Multi-Agent Online Learning with Imperfect Information

Zhengyuan Zhou
Department of Electrical Engineering, Stanford University, zyzhou@stanford.edu

Panayotis Mertikopoulos
Univ. Grenoble Alpes, CNRS, Grenoble INP, Inria, LIG, F-38000 Grenoble, France, panayotis.mertikopoulos@image.fr

Nicholas Bambos, Peter Glynn
Department of Management Science and Engineering, Stanford University bambos@stanford.edu, glynn@stanford.edu

Claire Tomlin
Department of Electrical Engineering and Computer Sciences, UC Berkeley, tomlin@eecs.berkeley.edu

Multi-agent online learning revolves around multiple interacting agents that make sequential decisions online, each concerned with maximizing their individual rewards (which, a priori, typically depend on the actions of all other players). In this paper, we consider a model of multi-agent online learning where the game is not known in advance, and the agents' feedback is subject to both noise and delays. Motivated by its strong no-regret properties, we first focus on a class of learning algorithms known as online mirror descent (OMD), and we show that, even in the presence of noise, the induced sequence of play converges to Nash equilibria in a wide class of continuous games, provided that the feedback delays faced by the agents are synchronous and bounded. Subsequently, to tackle fully decentralized, asynchronous environments with unbounded feedback delays, we propose a variant of OMD which we call delayed mirror descent (DMD), and which relies on the repeated leveraging of past information. With this modification, the algorithm converges to Nash equilibria almost surely, even in noisy environments with no feedback synchronicity assumptions, and with feedback delays growing at a superlinear rate relative to the game's horizon.

Key words: Mirror descent; multi-agent learning; online decision making; decision making with imperfect data; stochastic approximation.

MSC2000 subject classification: Primary 91A26, 37N40; secondary 91A22, 90C25, 68T05.

History: A short conference version of this paper is presented at NIPS 2017. This paper was a 2017 INFORMS George Nicholson student paper award finalist and an Applied Society Probability student paper award finalist.

---

1. Introduction Online learning is a broad and powerful theoretical framework with widespread applications in machine learning, data science and operations research [6, 45, 3, 19]. In its most basic form, the prototypical online learning setting may be described as follows: At each round $t = 0, 1, \dots$, an agent selects an action x^t from a set of possible actions $\mathcal{X}$, and obtains a reward $u^t(x^t)$ based on some *a priori* unknown and (possibly) time-varying reward function $u^t(\cdot)$. Subsequently, the player receives some feedback (e.g., the past history of the reward functions or some restricted information thereof), and selects a new action x^{t+1} with the goal of maximizing the obtained reward. Aggregating over the stages of this online decision process, this is usually quantified by asking that the player's *regret* $\mathrm{Reg}(T) \equiv \max_{x\in\mathcal{X}} \sum_{t=1}^{T} \left[u^t(x) - u^t(x^t)\right]$

grow sublinearly with the horizon of play T, a property known as "no regret".

One of the most widely used algorithmic schemes for learning in this context is the class of online mirror descent (OMD) algorithms [44]. Tracing its origins to convex optimization [34, 35], OMD proceeds by taking a gradient step in a dual space (where gradients live) and projecting back to the primal (decision) space via a suitably chosen *mirror map*. In particular, OMD and its variants includes as special cases several seminal learning algorithms, such as Zinkevich's online gradient descent (OGD) scheme [52], the exponentia/multiplicative weights algorithm [26, 1], matrix regularization schemes [21], and many more.[1]

When each u^t is concave,[2] OMD enjoys a sublinear $\mathcal{O}(\sqrt{T})$ regret bound which is known to be universally tight. Further, the desirable property of OMD still holds even in the presence of delays. For instance, the authors of [39] recently considered a general delay model on the feedback, where the gradient at round t is only available at round $t + d^t - 1$, with d^t being the delay associated with the agent's action at round t. [39] then considered a natural extension of OMD under delays: updating the set of gradients as they are received. If the total delay after time T is $D(T) = \sum_{t=1}^{T} d^t$, then OMD enjoys an $\mathcal{O}(D(T)^{1/2})$ regret bound [39]. This natural extension has several strengths: no assumption is made on how the gradients are received (the delayed gradients can be received out-of-order); further, as pointed out in [39], a gradient does not need to be timestamped by the round from which it originates – in contrast to the earlier pooling strategies of [20, 13]. As such, OMD provides a broad and powerful algorithmic framework to make sequential online decisions in model-agnostic regimes, where the underlying environment is rapidly changing or adversarial.

The above concerns single-agent settings, where a sole decision maker interacts with an unknown and arbitrarily changing environment. Albeit powerful, this black-box abstraction cannot capture finer structures and details that arise in multi-agent environments where several optimizing players are engaged in a concurrent decision process, with the actions of one player affecting the rewards of another [9, 47, 8, 17]. In many cases of practical interest, the agents' reward structure can be modeled as a continuous game [15], where each agent's reward is determined by the joint action of all the agents via a fixed (but potentially unknown) mechanism – the *stage game*. In this case, from a unilateral viewpoint, the reward function of any given player varies with time as the game's players change their actions from one stage to the next. Since the regret bounds of OMD raise high expectations in terms of performance guarantees, it is natural to assume that agents adopt some variant of OMD when faced with such online decision processes.

As a result, several natural questions arise: If each agent adopts an OMD policy to minimize their individual regret, what is the resulting evolution of the players' joint action profile? Does it converge? And, if so,

[1] Several variants of this class also exist and, perhaps unsurprisingly, they occur with a variety of different names – such as "Follow-the-Regularized-Leader" [22], dual averaging [35, 50], and so on.

[2] Usually, the term "descent" refers to the long-standing tradition of objective function minimization in optimization; in such cases, the benchmark assumption is convex. In a slight abuse of terminology, we are using the term "descent" for the maximization of concave functions, to avoid introducing unwieldy new algorithmic names.

under what assumptions would the agents' long-term behavior be represented by a Nash equilibrium of the underlying game?

1.1. Related Work Despite the fact that multi-agent/game-theoretic learning has received significant scrutiny in the literature, the questions raised above are still open for several reasons.

First, in general, joint convergence of no-regret learning does not hold. In fact, even in (mixed extensions of) finite games, a class of games where each agent has a finite number of actions and the rewards for agents are specified by a matrix that records the reward for each agent under each joint action, OMD (and no-regret learning in general) can fail to converge [29]. Even worse, in general, no-regret learning does not necessarily eliminate dominated strategies: there exist games whose Hannan set (the limit set of no-regret learning policies) contains strategies that assign weight *only* to strictly dominated strategies [48]. In such cases, an agent applying the no-regret learning algorithm will select a strictly inferior action in equilibrium regardless of the other agents' actions. More generally, even in the absence of dominated strategies, if a Nash equilibrium is not reached, an agent can always obtain larger rewards by deviating from its current action. Consequently, establishing convergence to Nash equilibrium under no-regret learning algorithms for a broad and meaningful class of games has attracted considerable interest in the literature – for both finite and continuous games.

Much of the existing literature on this issue has focused on studying convergence in (mixed extensions of) finite games [47, 48, 5, 33]. More specifically, earlier work of game-theoretic learning (see [16] for a comprehensive review) has mainly focused on learning in finite games with dynamics that are not necessarily regret-less. The primary focus of [9] is convergence to coarser equilibrium notions (such as correlated or coarse correlated equilibria), where a fairly complete characterization is given. That being said, as pointed out in [9], convergence to Nash is a much more difficult problem: recent results of [49] have clearly highlighted the gap between (coarse) correlated equilibria obtained by no-regret learning processes and Nash equilibria, and more recent works have explored in depth the non-convergent behaviors that can emerge in this setting [27, 37, 30]. More positive results can be obtained in the class of potential games where, in a recent paper, the authors of [12] established the convergence of multiplicative weights and other regularized strategies in potential games with only payoff-based, bandit feedback.

However, much less is known beyond mixed extensions of finite games – i.e., in the case of *continuous* games. In the context of mixing in games with continuous action spaces, the authors of [38] provide a convergence analysis for a perturbed version of the multiplicative weights algorithm in potential games. In a pure-strategy setting, the network games considered in [28] belong to the much broader class of games known as concave games: each agent's reward function is individually concave. Therein, the dynamics investigated may lead to positive regret in the limit. Note that mixed extensions of finite games belong to the class of linear games (i.e. each agent's reward is individually linear in its own action), which belong to the class of concave games, which in turn belong to the class of continuous games. Another recent paper [2] studied a two-player continuous zero-sum game, and showed that if both players adopt a no-regret learning algorithm, then the empirical time-average of the joint action converges to Nash equilibria. However, barring a few recent exceptions, the territory of no-regret learning on concave games is not well understood (let alone in general games with continuous action sets). An exception to this is the recent paper [24] where the authors establish the convergence of mirror descent in concave potential games with perfect information and synchronous user updates – a result later extended to learning in monotone games [31].

Second, the convergence mode that is commonly adopted in the existing literature is that of ergodic convergence – i.e., convergence in the sense of time averages [9, 7, 10, 25]), rather than the convergence of the actual sequence of play (i.e. x^t). The former is convergence of the *empirical* frequency of play, while the latter is convergence of *actual* play: the latter implies the former, but not the other way round. In a game-theoretic context, convergence of the actual sequence of play is crucial for several reasons: *a*) convergence in the sense of time averages does not preclude that players might play subpar (e.g., strictly dominated) strategies infinitely often; *b*) the players' rewards are determined by their actions, not their time-averages, so ergodic convergence diminishes in predictive value if it is not accompanied by similar conclusions for the players' realized actions; and *c*) because there is no inherent averaging, the analysis of the actual sequence of play provides a much finer understanding of the evolution of the joint action. In fact, the difference between these two convergence modes was highlighted in [29], where it is shown that even though the time-average of continuous-time OMD converges to Nash equilibrium in bilinear zero-sum games, actual play orbits interior Nash equilibria in perpetuity. Some recent positive results in specific games do exist: [24, 25] studied nonatomic routing games (a special class of concave potential games) and established that multiplicative weights converge to Nash equilibria, while the authors of [37] established the convergence of the multiplicative weights algorithm to Nash equilibria under certain conditions in atomic non-splittable congestion games.

A third important point is that many of the works discussed above concern multi-agent online learning models with perfect information – in the sense that there are no errors or other corruption factors in the feedback that each agent receives at each round. However, due to the massive growth in scale in many of the operations systems captured by multi-agent models, a new set of challenges arises: *i*) Feedback may arrive with a significant delay (as agents might be spread across different geographical locations); *ii*) such feedback delays will inevitably be asynchronous due to the sheer number of agents; *iii*) the received feedback is susceptible to noise, measurement errors and information transmission losses; and *iv*) to make decisions online, a learning algorithm must collect and act on only minimal information from a high-dimensional space (e.g., joint action of all agents). This means that imperfect information, including both delays and noise (and possibly many other types of feedback impediments), is rapidly becoming the rule rather than the exception; however, this crucial feature has not been sufficiently addressed in the literature so far.

1.2. Our Contributions In this paper, we focus on the noise and delay/asynchronicity aspects of the challenges mentioned above. To begin with, we identify a broad class of (not necessarily concave) continuous games, which we call *λ-variationally stable*, and which contain several important classes of continuous games that arise in the literature. Within this class, we are able to treat convergence questions without having to consider highly structured frameworks, such as concave potential games (though our analysis also covers such games). For completeness, we also discuss several practical examples of games that are variationally stable in Sections 2 and 3.

In terms of learning, we focus on the widely used OMD family of learning algorithms, and we examine in depth its convergence properties in the presence of delays/asynchronicities and noise. This is accomplished by means of a tailor-made Lyapunov function, the so-called *λ-Fenchel coupling*, which serves as a "primal-dual" divergence measure between (dual) gradient steps and (primal) decision variables. In particular, we show that the sequence of play induced by OMD converges to the game's set of Nash equilibria in all stable games, provided that the delays of all players are synchronous and bounded (see Theorems 4.1 and 4.2). As an important feature of this result, it should be noted that players may be receiving gradients out-of-order and do not need to keep track of the stages from which a given bit of feedback originates from. Proceeding in a piecemeal fashion, we then consider additional feedback impediments in the form of noisy measurements and observation errors. In this case, we show that the sequence of play induced by OMD, now a stochastic process, converges to a Nash equilibrium with probability 1, under synchronous and bounded delays (Theorem 4.5).

Finally, to lift the requirement of synchronous delays, we introduce a modification of vanilla OMD, which we call *delayed mirror descent* (DMD), and which leverages past information repeatedly, even in rounds where players receive no feedback. Thanks to this modification, we show that, in stable games, the sequence of play induced by DMD converges to Nash equilibria almost surely, even when delays are not synchronized between players, and might even be unbounded relative to the horizon of the game (Theorem 5.1 and Lemma 5.1). Importantly, this result remains robust in the presence of noise: as shown in Theorem 5.2 the sequence of play induced by DMD converges to a Nash equilibrium as long as there is no systematic bias in the errors affecting the players' feedback process.

Our analysis draws tools from variational analysis, stochastic approximation, dynamical systems and martingale limit theory, and provides a flexible toolbox for examining the convergence of both OMD and DMD under a unified framework.

2. Problem Setup We formalize the multi-agent online learning under imperfect information framework.

2.1. Multi-Agent Reward Structure: Continuous Games We start with the definition of a continuous game, which provides a reward function for each player in an online learning process.

Definition 2.1. A continuous game $\mathcal{G}$ 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 $\{1, 2, \ldots, N\}$, $\mathcal{X}$ is the joint action space, with $\mathcal{X}_i$, a compact and convex subset of a finite-dimensional vector space $\mathbb{R}^{d_i}$, being the action space for player i. Finally, $u_i : \mathcal{X} \to \mathbf{R}$ is the reward function for player i, assuming to satisfy the following regularity conditions for all $i \in \mathcal{N}$:

1. $u_i(\mathbf{x})$ is continuous in $\mathbf{x}$.
2. u_i is continuously differentiable in x_i and the partial gradient $\nabla_{x_i} u_i(\mathbf{x})$ is Lipschitz continuous in $\mathbf{x}$.

Throughout the paper, $\mathbf{x}_{-i}$ denotes the joint action of all players but player i. Consequently, the joint action[3] $\mathbf{x}$ will frequently be written as $(x_i, \mathbf{x}_{-i})$. Note that we do not assume, as is typical in the single-agent online learning problem, that each reward function u_i is concave in x_i: such an assumption is crucial in obtaining good regret performance, but is fairly restrictive for the current multi-agent learning problem.

Next, we define two important quantities in the current context.

Definition 2.2. We denote by $\mathbf{v}(\mathbf{x})$ to be the column vector of all the partial gradients of the reward functions[4]: $\mathbf{v}(\mathbf{x}) = (v_1(\mathbf{x}), \ldots, v_N(\mathbf{x}))$, where $v_i(\mathbf{x}) \triangleq \nabla_{x_i} u_i(\mathbf{x})$.

Definition 2.3. Given a continuous game $(\mathcal{N}, \mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i, \{u_i\}_{i=1}^{N})$, $\mathbf{x}^* \in \mathcal{X}$ is called a Nash equilibrium if for each $i \in \mathcal{N}$, $u_i(x_i^*, \mathbf{x}_{-i}^*) \geq u_i(x_i, \mathbf{x}_{-i}^*), \forall x_i \in \mathcal{X}_i$.

Two comments on notational conventions used in this paper: First, $\langle \cdot, \cdot \rangle$ denotes inner product. However, when there is no confusion (particularly in the proofs), we abbreviate it to normal multiplication for convenience. For instance, we write $v_i(\mathbf{x})(x_i - x_i^*)$ as a shorthand in replacement of $\langle v_i(\mathbf{x}), x_i - x_i^* \rangle$. Second, let $\|\cdot\|$ be a vector norm on $\mathcal{X}$. Then $\|\cdot\|^*$ denote its dual norm: $\|y\|^* = \max_{\|x\| \leq 1} \langle x, y \rangle$. We use the superscript $\|\cdot\|^*$ (rather than a subscript) because each agent has its own norm $\|\cdot\|_i$ on $\mathcal{X}_i$; hence we reserve the subscript to agent index.

2.2. Examples Below, we present some motivating examples for the class of games under consideration:

Example 2.1 (Cournot oligopolies). In the standard Cournot oligopoly model, there is a finite set $\mathcal{N} = \{1, \ldots, N\}$ of *firms*, each supplying the market with a quantity $x_i \in [0, C_i]$ of some good (or service) up to the firm's production capacity, given here by a positive scalar $C_i > 0$. This good is then priced as a decreasing function $P(x)$ of the total supply to the market, as determined by each firm's production; for concreteness, we focus on the standard linear model $P(x) = a - b \sum_i x_i$ where a and b are positive constants.

[3] Note that boldfaced letters are only used to denote joint actions. In particular, x_i is a vector even though it is not boldfaced.

[4] Note that per the second assumption in the definition of a continuous game, the gradient $\mathbf{v}(\mathbf{x})$ always exists and is a continuous function on the joint action space $\mathcal{X}$.

In this model, the utility of firm i (considered here as a player) is given by

$$u_i(x) = x_i P(x) - c_i x_i, \tag{2.1}$$

where c_i represents the marginal production cost of firm i, i.e., as the income obtained by producing x_i units of the good in question minus the corresponding production cost. Letting $\mathcal{X}_i = [0, C_i]$ denote the space of possible production values for each firm, the resulting game $\mathcal{G} \equiv \mathcal{G}(\mathcal{N}, \mathcal{X}, u)$ is easily seen to be a continuous game in the sense described above.

Example 2.2 (Kelly auctions). Consider a service provider with a number of splittable resources $s \in \mathcal{S} = \{1, \dots, S\}$ (representing e.g., bandwidth, server time, ad space on a website, etc.). These resources can be leased to a set of N bidders (players) who can place monetary bids $x_{is} \geq 0$ for the utilization of each resource $s \in \mathcal{S}$ up to each player's total budget b_i, i.e., $\sum_{s \in \mathcal{S}} x_{is} \leq b_i$. A popular – and widely used – mechanism to allocate resources in this case is the so-called *Kelly mechanism* [23] whereby resources are allocated proportionally to each player's bid, i.e., player i gets

$$\rho_{is} = \frac{q_s x_{is}}{c_s + \sum_{j \in \mathcal{N}} x_{js}} \tag{2.2}$$

units of the s-th resource (in the above, q_s denotes the available units of said resource and $c_s \geq 0$ is the "entry barrier" for bidding on it). A simple model for the utility of player i is then given by

$$u_i(x_i; x_{-i}) = \sum_{s \in \mathcal{S}} [g_i \rho_{is} - x_{is}], \tag{2.3}$$

where $x_i = (x_{is})_{s \in \mathcal{S}}$ is the bid vector of player i and g_i denotes the player's marginal gain from acquiring a unit slice of resources. As before, if we write $\mathcal{X}_i = \{x_i \in \mathbb{R}_+^S : \sum_{s \in \mathcal{S}} x_{is} \leq b_i\}$ for the space of possible bids of player i on the set of resources $\mathcal{S}$, we again obtain a continuous game in the sense of the previous section.

Example 2.3 (Congestion games). Consider a set of players $\mathcal{N}$ that share a set of *amenities* $s \in \mathcal{S}$, each associated with a nondecreasing convex *cost function* $c_s \colon \mathbb{R}_+ \to \mathbb{R}$ (for instance, these amenities could be links in an urban traffic network and their corresponding delay functions). Each player $i \in \mathcal{N}$ has a certain *load* $\rho_i > 0$ which is split over a collection $\mathcal{A}_i \subseteq 2^{\mathcal{S}}$ of subsets of amenities α_i of $\mathcal{S}$ – e.g., sets of links that form paths in the network. Then, the set of possible actions of player $i \in \mathcal{N}$ can be represented by the scaled simplex $\mathcal{X}_i = \rho_i \Delta(\mathcal{A}_i) = \{x_i \in \mathbb{R}_+^{\mathcal{A}_i} : \sum_{\alpha_i \in \mathcal{A}_i} x_{i\alpha_i} = \rho_i\}$ of *load distributions* over $\mathcal{A}_i$.

