A. Motivating examples

A.1. A conceptual approach for non-convex games

We extend here the solution concept for non-convex m-player games with *smoothed local equilibrium* proposed by (Hazan et al., 2017) to be valid in our *stochastic composite* game setup. We emphasize that the guarantees we present in this section are also valid for when each player only has access to a stochastic first-order oracle, making it closer to practical use.

To model the multi-player setting, consider m problems of the form (P) corresponding to each of the players, where every player i observes her online part of her objective function

$$f_t^i(\mathbf{z}) := f(\mathbf{x}_t^1, \ldots, \mathbf{x}_t^{i-1}, \mathbf{z}, \ldots, \mathbf{x}_t^m), \tag{A.1}$$

and then decides on $\mathbf{x}_{t+1}^i$.

It is sometimes desirable to induce specific properties in the game, this is fully supported by our model (P). For example: (i) to incur risk-aversion, the regularizer of each player g^i can be chosen accordingly, e.g., L^1-norm; (ii) to ensure a meaningful solution, such as the *global minimax point* condition defined by (Jin et al., 2019), restriction of the decision set to a compact convex set can be applied.

In our non-convex setting, obtaining the global measure of Nash equilibrium is beyond reach, and may not exist at all; see Proposition 6 in (Jin et al., 2019). Thus, a different, local, measure for equilibrium is essential. This topic is already receiving much attention in the literature, for example, for a multi-player non-convex games, (Pang & Scutari, 2011) proposes the local *quasi-Nash equilibrium* measure defined using KKT conditions. In the case of a (two-players) minmax game (e.g., GANs) for example, local measure is defined as the stationarity (first-order condition) of both players in the very recent (Jin et al., 2019, Nouiehed et al., 2019). For additional details, we refer to the works alluded above.

We follow the smoothed local equilibrium approach in Section 6 in (Hazan et al., 2017) , and extend it here to our composite model. This approach comes naturally from assuming that the players take into account the behavior history of the other players. Other than that, it allows for a tractable notion of equilibrium.

The *smoothed local equilibrium* is defined for the joint cost function (A.1) as follows, where $S_{t,w}^i(\mathbf{x}) = \frac{1}{w}\sum_{j=t-w+1}^{t} f_j^i(\mathbf{x})$.

Definition A.1 (smoothed local equilibrium)**.** *Let $\eta > 0, w \geq 1$. For an m-player iterative game with cost functions as in (A.1), a joint strategy at iteration $t > 0$, $(\mathbf{x}_t^1, \ldots, \mathbf{x}_t^{i-1}, \mathbf{x}_t^i, \ldots, \mathbf{x}_t^m)$, is an ε-(η, w) smoothed local equilibrium with respect to the history of w-iterates if:*

$$\left\| \mathcal{P}_\eta^{g^i}(\mathbf{x}_t^i; \nabla S_{t,w}^i(\mathbf{x}_t^i)) \right\|^2 \leq \varepsilon \qquad \forall i \in [m]. \tag{A.2}$$

Denote by $\mathrm{Reg}_w^i(T)$ the local regret (cf. Eq. (5)) of the i-th player. We first derive a guarantee for when each player has access to a perfect first-order oracle (using Theorem 3.1).

Theorem A.1 (Equilibrium with perfect oracle)**.** *Let the sequence $(\mathbf{x}_t^1, \ldots, \mathbf{x}_t^{i-1}, \mathbf{x}_t^i, \ldots, \mathbf{x}_t^m)$, $t = 1, \ldots, T$ be generated by running Algorithm 1 for all players simultaneously with input $\eta > 0$ and $w = \lceil 2k(\delta^2 + c)\varepsilon^{-1/2} \rceil$, given that the online function is determined by (A.1). Suppose that $V_w[T] \leq cT$ for some $c > 0$. Then there exists $t^* \geq w$ such that (A.2) holds true.*

Proof. There exists a $t^* \geq w$ such that

$$\begin{aligned}\sum_{i=1}^{k} \left\| \mathcal{P}_\eta^{g^i}(\mathbf{x}_{t^*}^i; \nabla f_{t^*}^i(\mathbf{x}_{t^*}^i)) \right\|^2 &\leq \frac{1}{T-w} \sum_{i=1}^{k} \sum_{t=w}^{T} \left\| \mathcal{P}_\eta^{g^i}(\mathbf{x}_t^i; \nabla f_t^i(\mathbf{x}_t^i)) \right\|^2 \\ &\leq \frac{1}{T-w} \sum_{i=1}^{k} \mathrm{Reg}_w^i(T).\end{aligned}$$

