + \sum_{t=0}^{t_1^*} \frac{\eta_{t+1}^2\|\mathbf{v}(\mathbf{x}_t)\|^2}{\lambda} = \mathrm{I} + \mathrm{II}.$$

Before bounding terms I and II, we present two technical lemmas which are crucial for our subsequent analysis (see Lemma A.1 and A.2 in Bach & Levy, 2019, for a detailed proof).

Lemma 3.5 *For a sequence of numbers $a_0, a_1, \ldots, a_n \in [0, a]$ and $b \geq 0$, the following inequality holds:*

$$\begin{aligned}\sqrt{b + \sum_{i=0}^{n-1} a_i} - \sqrt{b} &\leq \sum_{i=0}^n \frac{a_i}{\sqrt{b + \sum_{j=0}^{i-1} a_j}} \\ &\leq \frac{2a}{\sqrt{b}} + 3\sqrt{a} + 3\sqrt{b + \sum_{i=0}^{n-1} a_i}.\end{aligned}$$

Lemma 3.6 *For a sequence of numbers $a_0, a_1, \ldots, a_n \in [0, a]$ and $b \geq 0$, the following inequality holds:*

$$\sum_{i=0}^n \frac{a_i}{b + \sum_{j=0}^{i-1} a_j} \leq 2 + \frac{4a}{b} + 2\log\left(1 + \sum_{i=0}^{n-1} \frac{a_i}{b}\right).$$

Bounding term I: By the definition of t_1^* and Lemma 3.1, we have $\beta_t = \beta_{t_1^*+1}$ for all $t > t_1^*$. Thus, we derive from the definition of η_t that

$$\mathrm{I} \leq D^2\sqrt{\beta_{T+1} + \sum_{j=0}^{T-1} \|\mathbf{v}(\mathbf{x}_j)\|^2} \leq D^2\sqrt{\beta_{t_1^*+1} + \sum_{j=0}^{T-1} \|\mathbf{v}(\mathbf{x}_j)\|^2},$$

since $\|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\| \leq D$ for all $t \geq 0$. Since the notion of λ-cocercivity implies the notion of $(1/\lambda)$-Lipschiz continuity, we have

$$\|\mathbf{v}(\mathbf{x}_t)\|^2 \leq \frac{\|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\|^2}{\lambda^2} \leq \frac{D^2}{\lambda^2}. \qquad (6)$$

Using the first inequality in Lemma 3.5, we have

$$\begin{aligned}\mathrm{I} &\leq D^2\sqrt{\beta_{t_1^*+1}} + \sum_{t=0}^T \frac{D^2\|\mathbf{v}(\mathbf{x}_t)\|^2}{\sqrt{\beta_{t_1^*+1} + \sum_{j=0}^{t-1}\|\mathbf{v}(\mathbf{x}_j)\|^2}} \\ &\leq D^2\sqrt{\beta_{t_1^*+1}} + \sum_{t=0}^{t_1^*} \frac{D^2\|\mathbf{v}(\mathbf{x}_t)\|^2}{\sqrt{\beta_{t_1^*+1} + \sum_{j=0}^{t-1}\|\mathbf{v}(\mathbf{x}_j)\|^2}} \\ &\quad + \sum_{t=t_1^*+1}^T D^2\eta_{t+1}\|\mathbf{v}(\mathbf{x}_t)\|^2. \qquad (7)\end{aligned}$$

Since $\eta_{t+1} \leq \lambda/2D^2$ for all $t > t_1^*$, we have

$$\sum_{t=t_1^*+1}^T D^2\eta_{t+1}\|\mathbf{v}(\mathbf{x}_t)\|^2 \leq \sum_{t=t_1^*+1}^T \frac{\lambda\|\mathbf{v}(\mathbf{x}_t)\|^2}{2}. \qquad (8)$$

Using the second inequality in Lemma 3.5,

$$\begin{aligned}&\sum_{t=0}^{t_1^*} \frac{D^2\|\mathbf{v}(\mathbf{x}_t)\|^2}{\sqrt{\beta_{t_1^*+1} + \sum_{j=0}^{t-1}\|\mathbf{v}(\mathbf{x}_j)\|^2}} \qquad (9)\\ &\leq \frac{2D^2}{\lambda^2\sqrt{\beta_{t_1^*+1}}} + \frac{3D}{\lambda} + 3\sqrt{\beta_{t_1^*+1} + \sum_{j=0}^{t_1^*-1}\|\mathbf{v}(\mathbf{x}_j)\|^2}.\end{aligned}$$

By the definition of η_t, we have

$$\sqrt{\beta_{t_1^*+1} + \sum_{j=0}^{t_1^*-1}\|\mathbf{v}(\mathbf{x}_j)\|^2} = \frac{1}{\eta_{t_1^*+1}} < \frac{2D^2+1}{\lambda}. \quad (10)$$

Putting Eq. (7)-(10) together yields that

$$\mathrm{I} \leq D^2\sqrt{\beta_{t_1^*+1}} + \frac{2D^2}{\lambda^2\sqrt{\beta_{t_1^*+1}}} + \frac{3 + 3D + 6D^2}{\lambda} + \sum_{t=t_1^*+1}^T \frac{\lambda\|\mathbf{v}(\mathbf{x}_t)\|^2}{2}.$$

Bounding term II: By the definition of η_t and noting that $\beta_t \geq \beta_1$ for all $t \geq 1$, we have

$$\text{II} \leq \frac{1}{\lambda}\left(\sum_{t=0}^{t_1^*} \frac{\|\mathbf{v}(\mathbf{x}_t)\|^2}{\beta_1 + \sum_{j=0}^{t-1}\|\mathbf{v}(\mathbf{x}_j)\|^2}\right).$$

Recalling Eq. (6), we can apply Lemma 3.6 with Eq. (10) to obtain that

$$\begin{aligned}\text{II} &\leq \frac{1}{\lambda}\left(2 + \frac{4D^2}{\lambda^2\beta_1} + 2\log\left(1 + \frac{1}{\beta_1}\sum_{j=0}^{t_1^*-1}\|\mathbf{v}(\mathbf{x}_j)\|^2\right)\right) \\ &\leq \frac{1}{\lambda}\left(2 + \frac{4D^2}{\lambda^2\beta_1} + 2\log\left(1 + \frac{8D^4+2}{\lambda^2\beta_1}\right)\right).\end{aligned}$$

Therefore, we conclude that

$$\begin{aligned}\sum_{t=0}^{T}\frac{\lambda\|\mathbf{v}(\mathbf{x}_t)\|^2}{2} &\leq D^2\sqrt{\beta_{t_1^*+1}} + \frac{2D^2}{\lambda^2\sqrt{\beta_{t_1^*+1}}} \\ &+ \frac{3+3D+6D^2}{\lambda} + \frac{1}{\lambda}\left(2 + \frac{4D^2}{\lambda^2\beta_1} + 2\log\left(1 + \frac{8D^4+2}{\lambda^2\beta_1}\right)\right),\end{aligned}$$

which implies that $\sum_{t=0}^{T}\|\mathbf{v}(\mathbf{x}_t)\|^2$ is bounded by a constant for all $T \geq 0$. By starting the sequence at a later index t_1^*, we have $\|\mathbf{v}(\mathbf{x}_{t+1})\| \leq \|\mathbf{v}(\mathbf{x}_t)\|$ for all $t \geq t_1^*$ and $\sum_{t \geq t_1^*}\|\mathbf{v}(\mathbf{x}_t)\|^2 < +\infty$. Using the same argument as in Theorem 3.3, we conclude that the adaptive OGD iterate $\mathbf{x}_t$ satisfies $\epsilon(\mathbf{x}_T) = o(1/T)$. □

4. Convergence under Imperfect Feedback with Relative Random Noise

In this section, we analyze the convergence of OGD learning under imperfect feedback with relative random noise (3). In particular, we show that the almost sure last-iterate convergence is guaranteed and the finite-time average-iterate convergence rate is $O(1/T)$ when $0 < \tau_t \leq \tau < +\infty$. More importantly, we get a finite-time last-iterate convergence rate when τ_t satisfies certain a summability condition (12).

4.1. Almost Sure Last-Iterate Convergence

In this subsection, we establish the almost sure last-iterate convergence under imperfect feedback with relative random noise. The appealing feature here is that the convergence results provably hold with a constant step size. The first and second lemmas provide two different key inequalities for $\mathbf{x}_t$ and $\mathbb{E}[\epsilon(\mathbf{x}_t)]$ respectively.

Lemma 4.1 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and let the set of Nash equilibria, $\mathcal{X}^*$, be nonempty. Under the noisy model (2), the noisy OGD iterate $\mathbf{x}_t$ satisfies the following inequality, for any Nash equilibrium $\mathbf{x}^* \in \mathcal{X}^*$:*

$$\begin{aligned}\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 &\leq \|\mathbf{x}_t - \mathbf{x}^*\|^2 + 2\eta_{t+1}^2\|\xi_{t+1}\|^2 \\ &-(2\lambda\eta_{t+1} - 2\eta_{t+1}^2)\|\mathbf{v}(\mathbf{x}_t)\|^2 + 2\eta_{t+1}(\mathbf{x}_t - \mathbf{x}^*)^\top\xi_{t+1}.\end{aligned} \quad (11)$$

Lemma 4.2 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and let the set of Nash equilibria, $\mathcal{X}^*$, be nonempty. Under the noisy model (2) with relative random noise (3) and the step-size sequence $\eta_t \in (0, \lambda)$, the noisy OGD iterate $\mathbf{x}_t$ satisfies $\mathbb{E}[\epsilon(\mathbf{x}_{t+1})] \leq \mathbb{E}[\epsilon(\mathbf{x}_t)] + \tau_t\|\mathbf{v}(\mathbf{x}_t)\|^2/\lambda\eta_{t+1}$.*

We are now ready to characterize the almost sure last-iterate convergence. Note that the condition imposed on τ_t is weak and $\eta_t = \eta \in [\underline{\eta}, \overline{\eta}]$ is allowed for all $t \geq 1$.