Given a load profile $x = (x_1, \dots, x_N)$, costs are determined based on the utilization of each amenity as follows: First, the *demand* w_s of the s-th amenity is defined as the total load $w_s = \sum_{i \in \mathcal{N}} \sum_{\alpha_i \ni s} x_{i\alpha_i}$ on said amenity. This demand incurs a *cost* $c_s(w_s)$ per unit of load to each player utilizing amenity s, with

$c_s : \mathbb{R}_+ \to \mathbb{R}$ being here a nondecreasing convex function. Accordingly, the total cost to player $i \in \mathcal{N}$ is given by

$$c_i(x) = \sum_{\alpha_i \in \mathcal{A}_i} x_{i\alpha_i} c_{i\alpha_i}(x), \tag{2.4}$$

where $c_{i\alpha_i}(x) = \sum_{s \in \alpha_i} c_s(w_s)$ denotes the cost incurred to player i by the utilization of the set of amenities $\alpha_i \subseteq \mathcal{S}$. The resulting game is called an *atomic splittable congestion game* and is easily seen to adhere to the above framework.

2.3. Online Mirror Descent on Continuous Games under Delays We extend the general single-agent online learning delay model [39] to the multi-agent case. Algorithm 1 gives multi-agent OMD learning under delays. Several comments are in order here. First, each $h_i(\cdot)$ is a regularizer on $\mathcal{X}_i$, as defined next.

Algorithm 1 Multi-Agent Online Mirror Descent under Delays

```
Each player i chooses an arbitrary initial y_i^0 ∈ R^{d_i}.
for t = 0, 1, 2, ... do
    for i = 1, ..., N do
        x_i^t = arg max_{x_i ∈ X_i} {⟨y_i^t, x_i⟩ − h_i(x_i)}
        y_i^{t+1} = y_i^t + α^t Σ_{s ∈ G_i^t} v_i(x^s)
    end for
end for
```

DEFINITION 2.4. Let $\mathcal{D}$ be a compact and convex subset of $\mathbf{R}^m$. We say that $g : \mathcal{D} \to \mathbf{R}$ is a regularizer (with respect to some vector norm $\|\cdot\|$) if g is continuous and K-strongly convex with respect to $\|\cdot\|$: there exists some $K > 0$ such that $\forall t \in [0,1], \forall \mathbf{d}, \mathbf{d}' \in \mathcal{D}$: $g(t\mathbf{d} + (1-t)\mathbf{d}') \le tg(\mathbf{d}) + (1-t)g(\mathbf{d}') - \frac{1}{2}Kt(1-t)\|\mathbf{d}' - \mathbf{d}\|^2$.

Second, the gradient step size α^t in Algorithm 1 can be any positive and non-increasing sequence that satisfies the standard summability assumption: $\sum_{t=0}^{\infty} \alpha^t = \infty, \sum_{t=0}^{\infty} (\alpha^t)^2 < \infty$.

Third, some words on the delay model: in Algorithm 1, $\mathcal{G}_i^t$ denotes the set of rounds whose gradients become available for player i at the current round t. Denote player i's delay of the gradient at round s to be d_i^s (a positive integer), then this gradient $v_i(\mathbf{x}^s)$ will be available at round $s + d_i^s - 1$: $s \in \mathcal{G}_i^{s+d_i^s-1}$. In particular, if $d_i^s = 1, \forall s$, then this corresponds to the case where player i doesn't have any delays. Note that each player can receive out-of-order gradients: this can happen if the gradient at an earlier round has a larger delay than that of the gradient at a later round. Finally, note that the data (i.e. feedback) each player needs to collect and act upon is only of dimension d_i (i.e. dimension of each player's actions space), rather than the dimension of the joint action space.

2.4. Online Mirror Descent on Continuous Games under Both Delays and Noise We can further incorporate noisy feedback, the second feature of imperfect information, into the existing framework. Specifically, player i, instead of receiving a perfect gradient as in Line 5 of Algorithm 1, only receives a noisy gradient $\tilde{v}_i(\mathbf{X}^s) = v_i(\mathbf{X}^s) + \xi_i^{s+1}$ (see Algorithm 2), where $\{\xi^t\}_{t=1}^{\infty} = \{(\xi_i^t)_{i=1}^N\}_{t=1}^{\infty}$ is some noise process. Note that the iterates X_i^t and Y_i^t are now capitalized to make explicit the fact that they are now random variables. It is important to point out that the random variables X_i^t and Y_i^t are adapted to $y^0, \xi^1, \dots, \xi^t$, rather than $y_i^0, \xi_i^1, \dots, \xi_i^t$. We make the following assumption on the noise process:

Algorithm 2 Multi-Agent Online Mirror Descent under Delays and Noise

---

```
Each player i chooses an arbitrary initial y_i^0 ∈ R^{d_i}.
for t = 0, 1, 2, ... do
    for i = 1, ..., N do
        X_i^t = arg max_{x_i ∈ X_i} {⟨Y_i^t, x_i⟩ − h_i(x_i)}
        Y_i^{t+1} = Y_i^t + α^t Σ_{s∈G_i^t} ṽ_i(X^s) ≜ Y_i^t + α^t Σ_{s∈G_i^t} (v_i(X^s) + ξ_i^{s+1})
    end for
end for
```

Assumption 2.1. *Let $\mathcal{F}^t$ be the canonical filtration induced by the random variables $\xi^1, \dots, \xi^t$.*

1. *The noisy gradients are conditionally unbiased:* $\forall t, \mathbf{E}[\xi^{t+1} \mid \mathcal{F}^t] = 0$, *a.s.*.
2. *The noisy gradients are bounded in second moments:* $\forall t, \mathbf{E}[(\|\xi^{t+1}\|^*)^2 \mid \mathcal{F}^t] \leq V$, *a.s., for some* $V > 0$.

3. λ-Variational Stability As mentioned in the introduction, convergence to Nash equilibria does not hold in general, and can easily fail even in mixed extensions of finite games. Consequently, the existing literature has focused on obtaining such results in specific finite games (and mixed extensions thereof). In this section, we define a broad class of continuous games, called λ-variationally stable games, study its structural properties, and give several subclasses of games that belong to this class. In subsequent sections, we establish last-iterate convergence results in this general class of games.

3.1. Definition and Properties

Definition 3.1. Given a continuous game $(\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$, a set $\mathcal{C} \subset \mathcal{X}$ is called λ-variationally stable for some $\lambda \in \mathbf{R}_{++}^N$ if $\mathcal{C}$ is non-empty and it holds that (with equality if and only if $\mathbf{x} \in \mathcal{C}$):

$$\sum_{i=1}^{N} \lambda_i \langle v_i(\mathbf{x}), x_i - c_i \rangle \leq 0, \forall \mathbf{x} \in \mathcal{X}, \forall \mathbf{c} \in \mathcal{C}.$$

As the next lemma indicates, a variationally stable set has special structures:

Lemma 3.1. *In a continuous game* $(\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$*, if* $\mathcal{C}$ *is a non-empty* λ*-variationally stable set, then* $\mathcal{C}$ *is a closed and convex set of all Nash equilibria of the game.*

Proof: First we show that any element $\mathbf{x}^* \in \mathcal{C}$ is a Nash equilibrium. For any $i \in \mathcal{N}$, take any $x_i \in \mathcal{X}_i$ and any $\tau \in (0,1]$, set $\mathbf{x} \triangleq (x_1^*, \ldots, x_{i-1}^*, (1-\tau)x_i^* + \tau x_i, x_{i+1}^*, \ldots, x_N^*) = \mathbf{x}^* + \tau(x_i - x_i^*)\mathbf{e}_i$, where $\mathbf{e}_i$ is the j-th unit vector in the standard basis. By convexity of $\mathcal{X}_i$, we have $\mathbf{x} \in \mathcal{X}$. We then also have

$$\frac{d}{d\tau} u_i(x_i^* + \tau(x_i - x_i^*); \mathbf{x}_{-i}) = uv_i(\mathbf{x})(x_i - x_i^*). \tag{3.1}$$

By applying the variational stability condition to the profiles $\mathbf{x}^*$ and $\mathbf{x}$, it follows that the RHS of the above equation is strictly negative for all $\tau > 0$. In turn, this implies that $u_i(\mathbf{x}) \le u_i(\mathbf{x}^*)$, i.e. $\mathbf{x}^*$ is a Nash equilibrium.

Next, we show that $\mathcal{C}$ is closed. Take any convergent sequence $\{\mathbf{x}^j\}_{j=0}^\infty$ in $\mathcal{C}$: $\mathbf{x}^j \in \mathcal{C}, \lim_{j\to\infty} \mathbf{x}^j = \mathbf{x}^*$. Then, for any $\mathbf{x} \in \mathcal{X}$, we have $\sum_{i=1}^N \lambda_i v_i(\mathbf{x})(x_i - x_i^j) \le 0, \forall j = 0, 1, \ldots$. Therefore, by continuity, it follows that $\lim_{j\to\infty} \sum_{i=1}^N \lambda_i v_i(\mathbf{x})(x_i - x_i^j) = \sum_{i=1}^N \lambda_i v_i(\mathbf{x})(x_i - x_i^*) \le 0, \forall \mathbf{x} \in \mathcal{C}$, thereby implying $\mathbf{x}^* \in \mathcal{C}$. Since $\{\mathbf{x}^j\}_{j=0}^\infty$ is any sequence in $\mathcal{C}$, $\mathcal{C}$ contains all its limit points and is therefore closed.

To see that $\mathcal{C}$ is convex, take any $\mathbf{x}^*, \mathbf{y}^* \in \mathcal{C}$ and any $\tau \in [0,1]$. For any $\mathbf{x} \in \mathcal{X}$, we have

$$\sum_{i=1}^N \lambda_i v_i(\mathbf{x})(x_i - (\tau x_i^* - (1-\tau)y_i^*)) = \tag{3.2}$$

$$\tau \sum_{i=1}^N \lambda_i v_i(\mathbf{x})(x_i - x_i^*) + (1-\tau) \sum_{i=1}^N \lambda_i v_i(\mathbf{x})(x_i - y_i^*) \le 0, \tag{3.3}$$

thereby establishing that $\tau \mathbf{x}^* + (1-\tau)\mathbf{y}^* \in \mathcal{C}$.

Finally, to see that $\mathcal{C}$ contains all Nash equilibria of the game, assume $\mathbf{z}^* \notin \mathcal{C}$ is a Nash equilibrium. Then:

$$\sum_{i=1}^N \lambda_i v_i(\mathbf{z}^*)(x_i - z_i^*) \le 0, \forall \mathbf{x} \in \mathcal{X}. \tag{3.4}$$

Take an arbitrary $\mathbf{x}^* \in \mathcal{C}$. Since $\mathcal{C}$ is λ-variational stable and $\mathbf{z}^* \notin \mathcal{C}$, we have $\sum_{i=1}^N \lambda_i v_i(\mathbf{z}^*)(z_i^* - x_i^*) < 0$, implying that $\sum_{i=1}^N \lambda_i v_i(\mathbf{z}^*)(x_i^* - z_i^*) > 0$, which contradicts Equation 3.4. ■

In view of Lemma 3.1, we can define a general class of games based on the structure of the Nash set.

Definition 3.2. A continuous game $(\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$ is said to be λ-variationally stable if its set $\mathcal{X}^*$ of all Nash equilibria is a λ-variationally stable set.

We emphasize that in a game setting, λ-variational stability is more general than an important concept called operator monotonicity in variational analysis. Specifically, $v(\cdot)$ is called a monotone operator [40] if the following holds (with equality if and only if $\mathbf{x} = \tilde{\mathbf{x}}$):

$$\langle v(\mathbf{x}) - v(\tilde{\mathbf{x}}), \mathbf{x} - \tilde{\mathbf{x}} \rangle \triangleq \sum_{i=1}^N \langle v_i(\mathbf{x}) - v_i(\tilde{\mathbf{x}}), x_i - \tilde{x}_i \rangle \le 0, \forall \mathbf{x}, \tilde{\mathbf{x}} \in \mathcal{X}. \tag{3.5}$$

It is known that if $v(\cdot)$ is monotone (which defines an important and broad class of games referred to as strictly diagonal concave games studied in [42]), then a unique Nash equilibrium $\mathbf{x}^*$ exists and (per the property of a Nash equilibrium) satisfies $\langle v(\mathbf{x}^*), \mathbf{x}-\mathbf{x}^*\rangle \leq 0$. Consequently, by expanding Equation 3.5, it then follows that $\langle v(\mathbf{x}), \mathbf{x}-\mathbf{x}^*\rangle \leq \langle v(\mathbf{x}^*), \mathbf{x}-\mathbf{x}^*\rangle \leq 0$, where equality is achieved if and only if $\mathbf{x} = \mathbf{x}^*$. This implies that when $v(\mathbf{x})$ is a monotone operator, the singleton set of the unique Nash equilibrium is **1**-variationally stable. The converse is not true: when $v(\mathbf{x})$ is not a monotone operator, we can still have a unique Nash equilibrium[5] that is λ-variationally stable, or more generally, have a λ-variationally stable set of Nash equilibria.

3.2. A Simple Sufficient Condition for Variational Stability We now give a condition ensuring that a unique Nash equilibrium exists and is λ-variationally stable.

LEMMA 3.2. *Given a continuous game $(\mathcal{N}, \mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i, \{u_i\}_{i=1}^{N})$, where each u_i is twice continuously differentiable. For each $\mathbf{x} \in \mathcal{X}$, define the λ-weighted Hessian matrix $H^{\lambda}(\mathbf{x})$ as follows:*

$$H_{ij}^{\lambda}(\mathbf{x}) = \frac{1}{2}\lambda_i \nabla_{x_j} v_i(\mathbf{x}) + \frac{1}{2}\lambda_j (\nabla_{x_i} v_j(\mathbf{x}))^{\top}. \tag{3.6}$$

If $H^{\lambda}(\mathbf{x})$ is negative-definite for every $\mathbf{x} \in \mathcal{X}$, then $\mathbf{x}^*$ is the unique λ-variationally stable Nash equilibrium.

REMARK 3.1. It is important to note that the Hessian matrix so defined is a block matrix: each $H_{ij}^{\lambda}(\mathbf{x})$ is a $d_i \times d_j$ matrix: $H_{ij}^{\lambda}(\mathbf{x}) = \frac{1}{2}\lambda_i \nabla_{x_j} \nabla_{x_i} u_i(\mathbf{x}) + \frac{1}{2}\lambda_j (\nabla_{x_i} \nabla_{x_j} u_j(\mathbf{x}))^{\top}$. The proofs to both lemmas in this section are omitted due to space limitation.

Proof: Per Thereom 6 of [43], the assumption in the lemma implies that:

$$\sum_{i=1}^{N} \lambda_i \langle v_i(\mathbf{x}) - v_i(\tilde{\mathbf{x}}), x_i - \tilde{x}_i \rangle \leq 0, \forall \mathbf{x}, \tilde{\mathbf{x}} \in \mathcal{X}, \tag{3.7}$$

where equality holds if and only if $\mathbf{x} = \tilde{\mathbf{x}}$. Per Theorem 2 of [43], this inequality then implies that there exists a unique Nash equilibrium $\mathbf{x}^*$. Plug $\mathbf{x}^*$ into Inequality 3.7 for $\tilde{\mathbf{x}}$, we have that for any $\mathbf{x} \in \mathcal{X}$:

$$\sum_{i=1}^{N} \lambda_i \langle v_i(\mathbf{x}), x_i - x_i^* \rangle \leq \sum_{i=1}^{N} \lambda_i \langle v_i(\mathbf{x}^*), x_i - x_i^* \rangle \leq 0,$$

where the second inequality follows from the fact that $\mathbf{x}^*$ is a Nash equilibrium. Furthermore, both equality are achieved if and only if $\mathbf{x} = \mathbf{x}^*$. This implies that $\{\mathbf{x}^*\}$ is λ-variationally stable. ■

[5] When a continuous game admits a unique Nash equilibrium $\mathbf{x}^*$, we shall say for convenience that $\mathbf{x}^*$ is λ-variationally stable if $\{\mathbf{x}^*\}$ is λ-variationally stable, although it should be kept in mind that variational stability is a property on a set.

3.3. Examples of λ-Variationally Stable Games We end this section with several important classes of games that satisfy the λ-variational stability criterion. This is by no means a comprehensive list. Our goal is to illustrate that many important and well-known classes of games, some of which are already quite broad, are in fact subclasses of λ-variationally stable games:

Example 3.1 (Cournot Oligopolies). With notation as in Example 2.1 and $\lambda = \mathbf{1}$, we have

$$H^{\mathbf{1}}_{ij}(\mathbf{x}) = \frac{1}{2}\frac{\partial v_i(\mathbf{x})}{\partial x_j} + \frac{1}{2}\frac{\partial v_j(\mathbf{x})}{\partial x_i} = -b\delta_{ij} - b, \tag{3.8}$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. Consequently, we get

$$H^{\mathbf{1}}_{ij}(\mathbf{x}) = -b(\mathbf{I} + \mathbf{1}_{N\times N}). \tag{3.9}$$

Since $\mathbf{1}_{N\times N} = \mathbf{1}\mathbf{1}^\top$, we trivially conclude that $H^{\mathbf{1}}_{ij}(\mathbf{x})$ is negative-definite for all $\mathbf{x}$. Hence, by Lemma 3.2, the game admits a unique Nash equilibrium that is $\mathbf{1}$-variationally stable.

Example 3.2 (Kelly auctions). Our second example of a variationally stable game is the Kelly auction of Example 2.2. To show this, consider the weighted social welfare function

$$U(x) = \sum_{i\in\mathcal{N}} g_i^{-1} u_i(x) = \sum_{i\in\mathcal{N}}\sum_{s\in\mathcal{S}} \frac{q_s x_{is}}{c_s + \sum_{j\in\mathcal{N}} x_{js}} - \sum_{i\in\mathcal{N}}\sum_{s\in\mathcal{S}} x_{is} = \sum_{s\in\mathcal{S}} q_s \frac{\sum_{i\in\mathcal{N}} x_{is}}{c_s + \sum_{i\in\mathcal{N}} x_{is}} - \sum_{i\in\mathcal{N}}\sum_{s\in\mathcal{S}} x_{is}. \tag{3.10}$$

Since the function $q(x) = x/(c+x)$ is strictly concave for all $c > 0$, it readily follows that

a) Each payoff function u_i is strictly concave in x_i and convex in x_{-i}.

b) The welfare function $U(x)$ is concave in x.