Thus, if each player has access to a perfect first-order oracle and $V_w[T] \leq cT$, then by Theorem 3.1

$$\sum_{i=1}^{k} \left\| \mathcal{P}_\eta^{g^i}(\mathbf{x}_{t^*}^i; \nabla f_{t^*}^i(\mathbf{x}_{t^*}^i)) \right\|^2 \leq \frac{1}{T-w} \sum_{i=1}^{k} \frac{2}{w^2} \left(T\delta^2 + V_w[T] \right) \leq \frac{2kT(\delta^2+c)}{(T-w)w^2}.$$

Consequently, by setting $T = w^2$ and $w = \lceil 2k(\delta^2+c)\varepsilon^{-1/2} \rceil$ we obtain

$$\sum_{i=1}^{k} \left\| \mathcal{P}_\eta^{g^i}(\mathbf{x}_{t^*}^i; \nabla f_{t^*}^i(\mathbf{x}_{t^*}^i)) \right\|^2 \leq \frac{2k(\delta^2+c)}{(w-1)w} \leq \varepsilon,$$

as desired. □

By similar arguments, we derive the guarantees for when players have access via a stochastic first-order oracle, only now we utilize Theorem 4.2; we implicitly assume here that all the conditions of Theorem 4.2 are satisfied.

Theorem A.2 (Equilibrium with stochastic first-order oracle)**.** *Suppose that the sequence* $(\mathbf{x}_t^1, \ldots, \mathbf{x}_t^{i-1}, \mathbf{x}_t^i, \ldots, \mathbf{x}_t^m)$, $t = 1, \ldots, T$ *is generated by running Algorithm 1 for all players simultaneously with input* $\eta > 0$ *and* $w = \lceil \frac{2k\left(\delta^2+7\sigma^2+6c\right)}{\sqrt{\varepsilon}} \rceil$, *given that the online function is determined by* (A.1)*. Suppose that* $V_w[T] \leq cT$ *for some* $c > 0$*. Then there exists* $t^* \geq w$ *such that* (A.2) *holds true in expectation.*

Proof. There exists a $t^* \geq w$ such that

$$\begin{aligned}\sum_{i=1}^{k} \left\| \mathcal{P}_\eta^{g^i}(\mathbf{x}_{t^*}^i; \nabla f_{t^*}^i(\mathbf{x}_{t^*}^i)) \right\|^2 &\leq \frac{1}{T-w} \sum_{i=1}^{k} \sum_{t=w}^{T} \left\| \mathcal{P}_\eta^{g^i}(\mathbf{x}_t^i; \nabla f_t^i(\mathbf{x}_t^i)) \right\|^2 \\ &\leq \frac{1}{T-w} \sum_{i=1}^{k} \mathrm{Reg}_w^i(T).\end{aligned}$$

Thus, by taking expectation and using the fact that $V_w[T] \leq cT$, we obtain from Theorem 4.2 that

$$\begin{aligned}\sum_{i=1}^{k} \mathbb{E} \left\| \mathcal{P}_\eta^{g^i}(\mathbf{x}_{t^*}^i; \nabla f_{t^*}^i(\mathbf{x}_{t^*}^i)) \right\|^2 &\leq \frac{1}{T-w} \sum_{i=1}^{k} 2 \left(\left(\frac{T}{w^2} \right) \left(\delta^2 + 7\sigma^2 \right) + \frac{6}{w^2} V_w[T] \right) \\ &= \frac{2kT\left(\delta^2+7\sigma^2+6c\right)}{(T-w)w^2}.\end{aligned}$$

Consequently, by setting $T = w^2$ and $w = \lceil \frac{2k\left(\delta^2+7\sigma^2+6c\right)}{\sqrt{\varepsilon}} \rceil$ we obtain

$$\sum_{i=1}^{k} \mathbb{E} \left\| \mathcal{P}_\eta^{g^i}(\mathbf{x}_{t^*}^i; \nabla f_{t^*}^i(\mathbf{x}_{t^*}^i)) \right\|^2 \leq \frac{2k\left(\delta^2+7\sigma^2+6c\right)}{(w-1)w} \leq \varepsilon,$$

as desired. □

A.2. The online traffic assignment problem

Referring to (Bertsekas & Gallager, 1992) and (Shakkottai & Srikant, 2008) for an introduction to the topic, the key objective in traffic assignment problems is the optimal allocation of traffic over a given network with variable traffic inflows. To state this precisely, consider a directed multi-graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$. Embedded in this network is a set of origin-destination (O/D) pairs $(o_i, d_i) \in \mathcal{V} \times \mathcal{V}$, $i \in \mathcal{N} = \{1, 2, \ldots, N\}$, each routing a (possibly random) quantity of traffic from o_i to d_i via a set of paths $\mathcal{P}_i$ in $\mathcal{G}$. Writing $\mathcal{K}_i = \Delta(\mathcal{P}_i)$ for the simplex spanned by $\mathcal{P}_i$, a *trafic allocation vector* for the i-th O/D pair is defined to be a vector $\mathbf{x}_i = (x_{i,p_i})_{p_i \in \mathcal{P}_i} \in \mathcal{K}_i$ with each x_{i,p_i} denoting the fraction of the traffic of the i-th O/D pair that is routed via p_i. Then, collectively, a *traffic allocation profile* is an ensemble $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$ of such vectors belonging to the product space $\mathcal{K} = \prod_i \mathcal{K}_i$.