Theorem 4.3 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and let the set of Nash equilibria, $\mathcal{X}^*$, be nonempty. Under the noisy model (2) with relative random noise (3) satisfying $\tau_t \in (0, \tau]$ for some $\tau < +\infty$ and the step-size sequence satisfying $0 < \underline{\eta} \leq \eta_t \leq \overline{\eta} < \lambda/(1+\tau)$ for all $t \geq 1$, the noisy OGD iterate $\mathbf{x}_t$ converges to $\mathcal{X}^*$ almost surely.*

Proof. We obtain the following inenquality by taking an expectation of both sides of Eq. (11) conditioned on $\mathcal{F}_t$ (cf. Lemma 4.1):

$$\begin{aligned}\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \mid \mathcal{F}_t] &\leq \|\mathbf{x}_t - \mathbf{x}^*\|^2 - (2\lambda\eta_{t+1} - 2\eta_{t+1}^2) \\ &\|\mathbf{v}(\mathbf{x}_t)\|^2 + 2\eta_{t+1}^2\mathbb{E}[\|\xi_{t+1}\|^2 \mid \mathcal{F}_t] + 2\eta_{t+1}\mathbb{E}[(\mathbf{x}_t - \mathbf{x}^*)^\top\xi_{t+1} \mid \mathcal{F}_t].\end{aligned}$$

Since the noisy model (2) is with relative random noise (3) satisfying $\tau_t \in (0, \tau)$ for some $\tau < +\infty$, we have $\mathbb{E}[(\mathbf{x}_t - \mathbf{x}^*)^\top\xi_{t+1} \mid \mathcal{F}_t] = 0$ and $\mathbb{E}[\|\xi_{t+1}\|^2 \mid \mathcal{F}_t] \leq \tau\|\mathbf{v}(\mathbf{x}_t)\|^2$. Therefore, we have

$$\begin{aligned}\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \mid \mathcal{F}_t] &\leq \|\mathbf{x}_t - \mathbf{x}^*\|^2 \\ &-2(\lambda - \overline{\eta} - \tau\overline{\eta})\eta_{t+1}\|\mathbf{v}(\mathbf{x}_t)\|^2.\end{aligned}$$

Since $\eta_t > 0$ and $\overline{\eta} < \lambda/(1+\tau)$, we let $M_t = \|\mathbf{x}_t - \mathbf{x}^*\|^2$ and obtain that M_t is an nonnegative supermartingale. Then Doob's martingale convergence theorem shows that M_n converges to an nonnegative and integrable random variable almost surely. Letting $M_\infty = \lim_{t \to +\infty} M_t$, it suffices to show that $M_\infty = 0$ almost surely. We assume to the contrary that there exists $m > 0$ such that $M_\infty > m$ with positive probability. Then $M_t > m/2$ for sufficiently large t with positive probability. Formally, there exists $\delta > 0$ such that

$$\text{Prob}(M_t > m/2 \text{ for sufficiently large } t) \geq \delta.$$

By the definition of M_t and recalling that $\mathbf{x}^* \in \mathcal{X}^*$ can be any Nash equilibrium, we let U be a $(m/2)$-neighborhood of $\mathcal{X}^*$ and obtain that $\mathbf{x}_t \notin U$ for sufficiently large t with positive probability. Since $\|\mathbf{v}(\mathbf{x})\| = 0$ if and only if $\mathbf{x} \in$

$\mathcal{X}^*$, there exists $c > 0$ such that $\|\mathbf{v}(\mathbf{x})\| \geq c$ for sufficiently large t with positive probability. Therefore, we conclude that $\mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2] \nrightarrow 0$ as $t \to +\infty$.

On the other hand, by taking the expectation of Eq. (19) and using the condition $\eta_t \geq \underline{\eta} > 0$ for all $t \geq 1$, we have

$$\mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2] \;\leq\; \frac{\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2] - \mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2]}{2(\lambda - \overline{\eta} - \tau\overline{\eta})\underline{\eta}}.$$

This implies that $\sum_{t=0}^{\infty} \mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2] < +\infty$ and hence $\mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2] \to 0$ as $t \to +\infty$ which contradicts the previous argument. This completes the proof. □

4.2. Finite-Time Convergence Rate: Time-Average and Last-Iterate

In this subsection, we focus on deriving two types of rates—time-average rates and last-iterate convergence rates—as formalized by the following theorems.

Theorem 4.4 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and let the set of Nash equilibria, $\mathcal{X}^*$, be nonempty. Under the noisy model (2) with relative random noise (3) satisfying $\tau_t \in (0, \tau]$ for some $\tau < +\infty$ and a step-size sequence satisfying $0 < \underline{\eta} \leq \eta_t \leq \overline{\eta} < \lambda/(1+\tau)$ for all $t \geq 1$, the noisy iterate $\mathbf{x}_t$ satisfies $\frac{1}{T+1}(\mathbb{E}[\sum_{t=0}^T \epsilon(\mathbf{x}_t)]) = O(1/T)$.*

Inspired by Lemma 4.2, we impose an intuitive condition on the variance ratio of the noise process $\{\tau_t\}_{t \geq 0}$. More specifically, $\{\tau_t\}_{t \geq 0}$ is nonincreasing and there exists a function $a : \mathbb{R}^+ \to \mathbb{R}^+$ satisfying $a(t) = o(1)$ such that

$$\frac{1}{T+1}\left(\sum_{t=0}^{T-1} \tau_t\right) \;=\; O(a(T)). \tag{12}$$

Remark 4.5 *The condition (12) is fairly mild. Indeed, the decaying rate $a(t)$ can be very slow while still guaranteeing a finite-time last-iterate convergence rate. For some typical examples, we have $a(t) = \log\log(t)/t$ if $\tau_t = 1/t\log(t)$ and $a(t) = \log(t)/t$ if $\tau_t = 1/t$. When $\tau_t = \Omega(1/t)$, we have $a(t) = \tau_t$, such as $a(t) = 1/\sqrt{t}$ if $\tau_t = 1/\sqrt{t}$ and $a(t) = 1/\log\log(t)$ if $\tau_t = 1/\log\log(t)$. Under this condition, we can derive a last-iterate convergence rate given the decaying rate of τ_t as $t \to +\infty$.*

Under the condition (12), a finite-time last-iterate convergence rate can be derived under certain step-size sequences.

Theorem 4.6 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and let the set of Nash equilibria, $\mathcal{X}^*$, be nonempty. Under the noisy model (2) with relative random noise (3) satisfying Eq. (12) and the step-size sequence satisfying $0 < \underline{\eta} \leq \eta_t \leq \overline{\eta} < \lambda/(1+\tau)$ for all $t \geq 1$, the noisy OGD iterate $\mathbf{x}_t$ satisfies*

$$\mathbb{E}[\epsilon(\mathbf{x}_T)] \;=\; \begin{cases} O(a(T)) & \text{if } a(T) = \Omega(1/T), \\ O(1/T) & \text{otherwise.} \end{cases}$$

Proof. Using Lemma 4.2 and $\eta_t \geq \underline{\eta} > 0$, we have

$$\mathbb{E}[\epsilon(\mathbf{x}_T)] \;\leq\; \mathbb{E}[\epsilon(\mathbf{x}_t)] + \frac{\sum_{j=t}^{T-1} \tau_j \|\mathbf{v}(\mathbf{x}_j)\|^2}{\lambda\underline{\eta}}.$$

Summing up this inequality over $t = 0, \ldots, T$ yields

$$(T+1)\mathbb{E}[\epsilon(\mathbf{x}_T)] \leq \sum_{t=0}^{T} \mathbb{E}[\epsilon(\mathbf{x}_t)] + \frac{\sum_{t=0}^{T-1}\sum_{j=t}^{T-1} \tau_j \|\mathbf{v}(\mathbf{x}_j)\|^2}{\lambda\underline{\eta}}.$$

Using Eq. (19) and $\eta_t \geq \underline{\eta} > 0$ for all $t \geq 1$, we have

$$\mathbb{E}[\|\mathbf{v}(\mathbf{x}_j)\|^2] \leq \frac{\mathbb{E}[\|\mathbf{x}_j - \mathbf{x}^*\|^2] - \mathbb{E}[\|\mathbf{x}_{j+1} - \mathbf{x}^*\|^2]}{2(\lambda - \overline{\eta} - \tau\overline{\eta})\underline{\eta}}.$$

Putting these pieces with the fact that $\{\tau_t\}_{t \geq 0}$ is an nonincreasing sequence yields that

$$\begin{aligned} \sum_{t=0}^{T-1}\sum_{j=t}^{T-1} \tau_j \|\mathbf{v}(\mathbf{x}_j)\|^2 \;&\leq\; \left(\sum_{t=0}^{T-1} \tau_t\right)\left(\sum_{t=0}^{T-1} \|\mathbf{v}(\mathbf{x}_t)\|^2\right) \\ &\leq\; \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - \overline{\eta} - \tau\overline{\eta})\underline{\eta}}\left(\sum_{t=0}^{T-1} \tau_t\right). \end{aligned}$$

Together with the fact that $\mathbb{E}[\sum_{t=0}^T \epsilon(\mathbf{x}_t)] = O(1)$ (cf. Theorem 4.4), we have

$$\mathbb{E}[\epsilon(\mathbf{x}_T)] \;\leq\; \frac{\sum_{t=0}^T \mathbb{E}[\epsilon(\mathbf{x}_t)]}{T+1} + \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{T+1}\left(\sum_{t=0}^{T-1} \tau_t\right).$$

This completes the proof. □

Algorithm 2 Adaptive Online Gradient Descent with Noisy Feedback Information

Initialization: $\mathbf{x}_0 \in \mathbb{R}^n$, $\beta > 0$ and $\eta_1 = 1/\beta$.
for $t = 0, 1, 2, \ldots$ **do**
  for $i = 1, 2, \ldots, N$ **do**
    $x_i^{t+1} = x_i + \eta_{t+1}\mathbf{v}(\mathbf{x}_t)$.
    $\eta_{t+2} = 1/\sqrt{\beta + \log(t+2) + (\sum_{j=0}^t \|\mathbf{x}_j - \mathbf{x}_{j+1}\|^2/\eta_{j+1}^2)}$.
  end for
end for

4.3. Adaptive OGD Learning

We study the convergence of Algorithm 2 under the noise model (2) with relative random noise (3) satisfying that there exists $a(t) = o(1)$ such that

$$\frac{\log(T+1)}{T+1}\left(\sum_{t=0}^{T-1}\tau_t\right) = O(a(T)). \tag{13}$$

Note that Eq. (13) is slightly stronger than Eq. (12).