With these properties in mind, let $\lambda_i = 1/g_i$ and note that

$$\begin{aligned}\nabla^2_{x_i} U(x) &= \sum_{k=1}^{N} \lambda_k \nabla^2_{x_i} u_k(x) = \lambda_i \nabla^2_{x_i} u_i(x) + \sum_{k\neq i} \lambda_k \nabla^2_{x_i} u_k(x) \\ &= 2H^{\lambda}_{ii}(x) - \lambda_i \nabla^2_{x_i} u_i(x) + \sum_{k\neq i} \lambda_k \nabla^2_{x_i} u_k(x)\end{aligned} \tag{3.11}$$

and

$$\begin{aligned}\nabla_{x_i}\nabla_{x_j} U(x) &= \sum_{k=1}^{N} \lambda_k \nabla_{x_i}\nabla_{x_j} u_k(x) = \lambda_i \nabla_{x_i}\nabla_{x_j} u_i(x) + \lambda_j \nabla_{x_j}\nabla_{x_i} u_j(x) + \sum_{k\neq i,j} \lambda_k \nabla_{x_i}\nabla_{x_j} u_k(x) \\ &= 2H^{\lambda}_{ij}(x) + \sum_{k\neq i,j} \lambda_k \nabla_{x_i}\nabla_{x_j} u_k(x).\end{aligned} \tag{3.12}$$

To proceed, write the terms in the expressions above as $M_{ij}(x) = \nabla_{x_i}\nabla_{x_j} U(x)$, $D_{ij}(x) = \delta_{ij}\nabla_{x_i}\nabla_{x_j} u_i(x)$, and $B^k_{ij}(x) = (1-\delta_{ik})(1-\delta_{jk})\nabla_{x_i}\nabla_{x_j} u_k(x)$, so

$$2H^{\lambda}(x) = M(x) + D(x) - \sum_{i=1}^{N} \lambda_i B^i(x). \tag{3.13}$$

Since $U(x)$ is concave in x, we will also have $M(x) \succcurlyeq 0$; by the strict concavity of $u_i(x)$ in x_i, we also get $D(x) \succ 0$; finally, since each u_i is convex in x_{-i}, it follows that $B^i(x) \preccurlyeq 0$ for all $i = 1, \ldots, N$, $x \in \mathcal{X}$ (simply note that $B^i(x)$ is the Hessian matrix of $u_i(x_i; x_{-i})$ with the variable x_i omitted). Putting all this together, we get $H^\lambda(x) \succ 0$, i.e., so the game is variationally stable by Lemma 3.2.

Example 3.3 (Concave potential games). A game $\mathcal{G} = (\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$ is called a potential game [32] if there exists a potential function $V : \mathcal{X} \to \mathbf{R}$ such that $u_i(x_i, \mathbf{x}_{-i}) - u_i(\tilde{x}_i, \mathbf{x}_{-i}) = V(x_i, \mathbf{x}_{-i}) - V_i(\tilde{x}_i, \mathbf{x}_{-i}), \forall i \in \mathcal{N}, \forall \mathbf{x} \in \mathcal{X}, \forall \tilde{x}_i \in \mathcal{X}_i$. A potential game is further called a concave potential game if the potential function $V(\cdot)$ is concave. Note that in a concave potential game, we have

$$
\begin{aligned}
H^\lambda_{ij}(\mathbf{x}) &= \frac{1}{2}\lambda_i \nabla_{x_j} v_i(\mathbf{x}) + \frac{1}{2}\lambda_j (\nabla_{x_i} v_j(\mathbf{x}))^\top \\
&= \frac{1}{2}\lambda_i \nabla_{x_j} \nabla_{x_i} V(\mathbf{x}) + \frac{1}{2}\lambda_j (\nabla_{x_i} \nabla_{x_j} V(\mathbf{x}))^\top. \qquad (3.14)
\end{aligned}
$$

Setting $\lambda = \mathbf{1}$, we obtain $H^{\mathbf{1}}(\mathbf{x}) = \nabla^2 V$, which is negative semi-definite when V is concave. This implies that in a concave potential game, $\mathcal{C} = \arg\max_{\mathbf{x} \in \mathcal{X}} V(\mathbf{x})$ is $\mathbf{1}$-variationally stable per Lemma 3.2.

Example 3.4 (Diagonally Strict Concave Games). This is a class of continuous games introduced by Rosen [42]. Specifically, a diagonally strict concave games satisfies the following two conditions:

1) It is a concave game: each $u_i(x_i, x_{-i})$ is individually concave in x_i.

2) $\sum_{i=1}^N \lambda_i \langle v_i(\mathbf{x}) - v_i(\tilde{\mathbf{x}}), x_i - \tilde{x}_i \rangle \le 0, \forall \mathbf{x}, \tilde{\mathbf{x}} \in \mathcal{X}$, for some positive scalars $\lambda_1, \lambda_2, \ldots, \lambda_N$.

Note that strictly diagonally concave games strictly include monotone games, which are concave games satisfying $\langle v(\mathbf{x}) - v(\tilde{\mathbf{x}}), \mathbf{x} - \tilde{\mathbf{x}} \rangle \le 0, \forall \mathbf{x}, \tilde{\mathbf{x}} \in \mathcal{X}$. Namely, a monotone game is a concave game where the joint gradient is a monotone operator.

Example 3.5 (Pseudo-Monotone Games). The recent work [51] relaxed the monotone operator assumption and introduced a broader class of games called pseudo-monotone games. A pseudo-monotone game is a concave game satisfying, $\forall \mathbf{x}, \tilde{\mathbf{x}} \in \mathcal{X}$: If $\langle v(\tilde{\mathbf{x}}), \mathbf{x} - \tilde{\mathbf{x}} \rangle \le 0$, then $\langle v(\mathbf{x}), \mathbf{x} - \tilde{\mathbf{x}} \rangle \le 0$.

To see that a pseudo-monotone game is λ-variationally stable, we start by noting an important characterization of a Nash equilibrium that is well-known in the literature (e.g. [14]): $\mathbf{x}^*$ is a Nash equilibrium of a concave game if and only if for every $i \in \mathcal{N}$, and every $\mathbf{x} \in \mathcal{X}$, $\langle v_i(\mathbf{x}^*), (x_i - x_i^*) \rangle \le 0$. Note that the "only if" direction holds even in a general continuous game.

Consequently, at a Nash equilibrium $\mathbf{x}^*$, we have $\langle v(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle \le 0$. Since a pseudo-monotone game is a concave game, its set of Nash equilibria is non-empty, and hence at least one Nash equilibrium $\mathbf{x}^*$ exists. The definition of a pseudo-monotone game then immediately implies $\langle v(\mathbf{x}), \mathbf{x} - \mathbf{x}^* \rangle \le 0$, thereby establishing the conclusion.

Albeit general, not every continuous game is λ-variationally stable. To see this, simply note that at the generality of an arbitrary continuous game, a Nash equilibrium may fail to exist. And, even if a game admits an equilibrium, this need not be necessarily stable: the class of atomic splittable congestion games of Example 2.3 satisfies Rosen's diagonal strict concavity condition when the set of amenities corresponds to a network with parallel links [36], but not necessarily otherwise.

4. Multi-Agent Online Mirror Descent under Imperfect Information We tackle the convergence problem of multi-agent OMD under imperfect information: both delays and noise. We start by defining (Section 4.1) an important divergence measure, λ-Fenchel coupling. We then establish its useful properties that play an indispensable role throughout. Building on this tool, we establish (Section 4.2) last-iterate convergence of multi-agent OMD to Nash equilibria under synchronous and bounded delays and extend (Section 4.3) the result to almost sure convergence to Nash equilibria under both delays and noise.

4.1. λ-Fenchel Coupling

Definition 4.1. Let $(\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$ be a continuous game. For each player i, let $h_i : \mathcal{X}_i \to \mathbf{R}$ be a regularizer with respect to the norm $\|\cdot\|_i$ that is K_i-strongly convex.

1. The convex conjugate function $h_i^* : \mathbb{R}^{d_i} \to \mathbf{R}$ of h_i is defined as: $h_i^*(y_i) = \max_{x_i \in \mathcal{X}_i}\{\langle x_i, y_i\rangle - h_i(x_i)\}$.
2. The mirror map $C_i : \mathbb{R}^{d_i} \to \mathcal{X}_i$ associated with regularizer h_i for player i is defined as: $C_i(y_i) = \arg\max_{x_i \in \mathcal{X}_i}\{\langle x_i, y_i\rangle - h_i(x_i)\}$. Further, define $C : \mathbf{R}^{\sum_{i=1}^N d_i} \to \mathcal{X}$, with $C(\mathbf{y}) = (C_1(y_1), \dots, C_N(y_N))$.
3. For a $\lambda \in \mathbb{R}_{++}^N$, the λ-Fenchel coupling $F^\lambda : \mathcal{X} \times \mathbb{R}^{\sum_{i=1}^N d_i} \to \mathbf{R}$ is defined as: $F^\lambda(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^N \lambda_i(h_i(x_i) - \langle x_i, y_i\rangle + h_i^*(y_i))$.

Note that although the domain of h_i is $\mathcal{X}_i \subset \mathbb{R}^{d_i}$, the domain of its conjugate (gradient space) h_i^* is $\mathbb{R}^{d_i}$. Since it is not directly relevant to the results presented in this work, we mention in passing that λ-Fenchel coupling generalizes the well-known Bregman divergence function: when $\lambda = \mathbf{1}$, λ-Fenchel coupling coincides in value with Bregman divergence if $\mathbf{x}$ is an inteior point in $\mathcal{X}$ (but not necessarily otherwise). The two key properties of λ-Fenchel coupling are:

Lemma 4.1. *For each $i \in \{1, \dots, N\}$, let $h_i : \mathcal{X}_i \to \mathbb{R}$ be a regularizer with respect to the norm $\|\cdot\|_i$ that is K_i-strongly convex and let $\lambda \in \mathbb{R}_{++}^N$. Then $\forall \mathbf{x} \in \mathcal{X}, \forall \tilde{\mathbf{y}}, \mathbf{y} \in \mathbb{R}^{\sum_{i=1}^N d_i}$:*

1. $F^\lambda(\mathbf{x}, \mathbf{y}) \geq \frac{1}{2}\sum_{i=1}^N K_i \lambda_i \|C_i(y_i) - x_i\|_i^2 \geq \frac{1}{2}(\min_i K_i\lambda_i)\sum_{i=1}^N \|C_i(y_i) - x_i\|_i^2$.
2. $F^\lambda(\mathbf{x}, \tilde{\mathbf{y}}) \leq F^\lambda(\mathbf{x}, \mathbf{y}) + \sum_{i=1}^N \lambda_i\langle \tilde{y}_i - y_i, C_i(y_i) - x_i\rangle + \frac{1}{2}(\max_i \frac{\lambda_i}{K_i})\sum_{i=1}^N (\|\tilde{y}_i - y_i\|_i^*)^2$, *where $\|\cdot\|_i^*$ is the dual norm of $\|\cdot\|_i$ (i.e. $\|y_i\|_i^* = \max_{\|x_i\|_i \leq 1}\langle x_i, y_i\rangle$).*

Remark 4.1. Since each space $\mathcal{X}_i$ is endowed with norm $\|\cdot\|_i$, we can define the induced aggregate norm $\|\cdot\|$ on the joint space $\mathcal{X}$ as follows: $\|\mathbf{x}\| = \sum_{i=1}^N \|x_i\|_i$. Similarly for the aggregate dual norm:

$\|\mathbf{y}\|^* = \sum_{i=1}^N \|y_i\|_i^*$. Finally, note that Part 1 of Lemma 4.1 implies that $F^\lambda(\mathbf{x},\mathbf{y}^t) \to 0 \Longrightarrow C(\mathbf{y}^t) \to \mathbf{x}$ as $t \to \infty$. We further assume for the rest of the paper that the mirror maps are regular in the following (very weak) sense: $C(\mathbf{y}^t) \to \mathbf{x} \Longrightarrow F^\lambda(\mathbf{x},\mathbf{y}^t) \to 0$ as $t \to \infty$. Unless one aims for pathological cases, mirror maps induced by typical regularizers are regular: examples include the commonly used Euclidean and entropic regularizers.

4.2. Convergence of Multi-Agent OMD to Nash under Synchronous and Bounded Delays

ASSUMPTION 4.1. *The delays are assumed to be:*

1. *Synchronous:* $\mathcal{G}_i^t = \mathcal{G}_j^t, \forall i, j, \forall t$.
2. *Bounded:* $d_i^t \le D, \forall i, \forall t$ *(for some positive integer* D*).*

THEOREM 4.1. *Let* $(\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$ *be a continuous game that admits* $\mathbf{x}^*$ *as the unique Nash equilibrium that is* λ*-variationally stable. Under Assumption 4.1,* $\mathbf{x}^t \to \mathbf{x}^*$*, where* $\mathbf{x}^t$ *is given in Algorithm 1.*

REMARK 4.2. The proof is somewhat long and involved. To aid the understanding and enhance the intuition, we break it down into four main steps, the details are omitted due to space limitation.

1. Since the delays are synchronous, we denote by $\mathcal{G}^t$ the common set and d^t the common delay at round t. The gradient update in OMD under delays can then be written as:

$$y_i^{t+1} = y_i^t + \alpha^t \sum_{s\in\mathcal{G}^t} v_i(\mathbf{x}^s) = y_i^t + \alpha^t \left\{ |\mathcal{G}^t| v_i(\mathbf{x}^t) + \sum_{s\in\mathcal{G}^t} \{v_i(\mathbf{x}^s) - v_i(\mathbf{x}^t)\} \right\}. \tag{4.1}$$

   Define $b_i^t = \sum_{s\in\mathcal{G}^t}\{v_i(\mathbf{x}^s) - v_i(\mathbf{x}^t)\}$. Using the bounded delay assumption, we establish, after a chain of inequalities, that $\|b_i^t\|_i^* \le \frac{LD^3V_{\max}}{K}\alpha^{t-D+1}$ where $K \triangleq \min_i K_i$, L is the Lipschitz constant for $v(\cdot)$, $V_{\max} \triangleq \max_{\mathbf{x}\in\mathcal{X}} \|v(\mathbf{x})\|^*$ and D is the bound on delays. This implies that $\lim_{t\to\infty} \|b_i^t\|_i^* = 0$ for each i.
2. Define $\mathbf{b}^t = (b_1^t, \ldots, b_N^t)$ and Claim 1 yields that $\lim_{t\to\infty} \mathbf{b}^t = \mathbf{0}$. Per Equation 4.1 in Claim 1, we can then write the joint OMD update of all players under delays as follows:

$$\mathbf{x}^t = C(\mathbf{y}^t), \tag{4.2}$$

$$\mathbf{y}^{t+1} = \mathbf{y}^t + \alpha^t \left\{ |\mathcal{G}^t| v(\mathbf{x}^t) + \mathbf{b}^t \right\}. \tag{4.3}$$

   Let $B(\mathbf{x}^*, \epsilon) \triangleq \{\mathbf{x} \in \mathcal{X} \mid \|\mathbf{x} - \mathbf{x}^*\| < \epsilon\}$ be the open ball centered around $\mathbf{x}^*$ with radius ϵ. Then, using λ-Fenchel coupling as an energy function and leveraging the handle on $\mathbf{b}^t$ given by Claim 1, we establish that for any $\epsilon > 0$, the iterate $\mathbf{x}^t$ will visit $B(\mathbf{x}^*, \epsilon)$ infinitely often, no matter what the initial point $\mathbf{x}^0$ is. This is accomplished by applying Statement 2 of Lemma 3.1 to expand Equation C.12, and subsequently by showing that, through a chain of inequalities, unless the above statement is true, the λ-Fenchel coupling, always non-negative, would go to $-\infty$.

3. Fix any $\delta > 0$ and consider the set $\tilde{B}(\mathbf{x}^*,\delta) \triangleq \{C(\mathbf{y}) \mid F^\lambda(\mathbf{x}^*,\mathbf{y}) < \delta\}$. In other words, $\tilde{B}(\mathbf{x}^*,\delta)$ is some "neighborhood" of $\mathbf{x}^*$, which contains every $\mathbf{x}$ that is an image of some $\mathbf{y}$ (under the mirror map $C(\cdot)$) that is within δ distance of $\mathbf{x}^*$ under the λ-Fenchel coupling "metric". Although $F^\lambda(\mathbf{x}^*,\mathbf{y})$ is not a metric, $\tilde{B}(\mathbf{x}^*,\delta)$ contains an open ball within it: for any $\delta > 0$, $\exists\epsilon(\delta) > 0$ such that: $B(\mathbf{x}^*,\epsilon) \subset \tilde{B}(\mathbf{x}^*,\delta)$.
4. For any "neighborhood" $\tilde{B}(\mathbf{x}^*,\delta)$, after long enough rounds, if $\mathbf{x}^t$ ever enters $\tilde{B}(\mathbf{x}^*,\delta)$, it will be trapped inside $\tilde{B}(\mathbf{x}^*,\delta)$ thereafter. Mathematically, the claim is that for any $\delta > 0$, $\exists T(\delta)$, such that for any $t \geq T(\delta)$, if $\mathbf{x}^t \in \tilde{B}(\mathbf{x}^*,\delta)$, then $\mathbf{x}^{\tilde{t}} \in \tilde{B}(\mathbf{x}^*,\delta), \forall \tilde{t} \geq t$. This is done by considering two possibilities:
   (a) Possibility 1: $\mathbf{x}^t \in B(\mathbf{x}^*,\epsilon(\frac{\delta}{2}))$;
   (b) Possibility 2: $\mathbf{x}^t \in \tilde{B}(\mathbf{x}^*,\delta) - B(\mathbf{x}^*,\epsilon(\frac{\delta}{2}))$;
   and using a different argument for each possibility to establish that $\mathbf{x}^{t+1} \in \tilde{B}(\mathbf{x}^*,\delta)$ in both cases.

It is now time to put all four elements above together. The significance of Claim 3 is that, since the iterate $\mathbf{x}^t$ will enter $B(\mathbf{x}^*,\epsilon)$ infinitely often (per Claim 2), $\mathbf{x}^t$ must enter $\tilde{B}(\mathbf{x}^*,\delta)$ infinitely often. It therefore follows that, per Claim 4, starting from iteration t, $\mathbf{x}^t$ will remain in $\tilde{B}(\mathbf{x}^*,\delta)$. Since this is true for any $\delta > 0$, we have $F^\lambda(\mathbf{x}^*,\mathbf{y}^t) \to 0$ as $t \to \infty$. Per Statement 1 in Lemma 4.1, this leads to that $\|C(\mathbf{y}^t) - \mathbf{x}^*\| \to 0$ as $t \to \infty$, thereby establishing that $\mathbf{x}^t = C(\mathbf{y}^t) \to \mathbf{x}^*$ as $t \to 0$.

In fact, the result generalizes straightforwardly to the set of Nash equilibria case, where proof is line-by-line identical, provided we redefine, in a standard way, every quantity that measures the distance between two points to the corresponding quantity that measures the distance between a point and a set (by taking the infimum over the distances between the point and a point in that set). We directly state the result below.

Theorem 4.2. *Let $(\mathcal{N},\mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$ be a continuous game that admits $\mathcal{X}^*$ as the λ-variationally stable set of Nash equilibria. Under Assumption 4.1,* $\lim_{t\to\infty} dist(\mathbf{x}^t,\mathcal{X}^*) = 0$*, where $dist(\cdot,\cdot)$ is the standard point-to-set distance function induced by the norm* $\|\cdot\|$.

4.3. Almost Sure Convergence of Multi-Agent OMD to Nash under Delays and Noise The additional feature of noisy gradients (Algorithm 2) introduces significant new challenges for establishing the convergence result. For instance, Claim 4 in Remark 4.2, a crucial step in establishing Theorem 4.1, no longer holds because, *a priori*, a single noisy gradient can potentially perturb the iterates out of the open neighborhood around the Nash set. Consequently, we take a different approach here: cast the OMD dynamics in a differential equation approximation framework and connect the iterates from the OMD algorithm to the solution from the differential equation. We begin with minimal mathematical preliminaries (see [4]).

Definition 4.2. A semiflow ϕ on a metric space (M,d) is a continuous map $\phi : \mathbf{R}_+ \times M \to M$: $(t,x) \to \phi_t(x)$, such that the semi-group properties hold: ϕ_0 = identity, $\phi_{t+s} = \phi_t \circ \phi_s$ for all $(t,s) \in \mathbf{R}_+ \times \mathbf{R}_+$.

Remark 4.3. A standard way to induce a semiflow is via an ordinary differential equation (ODE). Specifically, as mentioned in [4], if $F : \mathbf{R}^m \to \mathbf{R}^m$ is a continuous function and if the following ODE has a unique solution trajectory for each initial point $\tilde{x} \in \mathbf{R}^m$:

$$\begin{aligned} \frac{dx}{dt} &= F(x), \\ x(0) &= \tilde{x}, \end{aligned}$$

then $\phi_t(\tilde{x})$ defined by the solution trajectory $x(t) \in \mathbf{R}^m$ as follows is a semiflow: $\phi_t(\tilde{x}) \triangleq x(t)$ with $x(0) = \tilde{x}$. We say ϕ defined in this way is the semiflow induced by the corresponding ODE.

Definition 4.3. Let ϕ be a semiflow on the metric space (M, d). A continuous function $s : \mathbf{R}_+ \to M$ is an asymptotic pseudotrajectory for ϕ if for every $T > 0$, the following holds:

$$\lim_{t\to\infty} \sup_{0\le h\le T} d(s(t+h), \phi_h(s(t))) = 0. \tag{4.4}$$

Remark 4.4. By definition, that s is an asymptotic pseudotrajectory for ϕ means s and ϕ are very close for sufficiently large t. Specifically, for each $T > 0$, there is a large enough t_0, such that $\forall t > t_0$, the curve $s(t+h)$ approximates the trajectory $\phi_h(s(t))$ on the interval $h \in [0, T]$ with any pre-specified degree of accuracy.

We next state two martingale convergence theorems (adapted from [18]) that shall be useful: law of large number for martingales and Doob's martingale convergence theorem in order.

Theorem 4.3 (**Law of large number for martingales**). *Let $S^t = \sum_{k=0}^{t} X^k$ be a martingale adapted to the filtration $\mathcal{S}^t$ and $\{u^t\}_{t=0}^{\infty}$ be a increasing sequence of positive numbers with $\lim_{t\to\infty} u^t = \infty$. If $\exists p \in [1, 2]$ such that $\sum_{t=0}^{\infty} \frac{\mathbf{E}[|X^{t+1}|^p | \mathcal{S}^t]}{(u^t)^p} < \infty$, a.s. , then:*

$$\lim_{t\to\infty} \frac{S^t}{u^t} = 0, \ a.s..$$

Theorem 4.4 (**Doob's martingale convergence**). *Let S^t be a submartingale adapted to the filtration $\mathcal{S}^t$, where $t = 0, 1, 2, \ldots$. If S^t is l^1-bounded: $\sup_{t\ge0} \mathbf{E}[|S^t|] < \infty$, then S^t converges almost surely to a random variable S with $\mathbf{E}[|S|] < \infty$.*

With the above notation in place, we are now ready to state and establish the convergence result.

Theorem 4.5. *Let $(\mathcal{N}, \mathcal{X} = \prod_{i=1}^{N} \mathcal{X}_i, \{u_i\}_{i=1}^{N})$ be a continuous game that admits $\mathcal{X}^*$ as the λ-variationally stable set of Nash equilibria. Under Assumptions 4.1 and 2.1, $\mathbf{X}^t$ in Algorithm 2 converges to $\mathcal{X}^*$ a.s..*

Remark 4.5. The proof builds on the delays-only case, but is more involved and calls for a different proof strategy. We outline the main steps (again it suffices to consider the single Nash equilibrium case):

1. Using the same notation as Remark 4.2, OMD's joint gradient update in Algorithm 2 can be rewritten as:

$$Y^{t+1} = Y^t + \alpha^t \sum_{s\in\mathcal{G}^t} (v(X^s) + \xi^{s+1}) = Y^t + \alpha^t \left\{ |\mathcal{G}^t| v(X^t) + \sum_{s\in\mathcal{G}^t} \xi^{s+1} + B^t \right\}, \tag{4.5}$$

where $B_i^t = \sum_{s\in\mathcal{G}^t} \{v_i(X^s) - v_i(X^t)\}$ and $B^t = (B_1^t, \ldots, B_N^t)$. Similar to Claim 1 in Remark 4.2, we establish that $B^t \to \mathbf{0}$ almost surely as $t \to \infty$. Similar to Claim 2 in Remark 4.2, using λ-Fenchel coupling as the energy function, leveraging $B^t \to \mathbf{0}$, a.s., and using both the law of large number for martingales and Doob's martingale convergence theorem [18], we establish that any open ball around $\mathbf{x}^*$ is recurrent. Mathematically, for any $\epsilon > 0$ and any initial point $\mathbf{x}^0$, the iterate $\mathbf{X}^t$ visit $B(\mathbf{x}^*, \epsilon)$ infinitely often almost surely. Note that per Claim 3 in Remark 4.2, this then implies that for any $\delta > 0$, $\mathbf{X}^t$ must almost surely visit $\tilde{B}(\mathbf{x}^*, \delta)$ infinitely often, irrespective of the initial point.

2. Next we consider the ordinary differential equation (ODE) approximation of OMD as follows:

$$\dot{\mathbf{y}} = v(\mathbf{x}),$$
$$\mathbf{x} = C(\mathbf{y}).$$

This ODE is a mean approximation of Equation 4.5, where both the martingale noise term $\sum_{s\in\mathcal{G}^t} \xi^{s+1}$ and the negligible B^t term are removed, and where $|\mathcal{G}^t|$ is absorbed into the step-size. This can be written as $\dot{\mathbf{y}} = v(C(\mathbf{y}))$, which can be verified to admit a unique solution trajectory for any initial condition. Consequently, per Remark 4.3, this solution induces a semiflow[6], which we denote $\phi_t(\mathbf{y})$: it is the state at time t given it starts at $\mathbf{y}$. Note that we have used $\mathbf{y}$ as the initial point (as opposed to $\mathbf{y}^0$) to indicate that the semiflow representing the solution trajectory should be viewed as a function of the initial point $\mathbf{y}$.

3. We now relate the iterates generated by OMD to the above ODE's solution. First, we connect linearly the OMD iterates $Y^0, Y^1, Y^2, \ldots, Y^k, \ldots$ at times $0, \alpha^0, |\mathcal{G}^0|\alpha^0 + |\mathcal{G}^1|\alpha^1, \ldots, \sum_{l=0}^{k-1} |\mathcal{G}^l|\alpha^l, \ldots$ respectively to form a continuous, piecewise affine curve. Namely:

$$Y(t) = Y^k + (t - \sum_{l=0}^{k-1} \alpha^l) \frac{Y^{k+1} - Y^k}{\alpha^k}, \text{for } t \in [\sum_{l=0}^{k-1} |\mathcal{G}^l|\alpha^l, \sum_{l=0}^{k} |\mathcal{G}^l|\alpha^l), k = 0, 1, \ldots,$$

where we adopt the usual convention that $\sum_{l=0}^{-1} \alpha^l = 0$. We then show that $Y(t)$ (a random trajectory) is almost surely an asymptotic pseudotrajectory of the semi-flow ϕ induced by the above ODE. Mathematically, we establish that $\forall T > 0, \lim_{t\to\infty} \sup_{0\le h\le T} \|Y(t+h), \phi_h(Y(t))\|^* = 0$, a.s..

4. Having characterized the relation between the OMD trajectory (affine interpolation of the discrete OMD iterates) and the ODE trajectory (the semi-flow), we now turn to studying the latter (the semiflow given by the ODE trajectory). A desirable property of $\phi_t(\mathbf{y})$ is that the distance $F^\lambda(\mathbf{x}^*, \phi_t(\mathbf{y}))$ between the primal

[6] A crucial point to note is that since C may not be invertible, there may not exist a unique solution for $\mathbf{x}(t)$.

variable $\mathbf{x}^*$ and the dual variable $\phi_t(\mathbf{y})$) (as measured by the Lyapunov function λ-Fenchel coupling) can never increase as a function of t. We refer to this as the monotonicity property of λ-Fenchel coupling under the ODE trajectory, to be contrasted to the discrete-time dynamics, where such monotonicity is absent (even when perfect information on the gradient is available). More formally, we show that $\forall \mathbf{y}, \forall 0 \le s \le t$,

$$F^{\lambda}(\mathbf{x}^*, \phi_s(\mathbf{y})) \ge F^{\lambda}(\mathbf{x}^*, \phi_t(\mathbf{y})). \tag{4.6}$$

5. Continuing on the previous point, not only the distance $F^{\lambda}(\mathbf{x}^*, \phi_t(\mathbf{y}))$ can never increase as t increases, but also, provided that $\phi_t(\mathbf{y})$ is not too close to $\mathbf{x}^*$ (under the λ-Fenchel coupling divergence measure), $F^{\lambda}(\mathbf{x}^*, \phi_t(\mathbf{y}))$ will decrease no slower than linearly. This suggests that either $\phi_t(\mathbf{y})$ is already close to $\mathbf{x}^*$ (and hence $x(t) = C(\phi_t(\mathbf{y}))$ is close to $\mathbf{x}^*$), or their distance will be decreased by a meaningful amount in (at least) the ensuing short time-frame. We formalize this discussion into the following mathematical claim: $\forall \epsilon > 0, \forall \mathbf{y}, \exists s > 0$, such that:

$$F^{\lambda}(\mathbf{x}^*, \phi_s(\mathbf{y})) \le \max\{\frac{\epsilon}{2}, F^{\lambda}(\mathbf{x}^*, \mathbf{y}) - \frac{\epsilon}{2}\}. \tag{4.7}$$

6. Now consider an arbitrary fixed horizon T. If at time t, $F^{\lambda}(\mathbf{x}^*, \phi_0(Y(t)))$ is small, then by the monotonicity property in Claim 4, $F^{\lambda}(\mathbf{x}^*, \phi_h(Y(t)))$ will remain small on the entire interval $h \in [0,T]$. Since $Y(t)$ is an asymptotic pseudotrajectory of ϕ (Claim 3), $Y(t+h)$ and $\phi_h(Y(t))$ should be very close for $h \in [0,T]$, at least for t large enough. This means that $F^{\lambda}(\mathbf{x}^*, Y(t+h))$ should also be small on the entire interval $h \in [0,T]$, if λ-Fenchel coupling has a regular enough structure. It turns out that this is indeed the case. This can be made precise as follows: $\forall \epsilon, T > 0, \exists \tau(\epsilon, T) > 0$ such that $\forall t \ge \tau, \forall h \in [0,T]$:

$$F^{\lambda}(\mathbf{x}^*, Y(t+h)) < F^{\lambda}(\mathbf{x}^*, \phi_h(Y(t))) + \frac{\epsilon}{2}, \text{ a.s.}. \tag{4.8}$$

7. Finally, we are ready to put the above pieces together. Claim 6 gives us a way to control the amount by which the two λ-Fenchel coupling functions differ on the interval $[0,T]$. Claim 4 and Claim 5 together allow us to extend such control over successive intervals $[T,2T), [2T,3T), \ldots$, thereby establishing that, at least for t large enough, if $F^{\lambda}(\mathbf{x}^*, Y(t))$ is small, then $F^{\lambda}(\mathbf{x}^*, Y(t+h))$ will remains small $\forall h > 0$. As it turns out, this means that after long enough time, if $\mathbf{X}^t$ ever visits $\tilde{B}(\mathbf{x}^*, \epsilon)$, it will (almost surely) be forever trapped inside the neighborhood twice that size (i.e. $\tilde{B}(\mathbf{x}^*, 2\epsilon)$). Since Claim 1 ensures that $\mathbf{X}^t$ visits $\tilde{B}(\mathbf{x}^*, \epsilon)$ infinitively often (almost surely), the hypothesis is guaranteed to be true. Consequently, this leads to the following formal claim: $\forall \epsilon > 0, \exists \tau_0$ (a positive integer), such that:

$$F^{\lambda}(\mathbf{x}^*, Y(\tau_0 + h)) < \epsilon, \forall h \in [0, \infty), \text{ a.s.}. \tag{4.9}$$

To conclude, Equation (4.9) in Claim 7 implies that $F^{\lambda}(\mathbf{x}^*, Y^t) \to 0,$ a.s. as $t \to \infty$, where the OMD iterates Y^t are values at integer time points of the affine trajectory $Y(\tau)$. Per Statement 1 in Lemma 4.1, this leads to that $\|C(Y^t) - \mathbf{x}^*\| \to 0,$ a.s. as $t \to \infty$, thereby establishing that $X^t = C(Y^t) \to \mathbf{x}^*,$ a.s. as $t \to 0$.

5. Multi-Agent Delayed Mirror Descent under Imperfect Information The synchronous and bounded delay assumption in Assumption 4.1 is fairly strong[7]. In this section, by a simple modification of OMD, we propose a new learning algorithm called Delayed Mirror Descent (DMD), that allows the result to be generalized to cases with arbitrary asynchronous delays among players as well as unbounded delay growth.

5.1. Delayed Mirror Descent on Continuous Games The main idea for the modification is that when player i doesn't receive any gradient on round t, instead of not doing any gradient updates as in OMD, it uses the most recent set of gradients to perform updates. More formally, define the most recent information set[8]:

$$\tilde{\mathcal{G}}_i^t = \begin{cases} \mathcal{G}_i^t, \textbf{ if } \mathcal{G}_i^t \neq \emptyset \\ \tilde{\mathcal{G}}_i^{t-1}, \textbf{ if } \mathcal{G}_i^t = \emptyset. \end{cases}$$

With this definition, DMD is given in Algorithm 3: note that $\tilde{\mathcal{G}}_i^t$ is always non-empty by definition. As in OMD, the information each player needs to collect and act upon in DMD is of dimension d_i. Note also that if there is no delay, then DMD recovers OMD in Algorithm 1.

Algorithm 3 Multi-Agent Delayed Mirror Descent

1: Each player i chooses an arbitrary initial y_i^0.
2: **for** $t = 0, 1, 2, \ldots$ **do**
3:     **for** $i = 1, \ldots, N$ **do**
4:         $x_i^t = \arg\max_{x_i \in \mathcal{X}_i} \{\langle y_i^t, x_i \rangle - h_i(x_i)\}$
5:         $y_i^{t+1} = y_i^t + \frac{\alpha^t}{|\tilde{\mathcal{G}}_i^t|} \sum_{s \in \tilde{\mathcal{G}}_i^t} v_i(\mathbf{x}^s)$
6:     **end for**
7: **end for**

5.2. Main Delay Assumption Here we only make the following assumption on the delays:

Assumption 5.1. *For each player i,* $\lim_{t\to\infty} \sum_{s=\min \tilde{\mathcal{G}}_i^t}^{t} \alpha^s = 0$.

This assumption essentially requires delays to not grow too fast. In particular, delays can be arbitrarily asynchronous among agents. For concreteness, we next give two more explicit delay conditions that satisfy the main delay assumption. As made formal by the following lemma, if the delays are bounded (but possibly fully asynchronous), then Assumption 5.1 is satisfied. Furthermore, by appropriately choosing the sequence α^t, Assumption 5.1 can accommodate delays that grow **unbounded at a super-linear** rate.

[7] An important reason that vanilla OMD may fail to converge in the absence of such strong delays assumptions is that players do not take any actions when no gradient is received. This can potentially lead the joint OMD update off the convergence track when different players receive gradients at arbitrarily different times.

[8] There may not be any gradient information in the first few rounds due to delays. Without loss of generality, we can always start at the first round when there is non-empty gradient information, or equivalently, assume that some gradient is available at $t = 0$.

Lemma 5.1. *Let $\{d_i^s\}_{s=1}^{\infty}$ be the delay sequence for player i.*

1. *If each player i's delay is bounded (i.e. $\exists d \in \mathbb{Z}, d_i^s \le d, \forall s$), then Assumption 5.1 is satisfied for any positive, non-increasing, not-summable-but-square-summable sequence $\{\alpha^t\}$.*
2. *There exists a positive, non-increasing, not-summable-but-square-summable sequence (e.g. $\alpha^t = \frac{1}{t \log t \log\log t}$) such that if $d_i^s = O(s \log s), \forall i$, then Assumption 5.1 is satisfied.*

Remark 5.1. The proof, which is given in the appendix, indicates that one can also easily obtain slightly larger delay growth rates: $O(t \log t \log\log t), O(t \log t \log\log t \log\log\log t)$ and so on, by choosing the corresponding step size sequences. However, we believe such rates are only marginally larger and do not focus on them here. Further, it is conceivable that one can identify meaningfully larger delay growth rates that still satisfy Assumption 5.1, particularly under more restrictions on the degree of delay asynchrony among the players. We leave that for future work.

5.3. Convergence of Multi-Agent DMD to Nash under Asynchronous and Unbounded Delays

Theorem 5.1. *Let $(\mathcal{N}, \mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$ be a continuous game that admits $\mathcal{X}^*$ as the λ-variationally stable set of Nash equilibria. Under Assumption 5.1, $\lim_{t\to\infty} \mathrm{dist}(\mathbf{x}^t, \mathcal{X}^*) = 0$, where $\mathbf{x}^t$ is given in Algorithm 3.*

Remark 5.2. The proof uses a similar framework as the one in Remark 4.2, although the details are different. Building on the notation and arguments given in Remark 4.2, we again outline main ingredients. Once again, it suffices to consider the single Nash equilibrium case. Details are given in the appendix.

1. The individual gradient update in DMD can be rewritten as:

$$y_i^{t+1} = y_i^t + \frac{\alpha^t}{|\tilde{\mathcal{G}}_i^t|} \sum_{s \in \tilde{\mathcal{G}}_i^t} v_i(\mathbf{x}^s) = y_i^t + \alpha^t v_i(\mathbf{x}^t) + \alpha^t \sum_{s \in \tilde{\mathcal{G}}_i^t} \frac{v_i(\mathbf{x}^s) - v_i(\mathbf{x}^t)}{|\tilde{\mathcal{G}}_i^t|}.$$

   By defining: $b_i^t = \sum_{s \in \tilde{\mathcal{G}}_i^t} \frac{v_i(\mathbf{x}^s) - v_i(\mathbf{x}^t)}{|\tilde{\mathcal{G}}_i^t|}$, we can write player i's gradient update as:

$$y_i^{t+1} = y_i^t + \alpha^t (v_i(\mathbf{x}^t) + b_i^t).$$

   By bounding b_i^t's magnitude using the delay sequence, Assumption 5.1 allows us to establish that b_i^t has negligible impact over time. Mathematically, the claim is that $\lim_{t\to\infty} \|b_i^t\|_i^* = 0$.
2. The joint DMD update can be written as:

$$\mathbf{x}^t = C(\mathbf{y}^t), \tag{5.1}$$

$$\mathbf{y}^{t+1} = \mathbf{y}^t + \alpha^t (v(\mathbf{x}^t) + \mathbf{b}^t). \tag{5.2}$$

   Here again using λ-Fenchel coupling as the energy function and leveraging the handle on $\mathbf{b}^t$ given by Claim 1, we show that for any $\epsilon > 0$ the iterate $\mathbf{x}^t$ will visit $B(\mathbf{x}^*, \epsilon)$ infinitely often. Furthermore, per Claim 3 in Remark 4.2, $B(\mathbf{x}^*, \epsilon) \subset \tilde{B}(\mathbf{x}^*, \delta)$. This implies that $\mathbf{x}^t$ must enter $\tilde{B}(\mathbf{x}^*, \delta)$ infinitely often.

3. Again using λ-Fenchel coupling, we show that under multi-agent DMD, for any "neighborhood" $\tilde{B}(\mathbf{x}^*,\delta)$, after long enough iterations, if $\mathbf{x}^t$ ever enters $\tilde{B}(\mathbf{x}^*,\delta)$, it will be trapped inside $\tilde{B}(\mathbf{x}^*,\delta)$ thereafter.

Combining the above three elements, it follows that under multi-agent DMD, for any $\delta > 0$, starting from iteration t (depending possibly on δ), $\mathbf{x}^t$ will remain in $\tilde{B}(\mathbf{x}^*,\delta)$. Since this is true for any $\delta > 0$, we have $F^\lambda(\mathbf{x}^*,\mathbf{y}^t) \to 0$ as $t \to \infty$, thereby establishing that $\mathbf{x}^t = C(\mathbf{y}^t) \to \mathbf{x}^*$ as $t \to 0$.