Theorem 4.7 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and let the set of Nash equilibria, $\mathcal{X}^*$, be nonempty. Under the noise model (2) with relative random noise (3) satisfying Eq. (13), the adaptive noisy OGD iterate $\mathbf{x}_t$ satisfies*

$$\mathbb{E}[\epsilon(\mathbf{x}_T)] = \begin{cases} O(a(T)) & \text{if } a(T) = \Omega(\log(T)/T), \\ O(\log(T)/T) & \text{otherwise.} \end{cases}$$

The proof technique is new and can be interpreted as a novel combination of the techniques in Theorems 3.4 and 4.6. We refer the interested reader to the appendix for the details.

References

Arora, S., Hazan, E., and Kale, S. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

Bach, F. and Levy, K. Y. A universal algorithm for variational inequalities adaptive to smoothness and noise. In *COLT*, pp. 164–194, 2019.

Bloembergen, D., Tuyls, K., Hennes, D., and Kaisers, M. Evolutionary dynamics of multi-agent learning: a survey. *Journal of Artificial Intelligence Research*, 53:659–697, 2015.

Blum, A. On-line algorithms in machine learning. In *Online algorithms*, pp. 306–325. Springer, 1998.

Blum, A. and Mansour, Y. From external to internal regret. *Journal of Machine Learning Research*, 8:1307–1324, 2007.

Cesa-Bianchi, N. and Lugosi, G. *Prediction, Learning, and Games*. Cambridge university press, 2006.

Daskalakis, C., Ilyas, A., Syrgkanis, V., and Zeng, H. Training GANs with optimism. In *ICLR*, 2018.

Facchinei, F. and Pang, J.-S. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer Science & Business Media, 2007.

Hall, P. and Heyde, C. C. *Martingale Limit Theory and Its Application*. Academic Press, 2014.

Hazan, E. *Introduction to Online Convex Optimization*. Foundations and Trends(r) in Optimization Series. Now Publishers, 2016. ISBN 9781680831702. URL https://books.google.com/books?id=IFxLvgAACAAJ.

Hazan, E., Agarwal, A., and Kale, S. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, December 2007.

Kalai, A. and Vempala, S. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

Krichene, S., Krichene, W., Dong, R., and Bayen, A. Convergence of heterogeneous distributed learning in stochastic routing games. In *Communication, Control, and Computing (Allerton), 2015 53rd Annual Allerton Conference on*, pp. 480–487. IEEE, 2015.

Lam, K., Krichene, W., and Bayen, A. On learning how players learn: estimation of learning dynamics in the routing game. In *Cyber-Physical Systems (ICCPS), 2016 ACM/IEEE 7th International Conference on*, pp. 1–10. IEEE, 2016.

Mertikopoulos, P. and Zhou, Z. Learning in games with continuous action sets and unknown payoff functions. *Mathematical Programming*, 173(1-2):465–507, 2019.

Mertikopoulos, P., Belmega, E. V., Negrel, R., and Sanguinetti, L. Distributed stochastic optimization via matrix exponential learning. 65(9):2277–2290, May 2017.

Mertikopoulos, P., Papadimitriou, C., and Piliouras, G. Cycles in adversarial regularized learning. In *SODA*, pp. 2703–2717. SIAM, 2018a.

Mertikopoulos, P., Papadimitriou, C., and Piliouras, G. Cycles in adversarial regularized learning. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 2703–2717. SIAM, 2018b.

Monnot, B. and Piliouras, G. Limits and limitations of no-regret learning in games. *The Knowledge Engineering Review*, 32, 2017.

Palaiopanos, G., Panageas, I., and Piliouras, G. Multiplicative weights update with constant step-size in congestion games: Convergence, limit cycles and chaos. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems 30*, pp. 5872–5882. Curran Associates, Inc., 2017.

Polyak, B. T. *Introduction to Optimization*, volume 1. Optimization Software Inc., Publications Division, New York, 1987.

Quanrud, K. and Khashabi, D. Online learning with adversarial delays. In *Advances in Neural Information Processing Systems*, pp. 1270–1278, 2015.

Rosen, J. B. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society*, pp. 520–534, 1965.

Shalev-Shwartz, S. and Singer, Y. Convex repeated games and Fenchel duality. In *Advances in Neural Information Processing Systems 19*, pp. 1265–1272. MIT Press, 2007.

Shalev-Shwartz, S. et al. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012.

Shoham, Y. and Leyton-Brown, K. *Multiagent Systems: Algorithmic, Game-theoretic, and Logical Foundations*. Cambridge University Press, 2008.

Viossat, Y. and Zapechelnyuk, A. No-regret dynamics and fictitious play. *Journal of Economic Theory*, 148(2):825–842, 2013.

Zhou, Z., Mertikopoulos, P., Bambos, N., Glynn, P. W., and Tomlin, C. Countering feedback delays in multi-agent learning. In *NIPS '17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017.

Zhou, Z., Mertikopoulos, P., Athey, S., Bambos, N., Glynn, P. W., and Ye, Y. Learning in games with lossy feedback. In *Advances in Neural Information Processing Systems*, pp. 5134–5144, 2018.

Zhou, Z., Mertikopoulos, P., Moustakas, A., Bambos, N., and Glynn, P. Robust power management via learning and game design. *Operations Research*, 2020.

Zinkevich, M. Online convex programming and generalized infinitesimal gradient ascent. In *ICML '03: Proceedings of the 20th International Conference on Machine Learning*, pp. 928–936, 2003.

A. Convergence with Absolute Random Noise

In this section, we analyze the convergence of OGD-based learning on λ-cocercive games under imperfect feedback with absolute random noise (4). We first establish that OGD under noisy feedback converges almost surely in last-iterate to the set of Nash equilibria of a co-coercive game if $\sigma_t^2 \in (0, \sigma^2)$ for some $\sigma^2 < +\infty$ and the finite-time $O(1/\sqrt{T})$ convergence rate on $(1/T)\mathbb{E}[\sum_{t=0}^T \epsilon(\mathbf{x}_t)]$ under properly diminishing step-size sequences. We also present a finite-time convergence rate on $\mathbb{E}[\epsilon(\mathbf{x}_T)]$ if σ_t^2 satisfies certain conditions.

A.1. Almost Sure Last-Iterate Convergence

We start by developing a key iterative formula for $\mathbb{E}[\epsilon(\mathbf{x}_t)]$ in the following lemma.

Lemma A.1 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and let the set of Nash equilibria, $\mathcal{X}^*$, be nonempty. Under the noisy model (2) with absolute random noise (4) and letting the OGD-based learning run with a step-size sequence $\eta_t \in (0, \lambda)$, the noisy OGD iterate $\mathbf{x}_t$ satisfies*

$$\mathbb{E}[\epsilon(\mathbf{x}_{t+1})] \ \leq \ \mathbb{E}[\epsilon(\mathbf{x}_t)] + \frac{\sigma_t^2}{\lambda \eta_{t+1}}.$$

We are now ready to establish last-iterate convergence in a strong, almost sure sense. Note that the conditions imposed on σ_t^2 and η_t are minimal.

Theorem A.2 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and let the set of Nash equilibria, $\mathcal{X}^*$, be nonempty. Consider the noisy model (2) with absolute random noise (4) satisfying $\sigma_t^2 \in (0, \sigma^2]$ for some $\sigma^2 < +\infty$ and letting the OGD-based learning run with a step-size sequence satisfying*

$$\sum_{t=1}^{\infty} \eta_t \ = \ +\infty, \qquad \sum_{t=1}^{\infty} \eta_t^2 \ < \ +\infty.$$

Then the noisy OGD iterate $\mathbf{x}_t$ converges to $\mathcal{X}^*$ almost surely.

A.2. Finite-Time Convergence Rate: Time-Average and Last-Iterate

For completeness, we characterize two types of rates: the time-average and last-iterate convergence rate, as formalized by the following theorems.

Theorem A.3 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and let the set of Nash equilibria, $\mathcal{X}^*$, be nonempty. Under the noisy model (2) with absolute random noise (4) satisfying $\sigma_t^2 \in (0, \sigma^2]$ for some $\sigma^2 < +\infty$ and letting the OGD-based learning run with a step-size sequence $\eta_t = c/\sqrt{t}$ for some constant $c \in (0, \lambda)$, the noisy OGD iterate $\mathbf{x}_t$ satisfies*

$$\frac{1}{T+1}\left(\mathbb{E}\left[\sum_{t=0}^T \epsilon(\mathbf{x}_t)\right]\right) \ = \ O\left(\frac{\log(T)}{\sqrt{T}}\right).$$

Inspired by Lemma A.1, we impose an intuitive condition on the variance of noisy process $\{\sigma_t^2\}_{t \geq 0}$. More specifically, there exists a function $\alpha : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ satisfying $a(t) = o(1)$ and $a(t) = \Omega(1/t)$ such that

$$\frac{1}{T+1}\left(\sum_{t=0}^{T-1}(t+1)\sigma_t^2\right) \ = \ O(a(T)). \tag{14}$$

Under this condition, the noisy iterate generated by the OGD-based learning achieves the finite-time last-iterate convergence rate regardless of a sequence of possibly constant step-sizes η_t satisfying $0 < \underline{\eta} \leq \eta_t \leq \overline{\eta} < \lambda$ for all $t \geq 1$.