5.4. Almost Sure Convergence of Multi-Agent DMD to Nash under Delays and Noise Similar to Section 2.4 and adopting the same notation therein, when there are both delays and noise, multi-agent DMD is given in Algorithm 4. Characterizing convergence properties in the presence of both asynchronous delays and noise is more challenging. Here we only focus on the case where all delays are bounded.

Algorithm 4 Multi-Agent Delayed Mirror Descent under Noise

1: Each player i chooses an arbitrary initial y_i^0.
2: **for** $t = 0,1,2,\dots$ **do**
3:   **for** $i = 1,\dots,N$ **do**
4:     $X_i^t = \arg\max_{x_i \in \mathcal{X}_i} \{\langle Y_i^t, x_i\rangle - h_i(x_i)\}$
5:     $Y_i^{t+1} = Y_i^t + \frac{\alpha^t}{|\tilde{\mathcal{G}}_i^t|}\sum_{s\in\tilde{\mathcal{G}}_i^t} \tilde{v}_i(X^s) \triangleq Y_i^t + \frac{\alpha^t}{|\tilde{\mathcal{G}}_i^t|}\sum_{s\in\tilde{\mathcal{G}}_i^t}(v_i(X^s)+\xi_i^{s+1})$
6:   **end for**
7: **end for**

By using a similar framework in Remark 4.5, we obtain the following almost sure convergence result:

THEOREM 5.2. *Let $(\mathcal{N},\mathcal{X} = \prod_{i=1}^N \mathcal{X}_i, \{u_i\}_{i=1}^N)$ be a continuous game that admits $\mathcal{X}^*$ as the λ-variationally stable set of Nash equilibria. Under Assumption 2.1, if all delays are bounded (i.e. $d_i^t \le D, \forall i,t$), then $\mathbf{X}^t$ in Algorithm 4 converges to $\mathcal{X}^*$a.s..*

REMARK 5.3. One thing to note is that if $\mathcal{G}_i^t$ is always non-empty for every agent i (but otherwise fully asynchronous), then almost sure convergence to Nash would still hold even when delays grow unbounded at a $O(t\log t)$ rate in the presence of noise. Extending the convergence result to $O(t\log t)$ delays without this non-emptiness assumption (i.e. general arbitrary asynchrony in the presence of noise) is much harder and requires a substantially different argument in characterizing the behavior of the iterate dynamics. Nevertheless, we believe the result is still true and leave the full analysis for future work due to space limitation.

6. Discussion and Future Work In this paper, we have presented a framework of multi-agent online learning under imperfect information. To further demonstrate the broad applicability of this framework, we quickly discuss some consequences of our results, which can be applied to domains which may not appear at first sight directly related to the problem we are studying here. For space concerns, we will only focus on stochastic optimization. Specifically, a one-player game is an optimization problem, where each Nash equilibrium corresponds to an optimal solution. Consequently, specializing our result to the case where there is noise but no delay, we obtain that stochastic mirror descent converges to global optimal solutions almost surely for star-convex objectives. One step further, if we consider a different special case where each agent shares the same objective, we again obtain a stochastic optimization problem, but this time a distributed stochastic optimization problem where each agent is updating the global decision variables jointly. In this case, our result again gives almost sure convergence to global optimal solutions under fairly general conditions.

With the above said, we also believe much future work remains. For instance, it remains to extend the convergence result to both arbitrarily asynchronous and unbounded delays. We believe the result is still true, however, the proof will likely be substantially different. Another direction is to study other types of imperfect information such as feedback loss or duplicate feedback. Another thing to mention is that algorithms studied in this paper (OMD and/or DMD) are analyzed with respect to a family of admissible step-sizes. Sometimes an optimal step-size sequence can be chosen to optimize some additional criterion such as computational budget [46]. We leave this investigation and its implications on the applications for future work as well.

References

[1] Arora, Sanjeev, Elad Hazan, Satyen Kale. 2012. The multiplicative weights update method: A meta-algorithm and applications. *Theory of Computing* **8**(1) 121–164.

[2] Balandat, Maximilian, Walid Krichene, Claire Tomlin, Alexandre Bayen. 2016. Minimizing regret on reflexive banach spaces and learning nash equilibria in continuous zero-sum games. *arXiv preprint arXiv:1606.01261* .

[3] Bastani, Hamsa, Mohsen Bayati. 2015. Online decision-making with high-dimensional covariates .

[4] Benaïm, Michel. 1999. Dynamics of stochastic approximation algorithms. *Seminaire de probabilites XXXIII*. Springer, 1–68.

[5] Bloembergen, Daan, Karl Tuyls, Daniel Hennes, Michael Kaisers. 2015. Evolutionary dynamics of multi-agent learning: a survey. *Journal of Artificial Intelligence Research* **53** 659–697.

[6] Blum, Avrim. 1998. On-line algorithms in machine learning. *Online algorithms*. Springer, 306–325.

[7] Blum, Avrim, Eyal Even-Dar, Katrina Ligett. 2006. Routing without regret: On convergence to nash equilibria of regret-minimizing algorithms in routing games. *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*. ACM, 45–52.

[8] Busoniu, Lucian, Robert Babuska, Bart De Schutter. 2010. Multi-agent reinforcement learning: An overview. *Innovations in multi-agent systems and applications-1*. Springer, 183–221.

[9] Cesa-Bianchi, Nicolo, Gábor Lugosi. 2006. *Prediction, learning, and games*. Cambridge university press.

[10] Chien, Steve, Alistair Sinclair. 2007. Convergence to approximate nash equilibria in congestion games. *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 169–178.

[11] Coddington, A., N. Levinson. 1955. *Theory of Ordinary Differential Equations*. International series in pure and applied mathematics, R.E. Krieger. URL https://books.google.com/books?id=AUAbvgAACAAJ.

[12] Cohen, Johanne, Amélie Héliou, Panayotis Mertikopoulos. 2017. Learning with bandit feedback in potential games. *NIPS '17: Proceedings of the 31st International Conference on Neural Information Processing Systems*.

[13] Desautels, Thomas, Andreas Krause, Joel W Burdick. 2014. Parallelizing exploration-exploitation tradeoffs in gaussian process bandit optimization. *Journal of Machine Learning Research* **15**(1) 3873–3923.

[14] Facchinei, Francisco, Jong-Shi Pang. 2003. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer Science & Business Media.

[15] Fudenberg, D., D.K. Levine. 1998. *The Theory of Learning in Games*. EBSCO eBook Collection, MIT Press. URL https://books.google.com/books?id=G6vTQFluxuEC.

[16] Fudenberg, Drew, David K. Levine. 1998. *The Theory of Learning in Games*, *Economic learning and social evolution*, vol. 2. MIT Press, Cambridge, MA.

[17] Grover, Aditya, Maruan Al-Shedivat, Jayesh K Gupta, Yura Burda, Harrison Edwards. 2018. Learning policy representations in multiagent systems. *International Conference on Machine Learning*.

[18] Hall, P., C.C. Heyde. 1980. *Martingale limit theory and its application*. Probability and mathematical statistics, Academic Press. URL https://books.google.com/books?id=xxbvAAAAMAAJ.

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

[20] Joulani, Pooria, Andras Gyorgy, Csaba Szepesvári. 2013. Online learning under delayed feedback. *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*. 1453–1461.

[21] Kakade, Sham M., Shai Shalev-Shwartz, Ambuj Tewari. 2012. Regularization techniques for learning with matrices. *The Journal of Machine Learning Research* **13** 1865–1890.

[22] Kalai, Adam, Santosh Vempala. 2005. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences* **71**(3) 291–307.

[23] Kelly, Frank P., Aman K. Maulloo, David K. H. Tan. 1998. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society* **49**(3) 237–252.

[24] Krichene, Syrine, Walid Krichene, Roy Dong, Alexandre Bayen. 2015. Convergence of heterogeneous distributed learning in stochastic routing games. *Communication, Control, and Computing (Allerton), 2015 53rd Annual Allerton Conference on*. IEEE, 480–487.

[25] Lam, Kiet, Walid Krichene, Alexandre Bayen. 2016. On learning how players learn: estimation of learning dynamics in the routing game. *Cyber-Physical Systems (ICCPS), 2016 ACM/IEEE 7th International Conference on*. IEEE, 1–10.

[26] Littlestone, Nick, Manfred K Warmuth. 1994. The weighted majority algorithm. *INFORMATION AND COMPUTATION* **108** 212–261.

[27] Mehta, Ruta, Ioannis Panageas, Georgios Piliouras. 2015. Natural selection as an inhibitor of genetic diversity: Multiplicative weights updates algorithm and a conjecture of haploid genetics. *ITCS '15: Proceedings of the 6th Conference on Innovations in Theoretical Computer Science*.

[28] Menache, Ishai, Asuman Ozdaglar. 2011. Network games: Theory, models, and dynamics. *Synthesis Lectures on Communication Networks* **4**(1) 1–159.

[29] Mertikopoulos, Panayotis, Christos Papadimitriou, Georgios Piliouras. 2018. Cycles in adversarial regularized learning. *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 2703–2717.

[30] Mertikopoulos, Panayotis, Christos H. Papadimitriou, Georgios Piliouras. to appear. Cycles in adversarial regularized learning. *SODA '18: Proceedings of the 29th annual ACM-SIAM symposium on discrete algorithms*.

[31] Mertikopoulos, Panayotis, Zhengyuan Zhou. 2018. Learning in games with continuous action sets and unknown payoff functions. *Mathematical Programming* 1–43.

[32] Monderer, Dov, Lloyd S Shapley. 1996. Potential games. *Games and economic behavior* **14**(1) 124–143.

[33] Monnot, Barnabé, Georgios Piliouras. 2017. Limits and limitations of no-regret learning in games. *The Knowledge Engineering Review* **32**.

[34] Nemirovski, Arkadi Semen, David Berkovich Yudin. 1983. *Problem Complexity and Method Efficiency in Optimization*. Wiley, New York, NY.

[35] Nesterov, Yurii. 2009. Primal-dual subgradient methods for convex problems. *Mathematical Programming* **120**(1) 221–259.

[36] Orda, Ariel, Raphael Rom, Nahum Shimkin. 1993. Competitive routing in multi-user communication networks **1**(5) 614–627.

[37] Palaiopanos, Gerasimos, Ioannis Panageas, Georgios Piliouras. 2017. Multiplicative weights update with constant step-size in congestion games: Convergence, limit cycles and chaos. *NIPS '17: Proceedings of the 31st International Conference on Neural Information Processing Systems*.

[38] Perkins, Steven, Panayotis Mertikopoulos, David S. Leslie. 2017. Mixed-strategy learning with continuous action sets **62**(1) 379–384.

[39] Quanrud, Kent, Daniel Khashabi. 2015. Online learning with adversarial delays. *Advances in Neural Information Processing Systems*. 1270–1278.

[40] Rockafellar, R Tyrrell, Roger J-B Wets. 2009. *Variational analysis*, vol. 317. Springer Science & Business Media.

[41] Rockafellar, Ralph Tyrell. 2015. *Convex analysis*. Princeton university press.

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

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

[44] Shalev-Shwartz, Shai, Yoram Singer. 2007. Convex repeated games and Fenchel duality. *Advances in Neural Information Processing Systems 19*. MIT Press, 1265–1272.

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

[46] Shanbhag, Uday V, Jose H Blanchet. 2015. Budget-constrained stochastic approximation. *Winter Simulation Conference (WSC), 2015*. IEEE, 368–379.

[47] Shoham, Yoav, Kevin Leyton-Brown. 2008. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press.

[48] Viossat, Yannick, Andriy Zapechelnyuk. 2013. No-regret dynamics and fictitious play. *Journal of Economic Theory* **148**(2) 825–842.

[49] Viossat, Yannick, Andriy Zapechelnyuk. 2013. No-regret dynamics and fictitious play. *Journal of Economic Theory* **148**(2) 825–842.

[50] Xiao, Lin. 2010. Dual averaging methods for regularized stochastic learning and online optimization. *Journal of Machine Learning Research* **11**(Oct) 2543–2596.

[51] Zhu, Minghui, Emilio Frazzoli. 2016. Distributed robust adaptive equilibrium computation for generalized convex games. *Automatica* **63** 82–91.

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

Appendix. Missing Proofs

A. Proof of Lemma 4.1 For the first statement, note that by a well-known result in convex analysis (see [41]) when $x_i^* = C_i(y_i)$, it holds that y_i is a subgradient of h_i at the point x_i^*: $h_i(x_i) \geq h_i(x_i^*) + y_i(x_i - x_i^*)$. This then implies that, for any $\beta \in (0,1]$:

$$\lambda_i(h_i(x_i^*) + y_i(\beta(x_i - x_i^*))) \leq \lambda_i h_i(x_i^* + \beta(x_i - x_i^*)) = \lambda_i h_i((1-\beta)x_i^* + \beta x_i) \leq \tag{A.1}$$

$$\lambda_i\{(1-\beta)h_i(x_i^*) + \beta h_i(x_i) - \frac{1}{2}K_i\beta(1-\beta)\|x_i - x_i^*\|_i^2\}, \tag{A.2}$$

where the last inequality follows from the fact that h_i is K_i-strongly convex.

Connecting the first and the last term of the above inequality chain, we have:

$$\lambda_i y_i \beta(x_i - x_i^*) \leq \lambda_i(-\beta)h_i(x_i^*) + \lambda_i\beta h_i(x_i) - \frac{1}{2}K_i\lambda_i\beta(1-\beta)\|x_i - x_i^*\|_i^2.$$

Dividing both sides by β (since $\beta > 0$) and rearranging, we obtain:

$$\frac{1}{2}K_i\lambda_i(1-\beta)\|x_i - x_i^*\|_i^2 \leq \lambda_i h_i(x_i) - \lambda_i h_i(x_i^*) - \lambda_i y_i(x_i - x_i^*).$$

Taking the limit that β approaches 0 (from above) results: $\lambda_i h_i(x_i) - \lambda_i h_i(x_i^*) - \lambda_i y_i(x_i - x_i^*) \geq \frac{1}{2}K_i\lambda_i\|x_i - x_i^*\|_i^2$. Summing over all i's, we then obtain:

$$\sum_{i=1}^{N}\{\lambda_i h_i(x_i) - \lambda_i h_i(x_i^*) - \lambda_i y_i(x_i - x_i^*)\} \geq \sum_{i=1}^{N}\frac{1}{2}K_i\lambda_i\|x_i - x_i^*\|_i^2.$$

The conclusion then follows by noting that:

$$F^\lambda(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{N}\lambda_i(h_i(x_i) - x_i y_i + h_i^*(y_i)) = \sum_{i=1}^{N}\lambda_i(h_i(x_i) - x_i y_i + x_i^* y_i - h_i(x_i^*)) \tag{A.3}$$

$$\geq \sum_{i=1}^{N}\frac{1}{2}K_i\lambda_i\|x_i - x_i^*\|_i^2 \geq \frac{1}{2}(\min_i K_i\lambda_i)\sum_{i=1}^{N}\|C_i(y_i) - x_i\|_i^2, \tag{A.4}$$

where the last inequality follows by noting that $x_i^* = C_i(y_i)$.

For the second statement, we start by citing here a useful result in [40] (Theorem 12.60): For a proper, lower semi-continuous and convex function $f : \mathbf{R}^n \to \bar{\mathbf{R}}$ and a value $\sigma > 0$, f^* is σ-strongly convex (with respect to norm $\|\cdot\|^*$) if and only if f is differentiable and satisfies:

$$f(\tilde{\mathbf{x}}) \leq f(\mathbf{x}) + < \nabla f(\mathbf{x}), \tilde{\mathbf{x}} - \mathbf{x}) > + \frac{1}{2\sigma}\|\tilde{\mathbf{x}} - \mathbf{x}\|^2, \forall \mathbf{x}, \tilde{\mathbf{x}},$$

where $\bar{\mathbf{R}} = [-\infty, \infty]$. Note that in the original statement, only the Euclidean norm $\|\cdot\|_2$ is used (Definition 12.58 in [40] defined strong convexity implicitly in terms of the Euclidean norm), in which case $\|\cdot\|_2^* = \|\cdot\|_2$. However, as stated here, the same result holds true for any pair of norms $(\|\cdot\|, \|\cdot\|^*)$ by a straightforward adaptation of their proof.

Next, we note that in our case, each h_i is K_i-strongly convex with respect to norm $\|\cdot\|_i$ and since h_i is proper, lower semi-continuous and convex, it follows that $(h_i^*)^* = h_i$ (Theorem 11.1 in [40]). Further, it can be easily checked that h_i^* is proper, lower semi-continuous and convex (since it is a point-wise maximum of affine functions per its definition), it therefore follows that the K_i-strong convexity of $(h_i^*)^*$ (with respect to $\|\cdot\|_i^{**} = \|\cdot\|_i$) implies that h_i^* is differentiable and satisfies:

$$h_i^*(\tilde{y}_i) \le h_i^*(y_i) + (h_i^*)'(y_i)(\tilde{y}_i - y_i) + \frac{1}{2K_i}(\|\tilde{y}_i - y_i\|_i^*)^2, \forall y_i, \tilde{y}_i \tag{A.5}$$

$$= h_i^*(y_i) + C_i(y_i)(\tilde{y}_i - y_i) + \frac{1}{2K_i}(\|\tilde{y}_i - y_i\|_i^*)^2, \forall y_i, \tilde{y}_i \tag{A.6}$$

where the equality follows because $(h_i^*)'(y_i) = C_i(y_i)$.

Therefore, it then follows that upon substituting the preceding inequality for each $h_i^*(\tilde{y}_i)$ into $F^\lambda(\mathbf{x}, \tilde{\mathbf{y}}) = \sum_{i=1}^N \lambda_i(h_i(x_i) - x_i\tilde{y}_i + h_i^*(\tilde{y}_i))$, we have:

$$F^\lambda(\mathbf{x}, \tilde{\mathbf{y}}) \le \sum_{i=1}^N \lambda_i(h_i(x_i) - x_i\tilde{y}_i) + \sum_{i=1}^N \lambda_i\{h_i^*(y_i) + C_i(y_i)(\tilde{y}_i - y_i) + \frac{\lambda_i}{2K_i}(\|\tilde{y}_i - y_i\|_i^*)^2\} \tag{A.7}$$

$$= \sum_{i=1}^N \lambda_i\{h_i(x_i) + h_i^*(y_i) - x_i y_i + x_i(y_i - \tilde{y}_i) + C_i(y_i)(\tilde{y}_i - y_i))\} + \sum_{i=1}^N \frac{\lambda_i}{2K_i}(\|\tilde{y}_i - y_i\|_i^*)^2 \tag{A.8}$$

$$= F^\lambda(\mathbf{x}, \mathbf{y}) + \sum_{i=1}^N \lambda_i(\tilde{y}_i - y_i)(C_i(y_i) - x_i) + \sum_{i=1}^N \frac{\lambda_i}{2K_i}(\|\tilde{y}_i - y_i\|_i^*)^2 \tag{A.9}$$

$$\le F^\lambda(\mathbf{x}, \mathbf{y}) + \sum_{i=1}^N \lambda_i(\tilde{y}_i - y_i)(C_i(y_i) - x_i) + \frac{1}{2}(\max_i \frac{\lambda_i}{K_i}) \sum_{i=1}^N (\|\tilde{y}_i - y_i\|_i^*)^2. \tag{A.10}$$

■

B. Proof to Theorem 4.1 We present the details to each of the four steps outlined in Remark 4.2.

1. We start by fixing some notation. Let $\mathbf{y}^t = (y_1^t, \ldots, y_N^t), \mathbf{x}^t = (x_1^t, \ldots, x_N^t)$ be the iterates generated in Algorithm 1. Since $\mathcal{X}$ is compact and $v(\cdot)$ is continuous, $V_{\max}^i \triangleq \max_{x_i \in \mathcal{X}_i} \|v_i(\mathbf{x})\|_i^* < \infty, V_{\max} \triangleq \max_{\mathbf{x} \in \mathcal{X}} \|v(\mathbf{x})\|^* = \sum_{i=1}^N V_{\max}^i < \infty$. Since each $h_i(\cdot)$ is K_i strongly convex (with respect to $\|\cdot\|_i$), it follows from a well-known result in convex analysis [41] that the choice map $C(\cdot)$ is $\frac{1}{K}$-Lipschitz continuous, where $K \triangleq \min_i K_i$. Finally, since each $v_i(\cdot)$ is Lipschitz continuous, $v(\cdot)$ is Lipschitz continuous as well and denote the Lipschitz constant as L.