Theorem A.4 *Fix a λ-cocoercive game $\mathcal{G}$ with a continuous action space, $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathbb{R}^{n_i}, \{u_i\}_{i=1}^N)$, and let the set of Nash equilibria, $\mathcal{X}^*$, be nonempty. Under the noisy model (2) with absolute random noise (4) satisfying Eq. (14) and letting the OGD-based learning run with an nonincreasing step-size sequence satisfying $0 < \underline{\eta} \leq \eta_t \leq \overline{\eta} < \lambda$ for all $t \geq 1$, the noisy OGD iterate $\mathbf{x}_t$ satisfies*

$$\mathbb{E}[\epsilon(\mathbf{x}_T)] \ = \ O(a(T)).$$

B. Missing proofs in Section 3

B.1. Proof of Lemma 3.1

Since $\mathcal{X}_i = \mathbb{R}^{n_i}$, we have

$$\begin{aligned}
&\|x_{i,t+2} - x_{i,t+1}\|^2 \\
&= \|x_{i,t+1} - x_{i,t} + \eta(v_i(\mathbf{x}_{t+1}) - v_i(\mathbf{x}_t))\|^2 \\
&= \|x_{i,t+1} - x_{i,t}\|^2 + 2\eta(x_{i,t+1} - x_{i,t})^\top (v_i(\mathbf{x}_{t+1}) - v_i(\mathbf{x}_t)) + \eta^2\|v_i(\mathbf{x}_{t+1}) - v_i(\mathbf{x}_t)\|^2.
\end{aligned}$$

Expanding the right-hand side of the above inequality and summing up the resulting inequality over $i \in \mathcal{N}$ yields that

$$\begin{aligned}
\|\mathbf{x}_{t+2} - \mathbf{x}_{t+1}\|^2 &= \sum_{i\in\mathcal{N}} \|x_{i,t+2} - x_{i,t+1}\|^2 \\
&\leq \sum_{i\in\mathcal{N}} \left(\|x_{i,t+1} - x_{i,t}\|^2 + \eta^2\|v_i(\mathbf{x}_{t+1}) - v_i(\mathbf{x}_t)\|^2 + 2\eta(x_{i,t+1} - x_{i,t})^\top (v_i(\mathbf{x}_{t+1}) - v_i(\mathbf{x}_t))\right) \\
&= \|\mathbf{x}_{t+1} - \mathbf{x}_t\|^2 + 2\eta(\mathbf{x}_{t+1} - \mathbf{x}_t)^\top (\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t)) + \eta^2\|\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t)\|^2.
\end{aligned} \tag{15}$$

Since $\mathcal{G}$ is a λ-cocoercive game, we have

$$(\mathbf{x}_{t+1} - \mathbf{x}_t)^\top (\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t)) \leq -\lambda\|\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t)\|^2.$$

Plugging the above equation into Eq. (15) together with the condition $\eta \in (0, \lambda]$ yields that

$$\|\mathbf{x}_{t+2} - \mathbf{x}_{t+1}\|^2 \leq \|\mathbf{x}_{t+1} - \mathbf{x}_t\|^2.$$

Using the update formula in Eq. (1), we have $\|\mathbf{v}(\mathbf{x}_{t+1})\| \leq \|\mathbf{v}(\mathbf{x}_t)\|$ for all $t \geq 0$.

Then we proceed to bound $\sum_{t=0}^{+\infty} \|\mathbf{v}(\mathbf{x}_t)\|^2$. Indeed, for any $x_i \in \mathcal{X}_i$, we have

$$(x_i - x_{i,t+1})^\top (x_{i,t+1} - x_{i,t} - \eta v_i(\mathbf{x}_t)) = 0.$$

Applying the equality $a^\top b = (\|a+b\|^2 - \|a\|^2 - \|b\|^2)/2$ yields that

$$(x_{i,t+1} - x_i)^\top v_i(\mathbf{x}_t) = \frac{1}{2\eta}\left(\|x_{i,t} - x_{i,t+1}\|^2 + \|x_i - x_{i,t+1}\|^2 - \|x_i - x_{i,t}\|^2\right).$$

Summing up the resulting inequality over $i \in \mathcal{N}$ yields that

$$(\mathbf{x}_{t+1} - \mathbf{x})^\top \mathbf{v}(\mathbf{x}_t) = \frac{1}{2\eta}\left(\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2 + \|\mathbf{x} - \mathbf{x}_{t+1}\|^2 - \|\mathbf{x} - \mathbf{x}_t\|^2\right), \quad \forall \mathbf{x} \in \mathcal{X}.$$

Letting $\mathbf{x} = \mathbf{x}^* \in \mathcal{X}^*$, we have

$$(\mathbf{x}_{t+1} - \mathbf{x}^*)^\top \mathbf{v}(\mathbf{x}_t) = \frac{1}{2\eta}\left(\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2 + \|\mathbf{x}^* - \mathbf{x}_{t+1}\|^2 - \|\mathbf{x}^* - \mathbf{x}_t\|^2\right). \tag{16}$$

Furthermore, we have

$$(\mathbf{x}_{t+1} - \mathbf{x}^*)^\top \mathbf{v}(\mathbf{x}_t) = (\mathbf{x}_t - \mathbf{x}^*)^\top \mathbf{v}(\mathbf{x}_t) + (\mathbf{x}_{t+1} - \mathbf{x}_t)^\top \mathbf{v}(\mathbf{x}_t).$$

Since $\mathcal{G}$ is a λ-cocoercive game and $\mathbf{v}(\mathbf{x}^*) = 0$, we have

$$\begin{aligned}
(\mathbf{x}_t - \mathbf{x}^*)^\top \mathbf{v}(\mathbf{x}_t) &= (\mathbf{x}_t - \mathbf{x}^*)^\top (\mathbf{v}(\mathbf{x}_t) - \mathbf{v}(\mathbf{x}^*)) \leq -\lambda\|\mathbf{v}(\mathbf{x}_t) - \mathbf{v}(\mathbf{x}^*)\|^2 \\
&= -\lambda\|\mathbf{v}(\mathbf{x}_t)\|^2.
\end{aligned}$$

By Young's inequality we have

$$(\mathbf{x}_{t+1} - \mathbf{x}_t)^\top \mathbf{v}(\mathbf{x}_t) \leq \frac{\lambda\|\mathbf{v}(\mathbf{x}_t)\|^2}{2} + \frac{\|\mathbf{x}_{t+1} - \mathbf{x}_t\|^2}{2\lambda}.$$

Putting these pieces together yields that

$$(\mathbf{x}_{t+1} - \mathbf{x}^*)^\top \mathbf{v}(\mathbf{x}_t) \ \leq \ \frac{\|\mathbf{x}_{t+1} - \mathbf{x}_t\|^2}{2\lambda} - \frac{\lambda \|\mathbf{v}(\mathbf{x}_t)\|^2}{2}. \tag{17}$$

Plugging Eq. (17) into Eq. (16) together with the condition $\eta \in (0, \lambda]$ yields that

$$\lambda \|\mathbf{v}(\mathbf{x}_t)\|^2 \ \leq \ \frac{\|\mathbf{x}^* - \mathbf{x}_t\|^2 - \|\mathbf{x}^* - \mathbf{x}_{t+1}\|^2}{\eta}.$$

Summing up the above inequality over $t = 0, 1, 2, \ldots$ and using the boundedness of $\mathcal{X}$ yields that

$$\sum_{t=0}^{+\infty} \|\mathbf{v}(\mathbf{x}_t)\|^2 \ \leq \ \frac{\|\mathbf{x}^* - \mathbf{x}_0\|^2}{\eta\lambda}.$$

Note that $\mathbf{x}^* \in \mathcal{X}^*$ is chosen arbitrarily, we let $\mathbf{x}^* = \Pi_{\mathcal{X}^*}(\mathbf{x}_0)$ and conclude the desired inequality.

C. Missing Proofs in Section 4

C.1. Proof of Lemma 4.1

Using the update formula of $x_{i,t+1}$ in Eq. (1), we have the following for any $x_i^* \in \mathcal{X}_i^*$:

$$\|x_{i,t+1} - x_i^*\|^2 \ = \ \|x_{i,t} + \eta_{t+1} \hat{v}_{i,t+1} - x_i^*\|^2.$$

which implies that

$$\|x_{i,t+1} - x_i^*\|^2 \ = \ \|x_{i,t} - x_i^*\|^2 + \eta_{t+1}^2 \|\hat{v}_{i,t+1}\|^2 + 2\eta_{t+1}(x_{i,t} - x_i^*)^\top \hat{v}_{i,t+1}.$$

Summing up the above inequality over $i \in \mathcal{N}$ and rearranging yields that

$$\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \ = \ \|\mathbf{x}_t - \mathbf{x}^*\|^2 + 2\eta_{t+1}(\mathbf{x}_t - \mathbf{x}^*)^\top \hat{v}_{t+1} + \eta_{t+1}^2 \|\hat{v}_{t+1}\|^2. \tag{18}$$

Using Young's inequality, we have

$$\begin{aligned} \|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \ \ \leq \ \ & \|\mathbf{x}_t - \mathbf{x}^*\|^2 + 2\eta_{t+1}^2 \|\xi_{t+1}\|^2 \\ & + 2\eta_{t+1}^2 \|\mathbf{v}(\mathbf{x}_t)\|^2 + 2\eta_{t+1}(\mathbf{x}_t - \mathbf{x}^*)^\top (\mathbf{v}(\mathbf{x}_t) + \xi_{t+1}). \end{aligned}$$

Since $\mathbf{x}^* \in \mathcal{X}^*$ and $\mathcal{G}$ is a λ-cocoercive game, we have $\mathbf{v}(\mathbf{x}^*) = 0$ and

$$\begin{aligned} (\mathbf{x}_t - \mathbf{x}^*)^\top \mathbf{v}(\mathbf{x}_t) \ \ &= \ \ (\mathbf{x}_t - \mathbf{x}^*)^\top (\mathbf{v}(\mathbf{x}_t) - \mathbf{v}(\mathbf{x}^*)) \ \leq \ -\lambda \|\mathbf{v}(\mathbf{x}_t) - \mathbf{v}(\mathbf{x}^*)\|^2 \\ &= \ \ -\lambda \|\mathbf{v}(\mathbf{x}_t)\|^2. \end{aligned}$$

Putting these pieces yields the desired inequality.

C.2. Proof of Lemma 4.2

Using the same argument as in Lemma 4.1, we have

$$\mathbb{E}[\epsilon(\mathbf{x}_{t+1}) \mid \mathcal{F}_t] - \epsilon(\mathbf{x}_t) \ \leq \ \frac{\mathbb{E}[\|\xi_{t+1}\|^2 \mid \mathcal{F}_t]}{\lambda \eta_{t+1}} + \left(1 - \frac{\lambda}{\eta_{t+1}}\right) \mathbb{E}\left[\|\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t)\|^2 \mid \mathcal{F}_t\right].$$

Since the noisy model (2) is with relative random noise (4), we have $\mathbb{E}[\|\xi_{t+1}\|^2 \mid \mathcal{F}_t] \leq \tau_t \|\mathbf{v}(\mathbf{x}_t)\|^2$. Also, $\eta_t \in (0, \lambda)$ for all $t \geq 1$. Therefore, we conclude that

$$\mathbb{E}[\epsilon(\mathbf{x}_{t+1}) \mid \mathcal{F}_t] - \epsilon(\mathbf{x}_t) \ \leq \ \frac{\tau_t \|\mathbf{v}(\mathbf{x}_t)\|^2}{\lambda \eta_{t+1}}.$$

Taking an expectation of both sides yields the desired inequality.

C.3. Proof of Theorem 4.4

Using the same argument as in Theorem 4.3, we obtain that

$$\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \mid \mathcal{F}_t] \ \leq \ \|\mathbf{x}_t - \mathbf{x}^*\|^2 - 2(\lambda - \overline{\eta} - \tau\overline{\eta})\eta_{t+1}\|\mathbf{v}(\mathbf{x}_t)\|^2. \tag{19}$$

Taking an expectation of both sides of Eq. (19) and rearranging yields that

$$\mathbb{E}[\epsilon(\mathbf{x}_t)] \ \leq \ \frac{1}{2(\lambda - \overline{\eta} - \tau\overline{\eta})\eta_{t+1}} \left(\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2] - \mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2]\right). \tag{20}$$

Summing up the above inequality over $t = 0, 1, \ldots, T$ yields that

$$\mathbb{E}\left[\sum_{t=0}^T \epsilon(\mathbf{x}_t)\right] \ \leq \ \sum_{t=1}^T \left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right) \frac{\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2]}{2(\lambda - \overline{\eta} - \tau\overline{\eta})} + \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - \overline{\eta} - \tau\overline{\eta})\eta_1}.$$

On the other hand, we have $\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2] \ \leq \ \mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2]$. This implies that $\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2] \leq \|\mathbf{x}_0 - \mathbf{x}^*\|^2$ for all $t \geq 1$. Therefore, we conclude that

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=0}^T \epsilon(\mathbf{x}_t)\right] \ &\leq \ \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - \overline{\eta} - \tau\overline{\eta})\eta_1} + \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - \overline{\eta} - \tau\overline{\eta})} \sum_{t=1}^T \left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right) \\
&= \ \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - \overline{\eta} - \tau\overline{\eta})\eta_1} + \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - \overline{\eta} - \tau\overline{\eta})\eta_{T+1}} \\
&\leq \ \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{(\lambda - \overline{\eta} - \tau\overline{\eta})\underline{\eta}} \ = \ O(1).
\end{aligned}$$

This completes the proof.

C.4. Proof of Theorem 4.7

Since the step-size sequence $\{\eta_t\}_{t \geq 1}$ is decreasing and converges to zero, we define the first iconic time in our analysis as follows,

$$t^* \ = \ \max\left\{t \geq 0 \mid \eta_{t+1} > \frac{\lambda}{2(1+\tau)}\right\} \ < \ +\infty.$$

First, we claim that $\mathbb{E}[\|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\|] \leq D$ where $D = \max_{1 \leq t \leq t^*} \mathbb{E}[\|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\|]$. Indeed, it suffices to show that $\mathbb{E}[\|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\|] \leq \mathbb{E}[\|\mathbf{x}_{t^*} - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\|]$ holds for $t > t^*$. By the definition of t^*, we have $\eta_{t+1} < \lambda/(1+\tau)$ for all $t > t^*$. The desired inequality follows from Eq. (20) and the fact that $\mathbb{E}[\epsilon(\mathbf{x}_t)] \geq 0$ for all $t > t^*$.

Furthermore, we derive an upper bound for the term $\sum_{t=0}^T \|\mathbf{v}(\mathbf{x}_t)\|^2$. Using the update formula (cf. Eq. (1)) to obtain that

$$(\mathbf{x}_{t+1} - \mathbf{x}^*)^\top (\mathbf{v}(\mathbf{x}_t) + \xi_{t+1}) = \frac{1}{2\eta_{t+1}} \left(\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2 + \|\mathbf{x}^* - \mathbf{x}_{t+1}\|^2 - \|\mathbf{x}^* - \mathbf{x}_t\|^2\right).$$

Recall that $\mathcal{G}$ is λ-cocoercive and the noisy model is defined with relative random noise, we have

$$\mathbb{E}[(\mathbf{x}_t - \mathbf{x}^*)^\top (\mathbf{v}(\mathbf{x}_t) + \xi_{t+1}) \mid \mathcal{F}_t] \ \geq \ \lambda\|\mathbf{v}(\mathbf{x}_t)\|^2.$$

Using Young's inequality, we have

$$\mathbb{E}[(\mathbf{x}_{t+1} - \mathbf{x}_t)^\top (\mathbf{v}(\mathbf{x}_t) + \xi_{t+1}) \mid \mathcal{F}_t] \ \geq \ -\frac{\lambda\|\mathbf{v}(\mathbf{x}_t)\|^2}{2} - \frac{(1+\tau)\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}_t\|^2 \mid \mathcal{F}_t]}{\lambda}.$$

Putting these pieces together and taking an expectation yields that

$$\lambda\mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2] \ \leq \ \mathbb{E}\left[\frac{\|\mathbf{x}^* - \mathbf{x}_t\|^2 - \|\mathbf{x}^* - \mathbf{x}_{t+1}\|^2}{\eta_{t+1}}\right] + \mathbb{E}\left[\left(\frac{2(1+\tau)}{\lambda} - \frac{1}{\eta_{t+1}}\right)\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2\right]. \tag{21}$$

Recall that the step-size sequence $\{\eta_t\}_{t\geq 1}$ is nonincreasing and $\|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\| \leq D$, we let $\mathbf{x}^* = \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})$ in Eq. (21) and obtain that

$$\sum_{t=0}^{T} \lambda \mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2] \ \leq \ \mathbb{E}\left[\frac{D^2}{\eta_{T+1}}\right] + \sum_{t=0}^{T} \mathbb{E}\left[\left(\frac{2(1+\tau)}{\lambda} - \frac{1}{\eta_{t+1}}\right)\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2\right].$$

To proceed, we define the second iconic time as

$$t_1^* \ = \ \max\left\{t \geq 0 \mid \eta_{t+1} > \frac{\lambda}{4(1+\tau)D^2 + 2(1+\tau)}\right\} \ > \ t^*.$$

It is clear that $t_1^* < +\infty$ and $\eta_{t+1} \leq \lambda/(2+2\tau)$ for all $t > t_1^*$ which implies that $(2+2\tau)/\lambda - 1/\eta_{t+1} \leq 0$. Assume T sufficiently large without loss of generality, we have

$$\sum_{t=0}^{T} \lambda \mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2] \ \leq \ \mathbb{E}\left[\frac{D^2}{\eta_{T+1}}\right] + \frac{2(1+\tau)}{\lambda}\left(\sum_{t=0}^{t_1^*} \mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2]\right) \ = \ \mathrm{I} + \mathrm{II}.$$

We also use Lemmas A.1 and A.2 from Bach & Levy (2019) to bound terms I and II. For convenience, we present these two lemmas here:

Lemma C.1 *For a sequence of numbers $a_0, a_1, \ldots, a_n \in [0, a]$ and $b \geq 0$, the following inequality holds:*

$$\sqrt{b + \sum_{i=0}^{n-1} a_i} - \sqrt{b} \ \leq \ \sum_{i=0}^{n} \frac{a_i}{\sqrt{b + \sum_{j=0}^{i-1} a_j}} \ \leq \ \frac{2a}{\sqrt{b}} + 3\sqrt{a} + 3\sqrt{b + \sum_{i=0}^{n-1} a_i}.$$

Bounding term I: We derive from the definition of η_t and Jensen's inequality that

$$\mathrm{I} \ \leq \ D^2\sqrt{\beta + \log(T+1) + \sum_{j=0}^{T-1} \mathbb{E}\left[\frac{\|\mathbf{x}_j - \mathbf{x}_{j+1}\|^2}{\eta_{j+1}^2}\right]}.$$