Since $d^t \le D, \forall t$, it follows that $|\mathcal{G}^t| \le D$ and $\min \mathcal{G}^t \ge t - D + 1$, for otherwise at least one gradient comes from $D+1$ rounds before. Further, per the OMD update (first equality in Equation 4.1), we have:

$$\|\mathbf{y}^{t+1} - \mathbf{y}^t\|^* = \sum_{i=1}^{N} \|y_i^{t+1} - y_i^t\|_i^* = \sum_{i=1}^{N} \|\alpha^t \sum_{s \in \mathcal{G}^t} v_i(\mathbf{x}^s)\|_i^* \tag{B.1}$$

$$\le \alpha^t \sum_{i=1}^{N} \sum_{s \in \mathcal{G}^t} \|v_i(\mathbf{x}^s)\|_i^* \le \alpha^t \sum_{i=1}^{N} |\mathcal{G}^t| V_{\max}^i \le \alpha^t D V_{\max} \tag{B.2}$$

By definition, we can expand b_i^t as follows:

$$b_i^t = \sum_{s \in \mathcal{G}^t} \{v_i(\mathbf{x}^s) - v_i(\mathbf{x}^t)\} \le \sum_{s \in \mathcal{G}^t} L\|\mathbf{x}^s - \mathbf{x}^t\| = \sum_{s \in \mathcal{G}^t} L\|C(\mathbf{y}^s) - C(\mathbf{y}^t)\| \le \sum_{s \in \mathcal{G}^t} \frac{L}{K}\|\mathbf{y}^s - \mathbf{y}^t\|^* \tag{B.3}$$

$$\le \frac{L}{K} \sum_{s \in \mathcal{G}^t} \left\{\|\mathbf{y}^s - \mathbf{y}^{s+1}\|^* + \|\mathbf{y}^{s+1} - \mathbf{y}^{s+2}\|^* + \cdots + \|\mathbf{y}^{t-1} - \mathbf{y}^t\|^*\right\} \tag{B.4}$$

$$\le \frac{L}{K} \sum_{s \in \mathcal{G}^t} \left\{\alpha^s D V_{\max} + \alpha^{s+1} D V_{\max} + \cdots + \alpha^t D V_{\max}\right\} \tag{B.5}$$

$$= \frac{LDV_{\max}}{K} \sum_{s \in \mathcal{G}^t} \sum_{k=s}^{t} \alpha^k \le \frac{LDV_{\max}}{K} |\mathcal{G}^t| \sum_{s=\min \mathcal{G}^t}^{t} \alpha^s \le \frac{LD^2 V_{\max}}{K} \sum_{s=\min \mathcal{G}^t}^{t} \alpha^s \tag{B.6}$$

$$\le \frac{LD^2 V_{\max}}{K} \sum_{s=t-D+1}^{t} \alpha^s \le \frac{LD^3 V_{\max}}{K} \alpha^{t-D+1} \to 0, \text{as } t \to \infty, \tag{B.7}$$

where the first inequality in Equation B.3 follows from the fact that $v(\cdot)$ is L-Lipschitz continuous, the second inequality in Equation B.3 follows from the fact that $C(\cdot)$ is $\frac{1}{K}$-Lipschitz continuous, Equation B.5 follows from Equations B.1 and B.2 and the first inequality in Equation B.6 follows from that α^t's are non-negative and the second inequality in Equation B.7 follows from α^t is non-increasing. Lastly, the convergence to 0 in Equation B.7 follows from the fact that α^t is square-summable.

2. Fix an arbitrary $\epsilon > 0$. Assume for contradiction purposes that $\mathbf{x}^t$ only visits $B(\mathbf{x}^*, \epsilon)$ a finite number of times and hence let $t^0 - 1$ be the last time $\mathbf{x}^t$ is in $B(\mathbf{x}^*, \epsilon)$: $\forall t \ge t^0, \mathbf{x}^t \in \mathcal{X} - B(\mathbf{x}^*, \epsilon)$. Since $\mathcal{X} - B(\mathbf{x}^*, \epsilon)$ is a compact set and $v_i(\mathbf{x})$ is continuous in $\mathbf{x}$ and since by assumption $\sum_{i=1}^{N} \lambda_i v_i(\mathbf{x})(x_i - x_i^*) < 0, \forall \mathbf{x} \in \mathcal{X}, \mathbf{x} \ne \mathbf{x}^*$, it follows that there exists a $c_{\max}(\epsilon) < 0$ such that

$$\sum_{i=1}^{N} \lambda_i v_i(\mathbf{x})(x_i - x_i^*) \le c_{\max}(\epsilon), \forall \mathbf{x} \in \mathcal{X} - B(\mathbf{x}^*, \epsilon). \tag{B.8}$$

Per Claim 1, $\lim_{t \to \infty} \mathbf{b}^t = \mathbf{0}$, therefore, $\|\mathbf{b}^t\|^*$ is bounded and we denote $B_{\max} \triangleq \max_t \|\mathbf{b}^t\|^*$. Next denote $R = \max_{\mathbf{x} \in \mathcal{X}} \|\mathbf{x}\|$, $\lambda_{\max} \triangleq \max_i \lambda_i$ and $\beta^t \triangleq \max_i \frac{(\alpha^t)^2 \lambda_i}{2K_i}$ and note that $\sum_{t=1}^{\infty} \beta^t < \infty$. Using Lemma 4.1, we have $\forall t \ge t^0$:

$$F^{\lambda}(\mathbf{x}^*, \mathbf{y}^{t+1}) = F^{\lambda}(\mathbf{x}^*, \mathbf{y}^t + \alpha^t \{|\mathcal{G}^t| v(\mathbf{x}^t) + \mathbf{b}^t\}) \le \tag{B.9}$$

$$F^{\lambda}(\mathbf{x}^*, \mathbf{y}^t) + \sum_{i=1}^{N} \lambda_i (\alpha^t \{|\mathcal{G}^t| v_i(\mathbf{x}^t) + b_i^t\})(C_i(y_i^t) - x_i^*) + \beta^t (\||\mathcal{G}^t| v(\mathbf{x}^t) + \mathbf{b}^t\|^*)^2 = \tag{B.10}$$

$$F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\alpha^t\left\{|\mathcal{G}^t|\sum_{i=1}^N\lambda_i v_i(\mathbf{x}^t)(x_i^t-x_i^*)+\sum_{i=1}^N\lambda_i b_i^t(x_i^t-x_i^*)\right\}+\beta^t(\||\mathcal{G}^t|v(\mathbf{x}^t)+\mathbf{b}^t\|^*)^2 \tag{B.11}$$

$$\le F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\alpha^t\left\{|\mathcal{G}^t|c_{\max}(\epsilon)+\lambda_{\max}\|\mathbf{b}^t\|^*\|\mathbf{x}^t-\mathbf{x}^*\|\right\}+\beta^t\left\{2(\||\mathcal{G}^t|v(\mathbf{x}^t)\|^*)^2+2(\|\mathbf{b}^t\|^*)^2\right\} \tag{B.12}$$

$$\le F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\alpha^t\left\{|\mathcal{G}^t|c_{\max}(\epsilon)+\lambda_{\max}R\|\mathbf{b}^t\|^*\right\}+2\beta^t(D^2V_{\max}^2+B_{\max}^2) \tag{B.13}$$

$$\le F^\lambda(\mathbf{x}^*,\mathbf{y}^{t^0})+(\sum_{k=t^0}^t\alpha^k)\left\{\frac{\sum_{k=t^0}^t\alpha^k|\mathcal{G}^k|}{\sum_{k=t^0}^t\alpha^k}c_{\max}(\epsilon)+\lambda_{\max}R\frac{\sum_{k=t^0}^t\alpha^k\|\mathbf{b}^k\|^*}{\sum_{k=t^0}^t\alpha^k}\right\} \tag{B.14}$$

$$+2(\sum_{k=t^0}^t\beta^k)(D^2V_{\max}^2+B_{\max}^2), \tag{B.15}$$

where Equation (B.12) follows from Equation (B.8) and Equation (B.14) follows from telescoping.

Next, we show that:

$$1\le\lim_{t\to\infty}\frac{\sum_{k=t^0}^t\alpha^k|\mathcal{G}^k|}{\sum_{k=t^0}^t\alpha^k}\le D. \tag{B.16}$$

Partition the rounds into intervals $\{0,1,\ldots,D-1\},\{D,D+1,\ldots,2D-1\},\ldots$. Since each gradient is received exactly once with at most delay D, the gradients corresponding to the first interval will have been completely received by time $2D-1$ (i.e. by the end of the second interval). Similarly, the gradients corresponding to the l-th interval will have been all received by time $(l+1)D-1$. Consequently, since α^t is non-increasing, it follows that:

$$\sum_{k=0}^\infty\alpha^k|\mathcal{G}^k|\ge\sum_{l=2}^\infty D\alpha^{lD-1}\ge\sum_{k=2D-1}^\infty\alpha^k=\infty.$$

Consequently,

$$\lim_{t\to\infty}\frac{\sum_{k=t^0}^t\alpha^k|\mathcal{G}^k|}{\sum_{k=t^0}^t\alpha^k}=\lim_{t\to\infty}\frac{\sum_{k=0}^t\alpha^k|\mathcal{G}^k|}{\sum_{k=0}^t\alpha^k}=\lim_{t\to\infty}\frac{\sum_{k=0}^t\alpha^k|\mathcal{G}^k|}{\sum_{k=2D-1}^t\alpha^k}\ge 1.$$

Finally, $\frac{\sum_{k=t^0}^t\alpha^k|\mathcal{G}^t|}{\sum_{k=t^0}^t\alpha^k}\le D$ follows easily by noting that $|\mathcal{G}^t|\le D$.

Next, note that since $\lim_{t\to\infty}\mathbf{b}^t=\mathbf{0}$ and $\sum_{t=0}^\infty\alpha^t=\infty$, it follows that:

$$\frac{\sum_{k=t^0}^t\alpha^k\|\mathbf{b}^k\|^*}{\sum_{k=t^0}^t\alpha^k}\to 0,\ \text{as } t\to\infty. \tag{B.17}$$

Combining Equation B.16 and Equation B.17, we obtain:

$$(\sum_{k=t^0}^t\alpha^k)\left\{\frac{\sum_{k=t^0}^t\alpha^k|\mathcal{G}^k|}{\sum_{k=t^0}^t\alpha^k}c_{\max}(\epsilon)+\lambda_{\max}R\frac{\sum_{k=t^0}^t\alpha^k\|\mathbf{b}^k\|^*}{\sum_{k=t^0}^t\alpha^k}\right\}\to-\infty,\ \text{as } t\to\infty.$$

Since $\sum_{k=t^0}^\infty\beta^k<\infty$, Equation (B.14) implies that $F^\lambda(\mathbf{x}^*,\mathbf{y}^t)\to-\infty$, which contradicts the first statement in Lemma 4.1. The claim therefore follows.

3. Assume for contradiction purposes no $B(\mathbf{x}^*,\epsilon)$ is contained in $\tilde{B}(\mathbf{x}^*,\delta)$, which means that for any $\delta>0,\exists\mathbf{y}^l$, such that $\|Q(\mathbf{y}^l)-\mathbf{x}^*\|=\delta$ but $F^\lambda(\mathbf{x}^*,\mathbf{y}^l)\ge\epsilon$. This produces a sequence $\{\mathbf{y}^l\}_{l=0}^\infty$ such that $C(\mathbf{y}^l)\to\mathbf{x}^*$ but $F^\lambda(\mathbf{x}^*,\mathbf{y}^l)\ge\epsilon,\forall l$. This contradicts with the fact that the choice map $C(\cdot)$ is λ-Fenchel coupling conforming, because by definition it holds that if $C(\mathbf{y}^t)\to x$, then $F^\lambda(x,\mathbf{y}^t)\to 0$. Consequently, the claim follows.

4. Fix a given $\delta > 0$. Since β^t is summable and α^t is not summable but square summable, it follows that $\beta^t \to 0, \alpha^t \to 0, \frac{\alpha^t}{\beta^t} \to \infty$ as $t \to \infty$. There, denote
   (a) $T^1(\delta) = \arg\min_t \{t \mid \beta^s < \frac{\delta}{8(D^2 V_{\max}^2 + B_{\max}^2)}, \forall s \geq t\}$.
   (b) $T^2(\delta) = \arg\min_t \{t \mid c_{\max}(\epsilon(\frac{\delta}{2})) < -2\lambda_{\max} R \|\mathbf{b}^s\|^*, \forall s \geq t\}$.
   (c) $T^3(\delta) = \arg\min_t \{t \mid \alpha^s < \frac{\delta}{4\lambda_{\max} R B_{\max}}, \forall s \geq t\}$.
   (d) $T^4(\delta) = \arg\min_t \{t \mid \frac{\alpha^s}{\beta^s} > \frac{4(D^2 V_{\max}^2 + B_{\max}^2)}{-c_{\max}(\epsilon(\frac{\delta}{2}))}, \forall s \geq t\}$.

   We have $T^1(\delta) < \infty, T^2(\delta) < \infty$ (since $\lim_{t\to\infty} \|\mathbf{b}^t\|^* = 0$ and note that $c_{\max}(\epsilon(\frac{\delta}{2})) < 0$ by definition), $T^3(\delta) < \infty, T^4(\delta) < \infty$. Take

$$T(\delta) = \max(T^1(\delta), T^2(\delta), T^3(\delta), T^4(\delta)\}.$$

   Now take any $t \geq T(\delta)$. We show that if $\mathbf{x}^t \in \tilde{B}(\mathbf{x}^*, \delta)$, then $\mathbf{x}^{t+1} \in \tilde{B}(\mathbf{x}^*, \delta)$. To see that this statement holds, let $\mathbf{x}^t \in \tilde{B}(\mathbf{x}^*, \delta)$, which implies that $F^\lambda(\mathbf{x}^*, \mathbf{y}^t) < \delta$. Note that it suffices to consider $\mathcal{G}^t \neq \emptyset$, for otherwise $\mathbf{x}^{t+1} = \mathbf{x}^t$.

   Now there are two possibilities:
   (a) Possibility 1: $\mathbf{x}^t \in B(\mathbf{x}^*, \epsilon(\frac{\delta}{2}))$.
   (b) Possibility 2: $\mathbf{x}^t \in \tilde{B}(\mathbf{x}^*, \delta) - B(\mathbf{x}^*, \epsilon(\frac{\delta}{2}))$.

   Under Possibility 1, it follows

$$F^\lambda(\mathbf{x}^*, \mathbf{y}^{t+1}) \leq F^\lambda(\mathbf{x}^*, \mathbf{y}^t) + \alpha^t \sum_{i=1}^N \lambda_i \{|\mathcal{G}^t| v_i(\mathbf{x}^t) + b_i^t\}(x_i^t - x_i^*) + \beta^t (\||\mathcal{G}^t| v(\mathbf{x}_t) + \mathbf{b}^t\|^*)^2 \tag{B.18}$$

$$\leq F^\lambda(\mathbf{x}^*, \mathbf{y}^t) + \alpha^t \sum_{i=1}^N \lambda_i b_i^t (x_i^t - x_i^*) + \beta^t \left\{2(\||\mathcal{G}^t| v(\mathbf{x}_t)\|^*)^2 + 2(\|\mathbf{b}^t\|^*)^2\right\} \tag{B.19}$$

$$\leq F^\lambda(\mathbf{x}^*, \mathbf{y}^t) + \alpha^t \lambda_{\max} R B_{\max} + 2\beta^t (D^2 V_{\max}^2 + B_{\max}^2) \tag{B.20}$$

$$< F^\lambda(\mathbf{x}^*, \mathbf{y}^t) + \frac{\delta}{4\lambda_{\max} R B_{\max}} \lambda_{\max} R B_{\max} + \frac{2\delta}{8(D^2 V_{\max}^2 + B_{\max}^2)} (D^2 V_{\max^2} + B_{\max}^2) \tag{B.21}$$

$$\leq \frac{\delta}{2} + \frac{\delta}{4} + \frac{\delta}{4} = \delta, \tag{B.22}$$

   where the second inequality follows from λ-variational stability and the last inequality follows from the fact that $\mathbf{x}^t \in B(\mathbf{x}^*, \epsilon(\frac{\delta}{2})) \subset \tilde{B}(\mathbf{x}^*, \frac{\delta}{2})$ per Claim 2.

   Under Possibility 2, it follows from Equation B.13 that

$$F^\lambda(\mathbf{x}^*, \mathbf{y}^{t+1}) \leq F^\lambda(\mathbf{x}^*, \mathbf{y}^t) + \alpha^t \left\{|\mathcal{G}^t| c_{\max}(\epsilon(\frac{\delta}{2})) + \lambda_{\max} R \|\mathbf{b}^t\|^*\right\} + 2\beta^t (D^2 V_{\max}^2 + B_{\max}^2) \tag{B.23}$$

$$\leq F^\lambda(\mathbf{x}^*, \mathbf{y}^t) + 2\beta^t (D^2 V_{\max}^2 + B_{\max}^2) \left\{\frac{\alpha^t}{\beta^t} \frac{c_{\max}(\epsilon(\frac{\delta}{2})) + \lambda_{\max} R \|\mathbf{b}^t\|^*}{2(D^2 V_{\max}^2 + B_{\max}^2)} + 1\right\} \tag{B.24}$$

$$\leq F^\lambda(\mathbf{x}^*, \mathbf{y}^t) + 2\beta^t (V_{\max}^2 + B_{\max}^2) \left\{\frac{\alpha^t}{\beta^t} \frac{c_{\max}(\epsilon(\frac{\delta}{2}))}{4(V_{\max}^2 + B_{\max}^2)} + 1\right\} \tag{B.25}$$

$$< F^\lambda(\mathbf{x}^*, \mathbf{y}^t) < \epsilon, \tag{B.26}$$

where the second inequality follows from $|\mathcal{G}^t| \geq 1$ since it is not empty by assumption and $c_{\max} < 0$, the third inequality follows from $\lambda_{\max} R \|\mathbf{b}^t\|^* < -\frac{1}{2} c_{\max}(\epsilon(\frac{\delta}{2}))$ since $t \geq T^2(\delta)$ and the second-to-last inequality follows from $\frac{\alpha^t}{\beta^t} \frac{c_{\max}(\epsilon(\frac{\delta}{2}))}{4(V_{\max}^2 + B_{\max}^2)} + 1 < 0$ since $t \geq T^4(\delta)$.