Since $\mathbb{E}[\|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\|] \leq D$ for all $t \geq 0$ and the notion of λ-cocercivity implies the notion of $(1/\lambda)$-Lipschiz continuity, we have

$$\mathbb{E}\left[\frac{\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2}{\eta_{t+1}^2}\right] \ \leq \ (2+2\tau)\mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2] \ \leq \ \frac{(2+2\tau)\|\mathbf{x}_t - \Pi_{\mathcal{X}^*}(\mathbf{x}_{t^*})\|^2}{\lambda^2} \ \leq \ \frac{(2+2\tau)D^2}{\lambda^2}.$$

Using the first inequality in Lemma C.1, we have

$$\begin{aligned}
\mathrm{I} \ \ \leq \ \ & D^2\sqrt{\beta + \log(T+1)} + \sum_{t=0}^{T} \frac{D^2\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2/\eta_{t+1}^2]}{\sqrt{\beta + \log(T+1) + \sum_{j=0}^{t-1}\mathbb{E}[\|\mathbf{x}_j - \mathbf{x}_{j+1}\|^2/\eta_{j+1}^2]}} \\
\leq \ \ & D^2\sqrt{\beta + \log(T+1)} + \sum_{t=0}^{t_1^*} \frac{D^2\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2/\eta_{t+1}^2]}{\sqrt{\beta + \log(T+1) + \sum_{j=0}^{t-1}\mathbb{E}[\|\mathbf{x}_j - \mathbf{x}_{j+1}\|^2/\eta_{j+1}^2]}} \\
& + \sum_{t=t_1^*+1}^{T} D^2\eta_{t+1}\mathbb{E}\left[\frac{\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2}{\eta_{t+1}^2}\right] \\
\leq \ \ & D^2\sqrt{\beta + \log(T+1)} + \sum_{t=0}^{t_1^*} \frac{D^2\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2/\eta_{t+1}^2]}{\sqrt{\beta + \log(T+1) + \sum_{j=0}^{t-1}\mathbb{E}[\|\mathbf{x}_j - \mathbf{x}_{j+1}\|^2/\eta_{j+1}^2]}} \\
& + \sum_{t=t_1^*+1}^{T} (2+2\tau)D^2\eta_{t+1}\mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2]. \qquad (22)
\end{aligned}$$

Since $\eta_{t+1} \leq \lambda/[4(1+\tau)D^2]$ for all $t > t_1^*$, we have

$$\sum_{t=t_1^*+1}^{T} (2+2\tau)D^2\eta_{t+1}\mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2] \;\leq\; \sum_{t=t_1^*+1}^{T} \frac{\lambda\mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2]}{2}. \tag{23}$$

Using the second inequality in Lemma 3.5, we have

$$\begin{aligned} & \sum_{t=0}^{t_1^*} \frac{D^2\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2/\eta_{t+1}^2]}{\sqrt{\beta + \log(T+1) + \sum_{j=0}^{t-1}\mathbb{E}[\|\mathbf{x}_j - \mathbf{x}_{j+1}\|^2/\eta_{j+1}^2]}} \\ \leq \;\; & \frac{(4+4\tau)D^2}{\lambda^2\sqrt{\beta+\log(T+1)}} + \frac{3D\sqrt{2+2\tau}}{\lambda} + 3\sqrt{\beta + \log(T+1) + \sum_{j=0}^{t_1^*-1}\mathbb{E}\left[\frac{\|\mathbf{x}_j - \mathbf{x}_{j+1}\|^2}{\eta_{j+1}^2}\right]}. \end{aligned} \tag{24}$$

By the definition of η_t, we have

$$\begin{aligned} & \sqrt{\beta + \log(T+1) + \sum_{j=0}^{t_1^*-1}\mathbb{E}\left[\frac{\|\mathbf{x}_j - \mathbf{x}_{j+1}\|^2}{\eta_{j+1}^2}\right]} \\ \leq \;\; & \frac{1}{\eta_{t_1^*+1}} + \sqrt{\log(T+1)} \;<\; \frac{4(1+\tau)D^2 + 2(1+\tau)}{\lambda} + \sqrt{\log(T+1)}. \end{aligned} \tag{25}$$

Putting Eq. (23)-(25) together yields that

$$\begin{aligned} \mathrm{I} \;\; \leq \;\; & D^2\sqrt{\beta+\log(T+1)} + \frac{(4+4\tau)D^2}{\lambda^2\sqrt{\beta+\log(T+1)}} + \frac{3D\sqrt{2+2\tau}}{\lambda} + \frac{12(1+\tau)D^2 + 6(1+\tau)}{\lambda} \\ & + \sqrt{\log(T+1)} + \sum_{t=t_1^*+1}^{T}\frac{\lambda\mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2]}{2}. \end{aligned}$$

Bounding term II: Recalling that

$$\mathbb{E}\left[\frac{\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2}{\eta_{t+1}^2}\right] \;\leq\; \frac{(2+2\tau)D^2}{\lambda^2},$$

and $\eta_t \leq 1/\beta$ for all $t \geq 1$, we have

$$\mathbb{E}\left[\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2\right] \;\leq\; \frac{(2+2\tau)D^2}{\lambda^2\beta^2},$$

Putting these pieces together yields that

$$\mathrm{II} \;=\; \frac{2(1+\tau)}{\lambda}\left(\sum_{t=0}^{t_1^*}\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2]\right) \;\leq\; \frac{4(1+\tau)^2D^2t_1^*}{\lambda^3\beta^2}.$$

Therefore, we have

$$\begin{aligned} \sum_{t=0}^{T}\frac{\lambda\mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2]}{2} \;\; \leq \;\; & D^2\sqrt{\beta+\log(T+1)} + \sqrt{\log(T+1)} + \frac{(4+4\tau)D^2}{\lambda^2\sqrt{\beta+\log(T+1)}} \\ & + \frac{3D\sqrt{2+2\tau}}{\lambda} + \frac{12(1+\tau)D^2 + 6(1+\tau)}{\lambda} + \frac{4(1+\tau)^2D^2t_1^*}{\lambda^3\beta^2}. \end{aligned}$$

By the definition, we have $t_1^* < +\infty$ is uniformly bounded. To this end, we conclude that $\sum_{t=0}^{T}\mathbb{E}[\|\mathbf{v}(\mathbf{x}_t)\|^2] \leq C_1 + C_2\sqrt{\log(T+1)}$, where $C_1 > 0$ and $C_2 > 0$ are universal constants.

Finally, we proceed to bound the term $\epsilon(\mathbf{x}_T)$. Without loss of generality, we can start the sequence at a later index t_1^* since $t_1^* < +\infty$. This implies that $\eta_{t+1} \leq \lambda/2(1+\tau)$. Using the last equation in the proof of Lemma 4.2, we have

$$\mathbb{E}[\epsilon(\mathbf{x}_T)] \ \leq \ \mathbb{E}[\epsilon(\mathbf{x}_t)] + \sum_{j=t}^{T-1} \frac{\tau_j}{\lambda} \mathbb{E}\left[\frac{\|\mathbf{v}(\mathbf{x}_j)\|^2}{\eta_{j+1}}\right].$$

Summing up the above inequality over $t = t_1^*, \ldots, T+$ yields

$$(T - t_1^* + 1)\mathbb{E}[\epsilon(\mathbf{x}_T)] \ \leq \ \sum_{t=t_1^*}^{T} \mathbb{E}[\epsilon(\mathbf{x}_t)] + \frac{1}{\lambda}\left(\sum_{t=t_1^*}^{T-1}\sum_{j=t}^{T-1} \tau_j \mathbb{E}\left[\frac{\|\mathbf{v}(\mathbf{x}_j)\|^2}{\eta_{j+1}}\right]\right).$$

Since $\{\tau_t\}_{t \geq 0}$ is an nonincreasing sequence, we have

$$\sum_{t=t_1^*}^{T-1}\sum_{j=t}^{T-1} \tau_j \mathbb{E}\left[\frac{\|\mathbf{v}(\mathbf{x}_j)\|^2}{\eta_{j+1}}\right] \ \leq \ \left(\sum_{t=0}^{T-1} \tau_j\right)\left(\sum_{t=t_1^*}^{T-1} \mathbb{E}\left[\frac{\|\mathbf{v}(\mathbf{x}_j)\|^2}{\eta_{j+1}}\right]\right).$$

Using Eq. (19) and $\eta_{t+1} \leq \lambda/2(1+\tau)$ for all $t > t_1^*$, we have

$$\mathbb{E}\left[\frac{\|\mathbf{v}(\mathbf{x}_j)\|^2}{\eta_{j+1}}\right] \ \leq \ \mathbb{E}\left[\frac{\|\mathbf{x}_j - \mathbf{x}^*\|^2 - \|\mathbf{x}_{j+1} - \mathbf{x}^*\|^2}{\eta_{j+1}^2}\right].$$

Note that $\{\eta_t\}_{t \geq 0}$ is an nonnegative and nonincreasing sequence and $\mathbb{E}[\|\mathbf{x}_{j+1} - \mathbf{x}^*\|^2] \leq D^2$. Putting these pieces together yields that

$$\begin{aligned}
\sum_{t=t_1^*}^{T-1} \mathbb{E}\left[\frac{\|\mathbf{v}(\mathbf{x}_j)\|^2}{\eta_{j+1}}\right] \ &\leq \ \mathbb{E}\left[\frac{D^2}{\eta_T^2}\right] \ \leq \ D^2\left(\beta + \log(T) + \sum_{t=0}^{T} \mathbb{E}\left[\frac{\|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2}{\eta_{t+1}^2}\right]\right) \\
&\leq \ D^2\left(\beta + \log(T) + 2(1+\tau)\sum_{t=0}^{T} \mathbb{E}\left[\|\mathbf{v}(\mathbf{x}_t)\|^2\right]\right) \\
&= \ O(\log(T)).
\end{aligned}$$

Therefore, we conclude that

$$\mathbb{E}[\epsilon(\mathbf{x}_T)] \ \leq \ \frac{\sum_{t=t_1^*}^{T} \mathbb{E}[\epsilon(\mathbf{x}_t)]}{T - t_1^* + 1} + \frac{C\log(T+1)}{\lambda(T - t_1^* + 1)}\left(\sum_{t=0}^{T-1} \tau_t\right) \ \text{for some } C > 0.$$

This completes the proof.

D. Missing Proofs in Section B

D.1. Proof of Lemma A.1

By the definition of $\epsilon(\mathbf{x})$, we have

$$\begin{aligned}
\epsilon(\mathbf{x}_{t+1}) - \epsilon(\mathbf{x}_t) \ &= \ \|\mathbf{v}(\mathbf{x}_{t+1})\|^2 - \|\mathbf{v}(\mathbf{x}_t)\|^2 \ = \ (\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t))^\top(\mathbf{v}(\mathbf{x}_{t+1}) + \mathbf{v}(\mathbf{x}_t)) \\
&= \ 2(\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t))^\top \mathbf{v}(\mathbf{x}_t) + \|\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t)\|^2.
\end{aligned}$$

Using the update formula in Eq. (1), it holds that $\mathbf{v}(\mathbf{x}_t) = \eta_{t+1}^{-1}(\mathbf{x}_{t+1} - \mathbf{x}_t) - \xi_{t+1}$. Therefore, we have

$$\begin{aligned}
\epsilon(\mathbf{x}_{t+1}) - \epsilon(\mathbf{x}_t) \ &= \ \frac{2}{\eta_{t+1}}\left((\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t))^\top(\mathbf{x}_{t+1} - \mathbf{x}_t) - (\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t))^\top \xi_{t+1}\right) \\
&\quad + \|\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t)\|^2.
\end{aligned}$$

Since $\mathcal{G}$ is a λ-cocoercive game, we have

$$(\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t))^\top (\mathbf{x}_{t+1} - \mathbf{x}_t) \ \leq \ -\lambda \|\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t)\|^2.$$

Using Young's inequality, we have

$$-(\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t))^\top \xi_{t+1} \ \leq \ \frac{\lambda \|\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t)\|^2}{2} + \frac{\|\xi_{t+1}\|^2}{2\lambda}.$$

Putting these pieces together yields that

$$\epsilon(\mathbf{x}_{t+1}) - \epsilon(\mathbf{x}_t) \ \leq \ \frac{\|\xi_{t+1}\|^2}{\lambda \eta_{t+1}} + \left(1 - \frac{\lambda}{\eta_{t+1}}\right) \|\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t)\|^2. \tag{26}$$

Taking an expectation of Eq. (26) conditioned on $\mathcal{F}_t$ yields that

$$\mathbb{E}[\epsilon(\mathbf{x}_{t+1}) \mid \mathcal{F}_t] - \epsilon(\mathbf{x}_t) \ \leq \ \frac{\mathbb{E}[\|\xi_{t+1}\|^2 \mid \mathcal{F}_t]}{\lambda \eta_{t+1}} + \left(1 - \frac{\lambda}{\eta_{t+1}}\right) \mathbb{E}\left[\|\mathbf{v}(\mathbf{x}_{t+1}) - \mathbf{v}(\mathbf{x}_t)\|^2 \mid \mathcal{F}_t\right].$$

Since the noisy model (2) is with absolute random noise (4), we have $\mathbb{E}[\|\xi_{t+1}\|^2 \mid \mathcal{F}_t] \leq \sigma_t^2$. Also, $\eta_t \in (0, \lambda)$ for all $t \geq 1$. Therefore, we conclude that

$$\mathbb{E}[\epsilon(\mathbf{x}_{t+1}) \mid \mathcal{F}_t] - \epsilon(\mathbf{x}_t) \ \leq \ \frac{\sigma_t^2}{\lambda \eta_{t+1}}.$$

Taking the expectation of both sides yields the desired inequality.

D.2. Proof of Theorem A.2

Recalling Eq. (11) (cf. Lemma 4.1), we take the expectation of both sides conditioned on $\mathcal{F}_t$ and obtain that

$$\begin{aligned}\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \mid \mathcal{F}_t] \ \ &\leq \ \ \|\mathbf{x}_t - \mathbf{x}^*\|^2 - (2\lambda - 2\eta_{t+1})\eta_{t+1}\|\mathbf{v}(\mathbf{x}_t)\|^2 + 2\eta_{t+1}^2 \mathbb{E}[\|\xi_{t+1}\|^2 \mid \mathcal{F}_t] \\ &\quad + 2\eta_{t+1}\mathbb{E}[(\mathbf{x}_t - \mathbf{x}^*)^\top \xi_{t+1} \mid \mathcal{F}_t].\end{aligned}$$

Since the noisy model (2) is with absolute random noise (4) satisfying $\sigma_t^2 \in (0, \sigma^2]$ for some $\sigma^2 < +\infty$, we have $\mathbb{E}[(\mathbf{x}_t - \mathbf{x}^*)^\top \xi_{t+1} \mid \mathcal{F}_t] = 0$ and $\mathbb{E}[\|\xi_{t+1}\|^2 \mid \mathcal{F}_t] \leq \sigma^2$. Therefore, we have

$$\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \mid \mathcal{F}_t] \ \leq \ \|\mathbf{x}_t - \mathbf{x}^*\|^2 - (2\lambda - 2\eta_{t+1})\eta_{t+1}\|\mathbf{v}(\mathbf{x}_t)\|^2 + 2\eta_{t+1}^2\sigma^2.$$

Since $\sum_{t=1}^\infty \eta_t^2 < \infty$, we have $\eta_t \to 0$ as $t \to +\infty$. Without loss of generality, we assume $\eta_t \leq \lambda$ for all t. Then we have

$$\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \mid \mathcal{F}_t] \ \leq \ \|\mathbf{x}_t - \mathbf{x}^*\|^2 + 2\eta_{t+1}^2\sigma^2. \tag{27}$$

We let $M_t = \|\mathbf{x}_t - \mathbf{x}^*\|^2 + 2\sigma^2(\sum_{j>t} \eta_j^2)$ and obtain that M_t is an nonnegative supermartingale. Then Doob's martingale convergence theorem shows that M_n converges to an nonnegative and integrable random variable almost surely. Let $M_\infty = \lim_{t \to +\infty} M_t$, it suffices to show that $M_\infty = 0$ almost surely now.

We first claim that *every neighborhood U of $\mathcal{X}^*$ is recurrent*: there exists a subsequence $\mathbf{x}_{t_k}$ of $\mathbf{x}_t$ such that $\mathbf{x}_{t_k} \to \mathcal{X}^*$ almost surely. Equivalently, there exists a Nash equilibria $\mathbf{x}^* \in \mathcal{X}^*$ such that $\|\mathbf{x}_{t_k} - \mathbf{x}^*\|^2 \to 0$ almost surely. To this end, we can define M_t with such Nash equilibria. Since $\sum_{t=1}^\infty \eta_t^2 < \infty$, we have $\sum_{j>t} \eta_j^2 \to 0$ as $t \to +\infty$ and the following statement holds almost surely:

$$\lim_{k \to +\infty} M_{t_k} \ = \ \lim_{k \to +\infty} \|\mathbf{x}_{t_k} - \mathbf{x}^*\|^2 \ = \ 0.$$

Since the whole sequence is guaranteed to converge to M_∞ almost surely, we conclude that $M_\infty = 0$ almost surely.

Proof of the recurrence claim: Let U be a neighborhood of $\mathcal{X}^*$ and assume to the contrary that, $\mathbf{x}_t \notin U$ for sufficiently large t with positive probability. By starting the sequence at a later index if necessary and noting that $\sum_{t=1}^\infty \eta_t^2 < \infty$, we may assume that $\mathbf{x}_t \notin U$ and $\eta_t \leq \lambda/2$ for all t without loss of generality. Thus, there exists some $c > 0$ such that $\|\mathbf{v}(\mathbf{x}_t)\|^2 \geq c$ for all t. As a result, for all $\mathbf{x}^* \in \mathcal{X}^*$, we let $\psi_{t+1} = (\mathbf{x}_t - \mathbf{x}^*)^\top \xi_{t+1}$ and have

$$\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \ \leq \ \|\mathbf{x}_t - \mathbf{x}^*\|^2 - \lambda c \eta_{t+1} + 2\eta_{t+1}\psi_{t+1} + 2\eta_{t+1}^2\|\xi_{t+1}\|^2.$$

Summing up the above inequality over $t = 0, 1, \ldots, T$ together with $\theta_t = \sum_{j=1}^t \eta_j$ yields that

$$\|\mathbf{x}_{T+1} - \mathbf{x}^*\|^2 \ \leq \ \|\mathbf{x}_0 - \mathbf{x}^*\|^2 - \lambda c\theta_{T+1} + 2\theta_{T+1}\left[\frac{\sum_{t=1}^{T+1}\eta_t\psi_t}{\theta_{T+1}} + \frac{\sum_{t=1}^{T+1}\eta_t^2\|\xi_t\|^2}{\theta_{T+1}}\right]. \tag{28}$$

Since the noisy model (2) is with absolute random noise (4) satisfying $\sigma_t^2 \in (0, \sigma^2]$ for some $\sigma^2 < +\infty$, we have $\mathbb{E}[\psi_{t+1} \mid \mathcal{F}_t] = 0$. Furthermore, we obtain by taking the expectation of both sides of Eq. (27) that

$$\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2] \ \leq \ \mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2] + 2\eta_{t+1}^2\sigma^2,$$

and the following inequality holds true for all $t \geq 1$:

$$\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2] \leq \|\mathbf{x}_0 - \mathbf{x}^*\|^2 + 2\sigma^2\sum_{j=1}^{\infty}\eta_j^2 \ < \ +\infty.$$

Since $\|\mathbf{x}_t - \mathbf{x}^*\|^2 \geq 0$, we have $\|\mathbf{x}_t - \mathbf{x}^*\|^2 < +\infty$ almost surely. Therefore, $\mathbb{E}[\|\psi_{t+1}\|^2 \mid \mathcal{F}_t] \leq \sigma^2\|\mathbf{x}_t - \mathbf{x}^*\|^2 < +\infty$. Then the law of large numbers for martingale differences yields that $\theta_{T+1}^{-1}(\sum_{t=1}^{T+1}\eta_t\psi_t) \to 0$ almost surely (Hall & Heyde, 2014, Theorem 2.18). Furthermore, let $R_t = \sum_{j=1}^t \eta_j^2\|\xi_j\|^2$, then R_t is a submartingale and

$$\mathbb{E}[R_t] \ \leq \ \sigma^2\sum_{j=1}^t\eta_j^2 < \sigma^2\sum_{j=1}^{\infty}\eta_j^2 \ < \ +\infty.$$

From Doob's martingale convergence theorem, R_t converges to some random, finite value almost surely (Hall & Heyde, 2014, Theorem 2.5). Putting these pieces together with Eq. (28) yields that $\|\mathbf{x}_t - \mathbf{x}^*\|^2 \sim -\lambda c\tau_t \to -\infty$ almost surely, a contradiction. Therefore, we conclude that every neighborhood of $\mathcal{X}^*$ is recurrent.