■

C. Proof to Theorem 4.5 We prove in turn each of the 7 claims laid out in Remark 5.2.

1. Adopt the same notation in the proof to Claim 1 in Remark 4.2 and let $Y^t = (Y_1^t, \ldots, Y_N^t), X^t = (X_1^t, \ldots, X_N^t)$ be the iterates generated in Algorithm 2. That $B^t \to \mathbf{0}$ almost surely follows exactly the same argument (path-by-path) as that to $b^t \to \mathbf{0}$ in the proof for the first claim of Theorem 4.1. Fix an arbitrary $\epsilon > 0$. Assume for contradiction purposes that X^t only visits $B(\mathbf{x}^*, \epsilon)$ a finite number of times and hence let $t^0 - 1$ be the last time X^t is in $B(\mathbf{x}^*, \epsilon)$. Since $Y^{t+1} = Y^t + \alpha^t \sum_{s \in \mathcal{G}^t} (v(X^s) + \xi^{s+1}) = Y^t + \alpha^t \left\{ |\mathcal{G}^t| v(X^t) + \sum_{s \in \mathcal{G}^t} \xi^{s+1} + B^t \right\}$, we have that $\forall t \geq t^0$:

$$F^\lambda(\mathbf{x}^*, Y^{t+1}) = F^\lambda(\mathbf{x}^*, Y^t + \alpha^t \left\{ |\mathcal{G}^t| v(X^t) + \sum_{s \in \mathcal{G}^t} \xi^{s+1} + B^t \right\}) \tag{C.1}$$

$$\leq F^\lambda(\mathbf{x}^*, Y^t) + \sum_{i=1}^N \lambda_i (\alpha^t (|\mathcal{G}^t| v_i(X^t) + B_i^t + \sum_{s \in \mathcal{G}^t} \xi_i^{s+1})(C_i(Y_i^t) - x_i^*) + \beta^t (\| |\mathcal{G}^t| v(X^t) + B^t + \sum_{s \in \mathcal{G}^t} \xi^{s+1} \|^*)^2 \tag{C.2}$$

$$= F^\lambda(\mathbf{x}^*, Y^t) + \alpha^t |\mathcal{G}^t| \sum_{i=1}^N \lambda_i v_i(X^t)(X_i^t - x_i^*) + \alpha^t \sum_{i=1}^N \lambda_i (\sum_{s \in \mathcal{G}^t} \xi_i^{s+1})(X_i^t - x_i^*) + \alpha^t \sum_{i=1}^N \lambda_i B_i^t (X_i^t - x_i^*) \tag{C.3}$$

$$+ \beta^t (\| |\mathcal{G}^t| v(X^t) + B^t + \sum_{s \in \mathcal{G}^t} \xi^{s+1} \|^*)^2 \tag{C.4}$$

$$\leq F^\lambda(\mathbf{x}^*, Y^t) + \alpha^t |\mathcal{G}^t| b_{\max}(\epsilon) + \alpha^t \sum_{i=1}^N \lambda_i (\sum_{s \in \mathcal{G}^t} \xi_i^{s+1})(X_i^t - x_i^*) + \alpha^t \sum_{i=1}^N \lambda_i B_i^t (X_i^t - x_i^*) \tag{C.5}$$

$$+ 3\beta^t \left\{ D^2 (\|v(X^t)\|^*)^2 + (\|B^t\|^*)^2 + D \sum_{s \in \mathcal{G}^t} (\|\xi^{s+1}\|^*)^2 \right\} \tag{C.6}$$

$$\leq F^\lambda(\mathbf{x}^*, Y^{t^0}) + (\sum_{k=t^0}^t \alpha^k |\mathcal{G}^k|) b_{\max}(\epsilon) + \sum_{k=t^0}^t \alpha^k \sum_{i=1}^N \lambda_i (\sum_{s \in \mathcal{G}^k} \xi_i^{s+1})(X_i^k - x_i^*) + \sum_{k=t^0}^t \alpha^k \sum_{i=1}^N \lambda_i B_i^k (X_i^k - x_i^*) \tag{C.7}$$

$$+ 3 \sum_{k=t^0}^t \beta^k \left\{ D^2 (\|v(X^t)\|^*)^2 + (\|B^t\|^*)^2 + D \sum_{s \in \mathcal{G}^t} (\|\xi^{s+1}\|^*)^2 \right\} \tag{C.8}$$

$$= F^\lambda(\mathbf{x}^*, Y^{t^0}) + (\sum_{k=t^0}^t \alpha^k |\mathcal{G}^k|) \left\{ b_{\max}(\epsilon) + \sum_{i=1}^N \lambda_i (\sum_{k=t^0}^t \frac{\alpha^k}{\sum_{k=t^0}^t \alpha^k |\mathcal{G}^k|} (\sum_{s \in \mathcal{G}^k} \xi_i^{s+1})(X_i^k - x_i^*)) \right. \tag{C.9}$$

$$\left. + \sum_{i=1}^N (\lambda_i \sum_{k=t^0}^t \frac{\alpha^k}{\sum_{k=t^0}^t \alpha^k |\mathcal{G}^k|} B_i^{k+1} (X_i^k - x_i^*)) \right\} + 3 \sum_{k=t^0}^t \beta^k \left\{ D^2 (\|v(X^t)\|^*)^2 + (\|B^t\|^*)^2 + D \sum_{s \in \mathcal{G}^t} (\|\xi^{s+1}\|^*)^2 \right\} \tag{C.10}$$

$$\to -\infty, \text{ a.s.,} \tag{C.11}$$

where the last inequality holds because $\sum_{k=t^0}^{t} \frac{\alpha^k}{\sum_{k=t^0}^{t} \alpha^k |\mathcal{G}^k|} (\sum_{s\in\mathcal{G}^k} \xi_i^{s+1})(X_i^k - x_i^*)) \to 0$, a.s. as $t \to \infty$ by law of large numbers for martingales, $\sum_{k=t^0}^{t} \frac{\alpha^k}{\sum_{k=t^0}^{t} \alpha^k |\mathcal{G}^k|} B_i^{k+1}(X_i^k - x_i^*) \to 0$, a.s. as $t \to \infty$ (since $B^t \to 0$ a.s. as $t \to \infty$), $3\sum_{k=t^0}^{t} \beta^k \left\{D^2(\|v(X^t)\|^*)^2 + (\|B^t\|^*)^2 + D\sum_{s\in\mathcal{G}^t}(\|\xi^{s+1}\|^*)^2\right\} \to C$, a.s. as $t \to \infty$, for some random variable C that is almost surely finite by Doob's martingale convergence theorem and $\sum_{k=t^0}^{t} \alpha^k |\mathcal{G}^k| \to \infty$ as $t \to \infty$. This contradicts the first statement in Lemma 4.1 and the claim therefore follows.

2. Since each $h_i(\cdot)$ is K_i strongly convex, $h(\cdot) = (h_1(\cdot), \ldots, h_N(\cdot))$ is K-strongly convex, where $K = \min_i\{K_i\}$. By a standard result in convex analysis [41], $C(\cdot)$ is $\frac{1}{K}$-Lipschitz continuous. Since v is lipschitz continuous by assumption, $v(Q(\cdot))$ is Lipschitz continuous. Consequently, standard results in differential equations ([11]) imply that a unique solution exists for the ODE.
3. [4] gives sufficient conditions that ensure a random trajectory to be an asymptotic pseudotrajectory of a semiflow almost surely. We next state one set of sufficient conditions directly in the current context. If for some $q \geq 2$, the following list of conditions are satisfied:

   (a) $\sup_t \mathbf{E}[(\|\xi^{t+1}\|^*)^q] < \infty$;

   (b) $\sum_{n=0}^{\infty} (\alpha^t)^{1+\frac{q}{2}} < \infty$;

   (c) $\sup_t \|\mathbf{x}^t\| < \infty$;

   then the affinely interpolated process $Y(t)$ is an asymptotic pseudotrajectory of the semi-flow ϕ induced by the ODE almost surely: $\forall T > 0, \lim_{t\to\infty} \sup_{0\leq h\leq T} \|Y(t+h), \phi_h(Y(t))\|^* = 0$, a.s.. Choose $q = 2$, the above conditions can be easily verified: (a) holds by the assumption on the martingale noise; (b) holds since α^t is square summable; (c) holds since the decision space $\mathcal{X}$ is compact. Therefore the claim follows.
4. By a well-known result in variational analysis ([40]), each $h_i(\cdot)$ is differentiable and

$$\frac{dh_i^*(y_i)}{dy_i} = C_i(y_i). \tag{C.12}$$

   Note further that since $\phi_t(\mathbf{y})$ is the solution to the ODE (under the initial condition $\mathbf{y}$), we have $\frac{d\phi_t(\mathbf{y})}{dt} = v(\mathbf{x}(t))$. Written out component-wise, we have

$$\frac{d(\phi_t(\mathbf{y}))_i}{dt} = v_i(\mathbf{x}(t)). \tag{C.13}$$

   We can thus compute the derivative of λ-Fenchel coupling as follows:

$$\frac{F^\lambda(\mathbf{x}^*, \phi_t(\mathbf{y}))}{dt} = \frac{\sum_{i=1}^{N} \lambda_i \left\{h_i(x_i^*) - (\phi_t(\mathbf{y}))_i x_i^* + h_i^*((\phi_t(\mathbf{y}))_i)\right\}}{dt} \tag{C.14}$$

$$= \sum_{i=1}^{N} \lambda_i \left\{-\frac{d(\phi_t(\mathbf{y}))_i}{dt} x_i^* + C_i(y_i)\frac{d(\phi_t(\mathbf{y}))_i}{dt}\right\} \tag{C.15}$$

$$= \sum_{i=1}^{N} \lambda_i \left\{-v_i(\mathbf{x}(t))x_i^* + v_i(\mathbf{x}(t))C_i(y_i)\right\} \tag{C.16}$$

$$= \sum_{i=1}^{N} \lambda_i v_i(\mathbf{x}(t))(x_i(t) - x_i^*) \leq 0, \tag{C.17}$$

where the second equality follows from Equation (C.12), the third equality follows from Equation (C.13), and the last inequality follows from $\mathbf{x}^*$ is λ-variationally stable. The monotonicity property therefore follows.

5. For any given $\epsilon > 0$, pick an $\hat{\epsilon} > 0$ per Claim 2 in Remark 4.2 such that $B(\mathbf{x}^*, \hat{\epsilon}) \subset \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$. By Equation (C.17), we have

$$\frac{F^\lambda(\mathbf{x}^*, \phi_t(\mathbf{y}))}{dt} = \sum_{i=1}^{N} \lambda_i v_i(\mathbf{x}(t))(x_i(t) - x_i^*) < 0, \forall \mathbf{x}(t) \neq \mathbf{x}^*.$$

Since $\mathcal{X} - B(\mathbf{x}^*, \hat{\epsilon})$ is a compact set and each $v_i(\cdot)$ is a continuous function, we have

$$\sum_{i=1}^{N} \lambda_i v_i(\mathbf{x}(t))(x_i(t) - x_i^*) \leq -a_{\tilde{\epsilon}}, \forall \mathbf{x}(t) \in \mathcal{X} - B(\mathbf{x}^*, \hat{\epsilon}), \tag{C.18}$$

for some positive constant $a_{\tilde{\epsilon}}$.

Starting at $\mathbf{y}$, by time s, there are two possibilities. The first possibility is that $\mathbf{x}(s) \in \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$. In this case, by definition,

$$F^\lambda(\mathbf{x}^*, \phi_s(\mathbf{y})) < \frac{\epsilon}{2}. \tag{C.19}$$

The second possibility is that $\mathbf{x}(s) \notin \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$. This implies that $\mathbf{x}(t) \notin B(\mathbf{x}^*, \hat{\epsilon}), \forall t \in [0, s]$, because otherwise, since $B(\mathbf{x}^*, \hat{\epsilon}) \subset \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$, it must be that $\mathbf{x}(s_0) \in \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$ for some $s_0 \in [0, s]$. This then implies that, by the monotonicity property established in Claim 4, $F^\lambda(\mathbf{x}^*, \phi_s(\mathbf{y})) \leq F^\lambda(\mathbf{x}^*, \phi_{s_0}(\mathbf{y}))$, thereby leading to $\mathbf{x}(s) \in \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$, a contradiction.

Since $\mathbf{x}(t) \notin \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2}), \forall t \in [0, s]$, we have $\mathbf{x}(t) \notin B(\mathbf{x}^*, \tilde{\epsilon}), \forall t \in [0, s]$, leading to that Equation C.18 holds for $t \in [0, s]$. Therefore, taking $s = \frac{\epsilon}{2a_{\tilde{\epsilon}}}$, we obtain:

$$F^\lambda(\mathbf{x}^*, \phi_s(\mathbf{y})) \leq F^\lambda(\mathbf{x}^*, \mathbf{y}) - a_{\tilde{\epsilon}} s = F^\lambda(\mathbf{x}^*, \mathbf{y}) - \frac{\epsilon}{2}. \tag{C.20}$$

Equation (C.19) and Equation (C.20) together establish that:

$$F^\lambda(\mathbf{x}^*, \phi_s(\mathbf{y})) \leq \max\{\frac{\epsilon}{2}, F^\lambda(\mathbf{x}^*, \mathbf{y}) - \frac{\epsilon}{2}\}.$$

6. Let $R = \sup_{\mathbf{x} \in \mathcal{X}} \|\mathbf{x}\|$, which is finite since $\mathcal{X}$ is compact. By the definition of dual norm and denote $\lambda_{\max} = \max_i \lambda_i$, we have

$$\sum_{i=1}^{N} \lambda_i \left\{(Y_i(t+h) - \phi_h^i(Y(t+h))\right\} (C_i(\phi_h^i(Y(t+h))) - x_i^*) \leq \tag{C.21}$$

$$\sum_{i=1}^{N} \lambda_i \|(Y_i(t+h) - \phi_h^i(Y(t+h))\|_i^* \|(C_i(\phi_h^i(Y(t+h))) - x_i^*)\| \leq \lambda_{\max} R \|Y(t+h) - \phi_h(t+h)\|^*, \quad \text{(C.22)}$$

where $\phi_h^i(Y(t+h))$ is the i-th component of $\phi_h(Y(t+h))$.

Fix some $T > 0$ and define $K_\lambda = \max_i \frac{\lambda_i}{K_i}$ and $\delta = \frac{\sqrt{(\lambda_{\max}R)^2 + 2\epsilon K_\lambda} - \lambda_{\max}R}{4K_\lambda}$. Per Claim 3, we have

$$\forall T > 0, \lim_{t\to\infty} \sup_{0 \leq h \leq T} \|Y(t+h), \phi_h(Y(t))\|^* = 0, \text{a.s.}.$$

Consequently, choose $\tau(\delta, T)$ such that $\|Y(t+h), \phi_h(Y(t))\|^* < \delta, \forall t \geq \tau$. Expanding λ-Fenchel coupling, we obtain that $\forall t \geq \tau$ and $\forall h \in [0, T]$:

$$F^\lambda(\mathbf{x}^*, Y(t+h)) = F^\lambda(\mathbf{x}^*, \phi_h(Y(t)) + Y(t+h) - \phi_h(Y(t))) \leq F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) \quad \text{(C.23)}$$

$$+ \sum_{i=1}^{N} \lambda_i \left\{(Y_i(t+h) - \phi_h^i(Y(t))\right\} (C_i(\phi_h^i(Y(t))) - x_i^*) + K_\lambda(\|Y(t+h) - \phi_h(Y(t))\|^*)^2 \quad \text{(C.24)}$$

$$\leq F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) + \lambda_{\max} R \|Y(t+h) - \phi_h(t)\|^* + K_\lambda(\|Y(t+h) - \phi_h(Y(t))\|^*)^2 \quad \text{(C.25)}$$

$$\leq F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) + \lambda_{\max} R\delta + K_\lambda \delta^2 \quad \text{(C.26)}$$

$$\leq F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) + \lambda_{\max} R \frac{\sqrt{(\lambda_{\max}R)^2 + 2\epsilon K_\lambda} - \lambda_{\max}R}{4K_\lambda} + K_\lambda(\frac{\sqrt{(\lambda_{\max}R)^2 + 2\epsilon K_\lambda} - \lambda_{\max}R}{4K_\lambda})^2 \quad \text{(C.27)}$$

$$< F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) + \lambda_{\max} R \frac{\sqrt{(\lambda_{\max}R)^2 + 2\epsilon K_\lambda} - \lambda_{\max}R}{2K_\lambda} + K_\lambda(\frac{(\sqrt{(\lambda_{\max}R)^2 + 2\epsilon K_\lambda} - \lambda_{\max}R)^2}{4K_\lambda^2} \quad \text{(C.28)}$$

$$= F^\lambda(\mathbf{x}^*, \phi_h(Y(t))) + \frac{\epsilon}{2}, \quad \text{(C.29)}$$

where the first inequality follows from Equation (C.22) and the last equality follows from straightforward algebraic verification. The claim is therefore established.

7. We start by fixing an arbitrary $\epsilon > 0$. Per Claim 5, there exists an $s > 0$ (depending on ϵ) such that Equation (4.7) holds. Set the horizon $T = s$. Per Claim 6, there exists a τ (depending on both ϵ and T) such that Equation (4.8) holds $\forall t \geq \tau$. Now, per Claim 1, $\mathbf{X}^t$ visits $\tilde{B}(\mathbf{x}^*, \delta)$ infinitely often[9]. Therefore, pick an integer $\tau_0 \geq \tau$ such that $\mathbf{X}^{\tau_0} \in \tilde{B}(\mathbf{x}^*, \frac{\epsilon}{2})$. With this choice of τ_0, we know that by definition of $\tilde{B}$,

$$F^\lambda(\mathbf{x}^*, Y(\tau_0)) < \frac{\epsilon}{2}. \quad \text{(C.30)}$$

Our goal is to establish that $F^\lambda(\mathbf{x}^*, Y(\tau_0 + h)) < \epsilon, \forall h \in [0, \infty)$. To that end, partition the time $[0, \infty)$ into disjoint time intervals $[0, T), [T, 2T), \ldots, [nT, (n+1)T), \ldots$.

Per Claim 4, the monotonicity property given in Equation (4.6) implies that:

$$F^\lambda(\mathbf{x}^*, \phi_h(Y(\tau_0))) \leq F^\lambda(\mathbf{x}^*, \phi_0(Y(\tau_0))) = F^\lambda(\mathbf{x}^*, Y(\tau_0)) < \frac{\epsilon}{2}, \forall h \geq 0, \quad \text{(C.31)}$$

[9] All the statements made here are true almost surely. We will omit repeatedly saying that for convenience. Alternatively, one can think of this as a path-by-path argument and each ensuing statement is made for a particular sample path that comes from a probability 1 space.

where the equality follows from the semi-group property of a semiflow.

Per Equation (4.8), for $h \in [0,T)$, we then have:

$$F^{\lambda}(\mathbf{x}^*, Y(\tau_0+h)) < F^{\lambda}(\mathbf{x}^*, \phi_h(Y(\tau_0))) + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \tag{C.32}$$

where the last inequality follows from Equation (C.31).

Now assume inductively that Equation (C.32) holds for every $h \in [nT,(n+1)T)$, where n is a non-negative integer. We then have $\forall h \in [nT,(n+1)T)$:

$$F^{\lambda}(\mathbf{x}^*, Y(\tau_0+T+h)) < F^{\lambda}(\mathbf{x}^*, \phi_T(Y(\tau_0+h))) + \frac{\epsilon}{2} \le \max\{\frac{\epsilon}{2}, F^{\lambda}(\mathbf{x}^*, Y(\tau_0+h)) - \frac{\epsilon}{2}\} + \frac{\epsilon}{2} \le \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \tag{C.33}$$

where the first inequality follows from Equation (4.8), the second inequality follows from Equation (4.7), and the third inequality follows from the induction hypothesis $F^{\lambda}(\mathbf{x}^*, Y(\tau_0+h)) < \epsilon$. Consequently, Equation (C.32) holds for every $h \in [(n+1)T,(n+2)T)$, thereby completing the induction and establishing that:

$$F^{\lambda}(\mathbf{x}^*, Y(\tau_0+h)) < \epsilon, \forall h \in [0,\infty).$$

■

D. Proof of Lemma 5.1 For Statement 1, first note that since α^t is square-summable, we have $\lim_{t\to\infty}(\alpha^t)^2 = 0$, thereby implying $\lim_{t\to\infty}\alpha^t = 0$. This further leads to $\max_{s\in[t-d+1,t]}\alpha^s \to 0$, as $t\to\infty$ for any fixed d.