D.3. Proof of Theorem A.3

Since $\eta_t = c/\sqrt{t}$ for all $t \geq 1$, we have $\eta_t \to 0$ and $\eta_t \leq c$ for all $t \geq 1$. This implies that

$$\lambda\eta_{t+1} - \eta_{t+1}^2 \ \geq \ (\lambda - c)\eta_{t+1}. \tag{29}$$

Plugging Eq. (29) into Eq. (11) (cf. Lemma 4.1) yields that

$$\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \ \leq \ \|\mathbf{x}_t - \mathbf{x}^*\|^2 - 2(\lambda - c)\eta_{t+1}\|\mathbf{v}(\mathbf{x}_t)\|^2 + 2\eta_{t+1}(\mathbf{x}_t - \mathbf{x}^*)^\top\xi_{t+1} + 2\eta_{t+1}^2\|\xi_{t+1}\|^2.$$

Using the same argument as in Theorem A.2, we have

$$\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2 \mid \mathcal{F}_t] \ \leq \ \|\mathbf{x}_t - \mathbf{x}^*\|^2 - 2(\lambda - c)\eta_{t+1}\|\mathbf{v}(\mathbf{x}_t)\|^2 + 2\eta_{t+1}^2\sigma^2. \tag{30}$$

Taking the expectation of both sides of Eq. (30) and rearranging yields that

$$\mathbb{E}[\epsilon(\mathbf{x}_t)] \ \leq \ \frac{1}{2(\lambda - c)\eta_{t+1}}\left(\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2] - \mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2]\right) + \frac{\eta_{t+1}\sigma^2}{\lambda - c}.$$

Summing up the above inequality over $t = 0, 1, \ldots, T$ and using $\eta_t = c/\sqrt{T+1}$ yields that

$$\mathbb{E}\left[\sum_{t=0}^T\epsilon(\mathbf{x}_t)\right] \ \leq \ \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - c)\eta_1} + \sum_{t=1}^T\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right)\frac{\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2]}{2(\lambda - c)} + \frac{\sigma^2}{\lambda - c}\left(\sum_{t=1}^{T+1}\eta_t\right).$$

On the other hand, we have

$$\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2] \ \leq \ \mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2] + 2\eta_{t+1}^2\sigma^2.$$

This implies that the following inequality holds for all $t \geq 1$:

$$\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2] \ \leq \ \|\mathbf{x}_0 - \mathbf{x}^*\|^2 + 2\sigma^2\left(\sum_{j=1}^t\eta_j^2\right) \ \leq \ \|\mathbf{x}_0 - \mathbf{x}^*\|^2 + 2\sigma^2c^2\log(t+1).$$

Therefore, we conclude that

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=0}^{T} \epsilon(\mathbf{x}_t)\right] &\leq \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2 + 2\sigma^2 c^2 \log(T+1)}{2(\lambda - c)} \sum_{t=1}^{T} \left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right) + \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - c)\eta_1} + \frac{\sigma^2}{\lambda - c}\left(\sum_{t=1}^{T+1} \eta_t\right) \\
&\leq \frac{\sqrt{T+1}(\|\mathbf{x}_0 - \mathbf{x}^*\|^2 + 2\sigma^2 c^2 \log(T+1))}{2c(\lambda - c)} + \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2c(\lambda - c)} + \frac{\sigma^2 c\sqrt{T+1}}{\lambda - c} \\
&= O\left(\sqrt{T+1}\log(T+1)\right).
\end{aligned}$$

This completes the proof.

D.4. Proof of Theorem A.4

Using Lemma A.1, we have

$$\mathbb{E}[\epsilon(\mathbf{x}_T)] \;\leq\; \mathbb{E}[\epsilon(\mathbf{x}_t)] + \frac{1}{\lambda}\left(\sum_{j=t}^{T-1} \frac{\sigma_j^2}{\eta_{j+1}}\right) \;\leq\; \mathbb{E}[\epsilon(\mathbf{x}_t)] + \frac{1}{\lambda\underline{\eta}}\left(\sum_{j=t}^{T-1} \sigma_j^2\right).$$

Summing up the above inequality over $t = 0, 1, \ldots, T$ yields that

$$(T+1)\mathbb{E}[\epsilon(\mathbf{x}_T)] \;\leq\; \sum_{t=0}^{T} \mathbb{E}[\epsilon(\mathbf{x}_t)] + \frac{1}{\lambda}\left(\sum_{t=0}^{T-1}\sum_{j=t}^{T-1} \frac{\sigma_j^2}{\eta_{j+1}}\right) \;\leq\; \sum_{t=0}^{T} \mathbb{E}[\epsilon(\mathbf{x}_t)] + \frac{1}{\lambda\underline{\eta}}\left(\sum_{t=0}^{T-1} (t+1)\sigma_t^2\right).$$

On the other hand, the derivation in Theorem A.3 implies that

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=0}^{T} \epsilon(\mathbf{x}_t)\right] &\leq \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - \overline{\eta})\eta_1} + \sum_{t=1}^{T}\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right)\frac{\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2]}{2(\lambda - \overline{\eta})} + \frac{1}{\lambda - \overline{\eta}}\left(\sum_{t=1}^{T+1} \eta_t\sigma_t^2\right) \\
&\leq \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - \overline{\eta})\eta_1} + \sum_{t=1}^{T}\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right)\frac{\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2]}{2(\lambda - \overline{\eta})} + \frac{\overline{\eta}}{\lambda - \overline{\eta}}\left(\sum_{t=1}^{T+1} \sigma_t^2\right).
\end{aligned}$$

On the other hand, we have

$$\mathbb{E}[\|\mathbf{x}_{t+1} - \mathbf{x}^*\|^2] \;\leq\; \mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2] + 2\eta_{t+1}^2\sigma_t^2.$$

This implies that the following inequality holds for all $t \geq 1$:

$$\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}^*\|^2] \;\leq\; \|\mathbf{x}_0 - \mathbf{x}^*\|^2 + 2\overline{\eta}^2\left(\sum_{j=1}^{t} \sigma_j^2\right).$$

Therefore, we conclude that

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=0}^{T} \epsilon(\mathbf{x}_t)\right] &\leq \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2 + 2\overline{\eta}^2(\sum_{t=1}^{T}\sigma_t^2)}{2(\lambda - \overline{\eta})} \sum_{t=1}^{T}\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right) + \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - \overline{\eta})\eta_1} + \frac{\overline{\eta}}{\lambda - \overline{\eta}}\left(\sum_{t=1}^{T+1} \sigma_t^2\right) \\
&\leq \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2 + 2\overline{\eta}^2(\sum_{t=1}^{T}\sigma_t^2)}{2(\lambda - \overline{\eta})\underline{\eta}} + \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(\lambda - \overline{\eta})\underline{\eta}} + \frac{\overline{\eta}}{\lambda - \overline{\eta}}\left(\sum_{t=1}^{T+1} \sigma_t^2\right) \\
&\leq \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{(\lambda - \overline{\eta})\underline{\eta}} + \left(1 + \frac{\overline{\eta}}{\underline{\eta}}\right)\frac{\overline{\eta}}{\lambda - \overline{\eta}}\left(\sum_{t=1}^{T+1} \sigma_t^2\right) \\
&\leq \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{(\lambda - \overline{\eta})\underline{\eta}} + \left(1 + \frac{\overline{\eta}}{\underline{\eta}}\right)\frac{\overline{\eta}}{\lambda - \overline{\eta}}\left(\sum_{t=1}^{T+1} (t+1)\sigma_t^2\right).
\end{aligned}$$

Putting these pieces together yields that

$$
\begin{aligned}
\mathbb{E}[\epsilon(\mathbf{x}_T)] \quad & \leq \quad \frac{1}{T+1}\left[\sum_{t=0}^{T}\mathbb{E}[\epsilon(\mathbf{x}_t)] + \frac{1}{\lambda\underline{\eta}}\left(\sum_{t=0}^{T-1}(t+1)\sigma_t^2\right)\right] \\
& \leq \quad \frac{1}{T+1}\left(\frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{(\lambda - \overline{\eta})\underline{\eta}} + \left(\frac{1}{\lambda\underline{\eta}} + \left(1 + \frac{\overline{\eta}}{\underline{\eta}}\right)\frac{\overline{\eta}}{\lambda - \overline{\eta}}\right)\left(\sum_{t=1}^{T+1}(t+1)\sigma_t^2\right)\right) \\
& \overset{\text{Eq. (14)}}{=} \quad O(a(T)).
\end{aligned}
$$

This completes the proof.