Since the delay is bounded by d, $\mathcal{G}_i^t$ is non-empty at least once in every D consecutive rounds. When it is non-empty, the most delayed gradient must have occurred on or after round $t-d+1$. Putting these two pieces together, we conclude that on any round t, the most delayed gradient must have occurred on or after round $t-2d+1$. Consequently, we have $\min\mathcal{G}_i^t \ge t-2d+1$. This leads to

$$\lim_{t\to\infty}\sum_{s=\min\tilde{\mathcal{G}}_i^t}^{t}\alpha^s \le \lim_{t\to\infty}\sum_{s=t-2d+1}^{t}\alpha^s \le \lim_{t\to\infty} 2d \max_{s\in[t-2d+1,t]}\alpha^s = 0.$$

For Statement 2, take $\alpha^t = \frac{1}{t\log t\log\log t}$, which is obviously positive, non-increasing and square-summable. Since $\int_{s=4}^{t}\frac{1}{s\log s\log\log s}ds = \log\log\log t \to \infty$ as $t\to\infty$, α^t is not summable. Next, let $\tilde{\mathcal{G}}_i^t$ be given and let $\tilde{t}$ be the most recent round (up to and including t) such that $\mathcal{G}_i^{\tilde{t}}$ is not empty. This means:

$$\tilde{\mathcal{G}}_i^t = \mathcal{G}_i^{\tilde{t}}, \mathcal{G}_i^k = \emptyset, \forall k \in (\tilde{t},t]. \tag{D.1}$$

Note that since the gradient at time $\tilde{t}$ will be available at time $\tilde{t}+d_i^{\tilde{t}}-1$, it follows that

$$t-\tilde{t} \le d_i^{\tilde{t}}. \tag{D.2}$$

Note that this implies $\tilde{t} \to \infty$ as $t \to \infty$, because otherwise, $\tilde{t}$ is bounded, leading to the right-side $d_i^{\tilde{t}}$ being bounded, which contradicts to the left-side diverging to infinity.

Since $d_i^s = O(s \log s)$, it follows that $d_i^s \le K s \log s$ for some $K > 0$. Consequently, Equation D.2 implies: $t \le \tilde{t} + K\tilde{t}\log\tilde{t}$. Denote $s_{\min}^t = \min \tilde{\mathcal{G}}_i^t$, Equation D.1 implies that $s_{\min}^t = \min \mathcal{G}_i^{\tilde{t}}$, thereby yielding $s_{\min}^t + d_i^{s_{\min}^t} - 1 = \tilde{t}$. Therefore:

$$d_i^{s_{\min}^t} = \tilde{t} - s_{\min}^t + 1. \tag{D.3}$$

Equation (D.3) implies that $s_{\min}^t \to \infty$ as $t \to \infty$, because otherwise, the left-hand side of Equation (D.3) is bounded while the right-hand side goes to infinity (since $\tilde{t} \to \infty$ as $t \to \infty$ as established earlier).

With the above notation, it follows that:

$$\begin{aligned}
\lim_{t\to\infty} \sum_{s=\min \tilde{\mathcal{G}}_i^t}^{t} \alpha^s \le \lim_{t\to\infty} \sum_{s=s_{\min}^t}^{t} \alpha^s = \lim_{t\to\infty} \left\{ \sum_{s=s_{\min}^t}^{\tilde{t}} \alpha^s + \sum_{s=\tilde{t}+1}^{t} \alpha^s \right\} &\le \lim_{t\to\infty} \left\{ d_i^{s_{\min}^t} \alpha^{s_{\min}^t} + (\tilde{t}\log\tilde{t})\alpha^{\tilde{t}} \right\} && \text{(D.4)}\\
&= \lim_{t\to\infty} \left\{ \frac{d_i^{s_{\min}^t}}{(s_{\min}^t)\log(s_{\min}^t)\log\log(s_{\min}^t)} + \frac{K\tilde{t}\log\tilde{t}}{(\tilde{t}+1)\log(\tilde{t}+1)\log\log(\tilde{t}+1)} \right\} && \text{(D.5)}\\
&\le \lim_{t\to\infty} \left\{ \frac{K(s_{\min}^t)\log(s_{\min}^t)}{(s_{\min}^t)\log(s_{\min}^t)\log\log(s_{\min}^t)} + \frac{K\tilde{t}\log\tilde{t}}{(\tilde{t}+1)\log(\tilde{t}+1)\log\log(\tilde{t}+1)} \right\} && \text{(D.6)}\\
&\le \lim_{t\to\infty} \left\{ \frac{K}{\log\log(s_{\min}^t)} + \frac{K}{\log\log(\tilde{t}+1)} \right\} = 0. && \text{(D.7)}
\end{aligned}$$

where the second inequality follows from $\frac{1}{t \log t}$ is a decreasing sequence and the last equality follows from the assumption $d_i^s = o(s \log s)$ and that $s_{\min}^t \to \infty$ as $t \to \infty$. ■

E. Proof of Theorem 5.1 We first prove Claim 1. Using the same notation as Claim 1 in Remark 4.2, and per the DMD update, we have

$$\begin{aligned}
\|\mathbf{y}^{t+1} - \mathbf{y}^t\|^* = \sum_{i=1}^{N} \|y_i^{t+1} - y_i^t\|_i^* = \sum_{i=1}^{N} \|\frac{\alpha^t}{|\mathcal{G}_i^t|} \sum_{s\in\mathcal{G}_i^t} v_i(\mathbf{x}^s)\|_i^* &\le \alpha^t \sum_{i=1}^{N} \frac{1}{|\mathcal{G}_i^t|} \sum_{s\in\mathcal{G}_i^t} \|v_i(\mathbf{x}^s)\|_i^* && \text{(E.1)}\\
&\le \alpha^t \sum_{i=1}^{N} \frac{1}{|\mathcal{G}_i^t|} |\mathcal{G}_i^t| V_{\max}^i = \alpha^t V_{\max} && \text{(E.2)}
\end{aligned}$$

By definition, we can expand b_i^t as follows:

$$\begin{aligned}
b_i^t = \sum_{s\in\tilde{\mathcal{G}}_i^t} \frac{v_i(\mathbf{x}^s) - v_i(\mathbf{x}^t)}{|\tilde{\mathcal{G}}_i^t|} &\le \frac{1}{|\tilde{\mathcal{G}}_i^t|} \sum_{s\in\tilde{\mathcal{G}}_i^t} L\|\mathbf{x}^s - \mathbf{x}^t\| && \text{(E.3)}\\
&= \frac{1}{|\tilde{\mathcal{G}}_i^t|} \sum_{s\in\tilde{\mathcal{G}}_i^t} L\|C(\mathbf{y}^s) - C(\mathbf{y}^t)\| \le \frac{1}{|\tilde{\mathcal{G}}_i^t|} \sum_{s\in\tilde{\mathcal{G}}_i^t} \frac{L}{K}\|\mathbf{y}^s - \mathbf{y}^t\|^* && \text{(E.4)}\\
&\le \frac{L}{K}\frac{1}{|\tilde{\mathcal{G}}_i^t|} \sum_{s\in\tilde{\mathcal{G}}_i^t} \left\{ \|\mathbf{y}^s - \mathbf{y}^{s+1}\|^* + \|\mathbf{y}^{s+1} - \mathbf{y}^{s+2}\|^* + \cdots + \|\mathbf{y}^{t-1} - \mathbf{y}^t\|^* \right\} && \text{(E.5)}
\end{aligned}$$

$$\leq \frac{L}{K}\frac{1}{|\tilde{\mathcal{G}}_i^t|}\sum_{s\in\tilde{\mathcal{G}}_i^t}\left\{\alpha^s V_{\max}+\alpha^{s+1}V_{\max}+\cdots+\alpha^t V_{\max}\right\} \tag{E.6}$$

$$= \frac{LV_{\max}}{K}\frac{1}{|\tilde{\mathcal{G}}_i^t|}\sum_{s\in\tilde{\mathcal{G}}_i^t}\sum_{k=s}^{t}\alpha^k \leq \frac{LV_{\max}}{K}\frac{1}{|\tilde{\mathcal{G}}_i^t|}|\tilde{\mathcal{G}}_i^t|\sum_{s=\min\tilde{G}_i^t}^{t}\alpha^s \leq \frac{LV_{\max}}{K}\sum_{s=\min\tilde{G}_i^t}^{t}\alpha^s, \tag{E.7}$$

where the inequality in Equation E.3 follows from the fact that $v(\cdot)$ is L-Lipschitz continuous, the inequality in Equation E.4 follows from the fact that $C(\cdot)$ is $\frac{1}{K}$ -Lipschitz continuous, Equation E.6 follows from Equations E.1 and E.2 and the first inequality in Equation E.7 follows from that α^t's are non-negative. Finally, the above chain of inequalities immediately implies the result by noting that per Assumption 5.1,

$$\frac{LV_{\max}}{K}\sum_{s=\min\tilde{G}_i^t}^{t}\alpha^s \to 0.$$

To prove Claim 2, Fix an arbitrary $\epsilon > 0$. Assume for contradiction purposes that $\mathbf{x}^t$ only visits $B(\mathbf{x}^*,\epsilon)$ a finite number of times and hence let t^0-1 be the last time $\mathbf{x}^t$ is in $B(\mathbf{x}^*,\epsilon)$: $\forall t\geq t^0, \mathbf{x}^t\in\mathcal{X}-B(\mathbf{x}^*,\epsilon)$. Since $\mathcal{X}-B(\mathbf{x}^*,\epsilon)$ is a compact set and $v_i(\mathbf{x})$ is continuous in $\mathbf{x}$ and since by assumption $\sum_{i=1}^N \lambda_i v_i(\mathbf{x})(x_i-x_i^*) < 0, \forall\mathbf{x}\in\mathcal{X}, \mathbf{x}\neq\mathbf{x}^*$, it follows that there exists a $c_{\max}(\epsilon) < 0$ such that

$$\sum_{i=1}^N \lambda_i v_i(\mathbf{x})(x_i-x_i^*) \leq c_{\max}(\epsilon), \forall\mathbf{x}\in\mathcal{X}-B(\mathbf{x}^*,\epsilon). \tag{E.8}$$

Per Claim 1, $\lim_{t\to\infty}\mathbf{b}^t=\mathbf{0}$, therefore, $\|\mathbf{b}^t\|^*$ is bounded and we denote $B_{\max}\triangleq\max_t\|\mathbf{b}^t\|^*$. Next denote $R=\max_{\mathbf{x}\in\mathcal{X}}\|\mathbf{x}\|$, $\lambda_{\max}\triangleq\max_i\lambda_i$ and $\beta^t\triangleq\max_i\frac{(\alpha^t)^2\lambda_i}{2K_i}$ and note that $\sum_{t=1}^\infty\beta^t<\infty$. Using Lemma 4.1, we have $\forall t\geq t^0$:

$$F^\lambda(\mathbf{x}^*,\mathbf{y}^{t+1}) = F^\lambda(\mathbf{x}^*,\mathbf{y}^t+\alpha^t\{v(\mathbf{x}^t)+\mathbf{b}^t\}) \leq \tag{E.9}$$

$$F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\sum_{i=1}^N\lambda_i(\alpha^t\{v_i(\mathbf{x}^t)+b_i^t\})(C_i(y_i^t)-x_i^*)+\beta^t(\|v(\mathbf{x}^t)+\mathbf{b}^t\|^*)^2 = \tag{E.10}$$

$$F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\alpha^t\left\{\sum_{i=1}^N\lambda_i v_i(\mathbf{x}^t)(x_i^t-x_i^*)+\sum_{i=1}^N\lambda_i b_i^t(x_i^t-x_i^*)\right\}+\beta^t(\|v(\mathbf{x}^t)+\mathbf{b}^t\|^*)^2 \tag{E.11}$$

$$\leq F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\alpha^t\left\{c_{\max}(\epsilon)+\lambda_{\max}\|\mathbf{b}^t\|^*\|\mathbf{x}^t-\mathbf{x}^*\|\right\}+\beta^t\left\{2(\|v(\mathbf{x}^t)\|^*)^2+2(\|\mathbf{b}^t\|^*)^2\right\} \tag{E.12}$$

$$\leq F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\alpha^t\left\{c_{\max}(\epsilon)+\lambda_{\max}R\|\mathbf{b}^t\|^*\right\}+2\beta^t(V_{\max}^2+B_{\max}^2) \tag{E.13}$$

$$\leq F^\lambda(\mathbf{x}^*,\mathbf{y}^{t^0})+(\sum_{k=t^0}^{t}\alpha^i)\left\{b_{\max}(\epsilon)+\lambda_{\max}R\frac{\sum_{k=t^0}^t\alpha^k\|\mathbf{b}^k\|^*}{\sum_{k=t^0}^t\alpha^k}\right\}+2(\sum_{k=t^0}^{t}\beta^k)(V_{\max}^2+B_{\max}^2), \tag{E.14}$$

where Equation (E.12) follows from Equation (E.8) and Equation (E.14) follows from telescoping.

Consequently, since $\lim_{t\to\infty}\mathbf{b}^t=\mathbf{0}$ and $\sum_{t=0}^\infty\alpha^t=\infty$, it follows that:

$$\frac{\sum_{k=t^0}^t\alpha^k\|\mathbf{b}^k\|^*}{\sum_{k=t^0}^t\alpha^k}\to 0, \text{ as } t\to\infty,$$

therefore implying that:

$$(\sum_{k=t^0}^{t}\alpha^i)\left\{b_{\max}(\epsilon)+\lambda_{\max}R\frac{\sum_{k=t^0}^{t}\alpha^k\|\mathbf{b}^k\|^*}{\sum_{k=t^0}^{t}\alpha^k}\right\}\to-\infty,\ \text{as}\ t\to\infty.$$

Since $\sum_{k=t^0}^{\infty}\beta^k<\infty$, Equation (E.14) implies that $F^\lambda(\mathbf{x}^*,\mathbf{y}^t)\to-\infty$, which contradicts the first statement in Lemma 4.1. The claim therefore follows.

Finally, we establish Claim 3. Fix a given $\delta>0$. Since β^t is summable and α^t is not summable but square summable, it follows that $\beta^t\to0,\alpha^t\to0,\frac{\alpha^t}{\beta^t}\to\infty$ as $t\to\infty$. There, denote

1. $T^1(\delta)=\arg\min_t\{t\mid\beta^s<\frac{\delta}{8(V_{\max}^2+B_{\max}^2)},\forall s\ge t\}$.
2. $T^2(\delta)=\arg\min_t\{t\mid c_{\max}(\epsilon(\frac{\delta}{2}))<-2\lambda_{\max}R\|\mathbf{b}^s\|^*,\forall s\ge t\}$.
3. $T^3(\delta)=\arg\min_t\{t\mid\alpha^s<\frac{\delta}{4\lambda_{\max}RB_{\max}},\forall s\ge t\}$.
4. $T^4(\delta)=\arg\min_t\{t\mid\frac{\alpha^s}{\beta^s}>\frac{4(V_{\max}^2+B_{\max}^2)}{-c_{\max}(\epsilon(\frac{\delta}{2}))},\forall s\ge t\}$.

We have $T^1(\delta)<\infty,T^2(\delta)<\infty$ (since $\lim_{t\to\infty}\|\mathbf{b}^t\|^*=0$ and note that $c_{\max}(\epsilon(\frac{\delta}{2}))<0$ by definition), $T^3(\delta)<\infty,T^4(\delta)<\infty$. Take

$$T(\delta)=\max(T^1(\delta),T^2(\delta),T^3(\delta),T^4(\delta)\}.$$

Now take any $t\ge T(\delta)$. It suffices to show that if $\mathbf{x}^t\in\tilde{B}(\mathbf{x}^*,\delta)$, then $\mathbf{x}^{t+1}\in\tilde{B}(\mathbf{x}^*,\delta)$. To see that this statement holds, let $\mathbf{x}^t\in\tilde{B}(\mathbf{x}^*,\delta)$, which implies that $F^\lambda(\mathbf{x}^*,\mathbf{y}^t)<\delta$.

Now there are two possibilities:

1. Possibility 1: $\mathbf{x}^t\in B(\mathbf{x}^*,\epsilon(\frac{\delta}{2}))$.
2. Possibility 2: $\mathbf{x}^t\in\tilde{B}(\mathbf{x}^*,\delta)-B(\mathbf{x}^*,\epsilon(\frac{\delta}{2}))$.

Under Possibility 1, it follows

$$\begin{aligned}
F^\lambda(\mathbf{x}^*,\mathbf{y}^{t+1}) &\le F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\alpha^t\sum_{i=1}^{N}\lambda_i\{v_i(\mathbf{x}^t)+b_i^t\}(x_i^t-x_i^*)+\beta^t(\|v(\mathbf{x}_t)+\mathbf{b}^t\|^*)^2 &\text{(E.15)}\\
&\le F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\alpha^t\sum_{i=1}^{N}\lambda_i b_i^t(x_i^t-x_i^*)+\beta^t\left\{2(\|v(\mathbf{x}_t)\|^*)^2+2(\|\mathbf{b}^t\|^*)^2\right\} &\text{(E.16)}\\
&\le F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\alpha^t\lambda_{\max}RB_{\max}+2\beta^t(V_{\max}^2+B_{\max}^2) &\text{(E.17)}\\
&< F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\frac{\delta}{4\lambda_{\max}RB_{\max}}\lambda_{\max}RB_{\max}+\frac{2\delta}{8(V_{\max}^2+B_{\max}^2)}(V_{\max^2}+B_{\max}^2) &\text{(E.18)}\\
&\le\frac{\delta}{2}+\frac{\delta}{4}+\frac{\delta}{4}=\delta, &\text{(E.19)}
\end{aligned}$$

where the second inequality follows from λ-variational stability and the last inequality follows from the fact that $\mathbf{x}^t\in B(\mathbf{x}^*,\epsilon(\frac{\delta}{2}))\subset\tilde{B}(\mathbf{x}^*,\frac{\delta}{2})$ per Claim 2.

Under Possibility 2, it follows from Equation E.13 that

$$F^\lambda(\mathbf{x}^*,\mathbf{y}^{t+1})\le F^\lambda(\mathbf{x}^*,\mathbf{y}^t)+\alpha^t\left\{c_{\max}(\epsilon(\frac{\delta}{2}))+\lambda_{\max}R\|\mathbf{b}^t\|^*\right\}+2\beta^t(V_{\max}^2+B_{\max}^2) \tag{E.20}$$

$$\leq F^{\lambda}(\mathbf{x}^*,\mathbf{y}^t) + 2\beta^t(V_{\max}^2 + B_{\max}^2)\left\{\frac{\alpha^t}{\beta^t}\frac{c_{\max}(\epsilon(\frac{\delta}{2})) + \lambda_{\max}R\|\mathbf{b}^t\|^*}{2(V_{\max}^2 + B_{\max}^2)} + 1\right\} \tag{E.21}$$

$$\leq F^{\lambda}(\mathbf{x}^*,\mathbf{y}^t) + 2\beta^t(V_{\max}^2 + B_{\max}^2)\left\{\frac{\alpha^t}{\beta^t}\frac{c_{\max}(\epsilon(\frac{\delta}{2}))}{4(V_{\max}^2 + B_{\max}^2)} + 1\right\} \tag{E.22}$$

$$< F^{\lambda}(\mathbf{x}^*,\mathbf{y}^t) < \epsilon, \tag{E.23}$$

where the second inequality follows from $\lambda_{\max}R\|\mathbf{b}^t\|^* < -\frac{1}{2}c_{\max}(\epsilon(\frac{\delta}{2}))$ since $t \geq T^2(\delta)$ and the second-to-last inequality follows from $\frac{\alpha^t}{\beta^t}\frac{c_{\max}(\epsilon(\frac{\delta}{2}))}{4(V_{\max}^2+B_{\max}^2)} + 1 < 0$ since $t \geq T^4(\delta)$. ■