In this general context, the cost (delay, latency, etc.) of routing a certain amount of traffic via a given path p_i is a function $\ell_{p_i}(\mathbf{x};\boldsymbol{\lambda})$ of the chosen allocation profile $\mathbf{x}\in\mathcal{K}$ and the set of *traffic demands* $\boldsymbol{\lambda}=(\lambda_1,\dots,\lambda_N)$ of each O/D pair.These demands are typically assumed to follow a non-stationary probability distribution (e.g., accounting for diurnal variations in an urban traffic network), leading to the *online traffic assignment problem* (OnTAP) stated below:

$$\begin{aligned}\text{minimize}\quad & \ell_t(\mathbf{x})=\sum_{i\in\mathcal{N}}\sum_{p_i\in\mathcal{P}_i}x_{i,p_i}\ell_{p_i}(\mathbf{x};\boldsymbol{\lambda}_t)+\mu\|\mathbf{x}\|_1\\ \text{subject to}\quad & \mathbf{x}\in\mathcal{K}.\end{aligned}\tag{OnTAP}$$

In the above formulation, the sparsity-inducing L^1 term is intended to "robustify" solutions by minimizing the overall number of paths employed. The cost functions ℓ_{p_i} are sums of positive polynomials (described below), so they are smooth over $\mathcal{K}$ but may otherwise be non-convex. As such, (OnTAP) can be cast in the framework of (P) by taking $g=\delta_{\mathcal{K}}+\mu\|\cdot\|_1$ with $\delta_{\mathcal{K}}$ denoting the convex indicator of $\mathcal{K}$.

Let us now detail the definition of the cost functions ℓ_{p_i} for (OnTAP). For simplicity, we will suppress the O/D index $i\in\mathcal{N}$, i.e., we will treat the problem as a single-O/D one; this doesn't play a major role in the sequel and only serves to make the notation ligther.

To begin, given a traffic allocation vector $\mathbf{x}\in\mathcal{K}$ and an inflow rate λ, the *traffic load* carried by edge $e\in\mathcal{E}$ is defined to be the total traffic routed via the edge in question, i.e.,

$$y_e\equiv y_e(\mathbf{x};\lambda)=\lambda\sum_{p:p\ni e}x_p,\tag{A.3}$$

and we write $\mathbf{y}=(y_e)_{e\in\mathcal{E}}$ for the corresponding *load profile* on the network. Given all this, the cost (delay, latency, etc.) experienced by an infinitesimal traffic element traversing edge e is given by a non-decreasing continuous *cost function* $\ell_e\colon\mathbb{R}_+\to\mathbb{R}_+$; more precisely, if $\mathbf{y}\equiv\mathbf{y}(\mathbf{x};\lambda)$ is the load profile induced by a traffic allocation profile $\mathbf{x}\in\mathcal{K}$ and a traffic demand λ, the incurred cost on edge $e\in\mathcal{E}$ is simply $\ell_e(y_e)$. Hence, the associated cost for path $p\in\mathcal{P}$ will be

$$\ell_p(\mathbf{x};\lambda)\equiv\sum_{e\in p}\ell_e(y_e(\mathbf{x};\lambda))=\sum_{e\in p}\ell_e\left(\lambda\sum_{p':p'\ni e}x_{p'}\right).\tag{A.4}$$

In urban traffic networks, the cost functions ℓ_e are typically non-decreasing positive polynomials fitted to appropriate statistical data; a common choice is the so-called "quartic BPR" model $\ell_e(y_e)=a_e+b_ey_e^4$ of the US Bureau of Public Roads (BPR), but this is beyond our scope.

B. Regretfulness when $w=1$

For completeness, we provide a simple example for when the "standard" stationarity measure Eq. (4), obtained from the local regret when $w=1$, fails. The bound $O(T/w^2)$ established by (Hazan et al., 2017) (cf. Theorem 2.7) is proved via a similar example.

Suppose that $g(x)=\delta_{[-1,1]}(x)$ is the indicator function for the set $[-1,1]$, and that

$$f_t(x)=\begin{cases}-x & \text{with probability } 0.5,\\ x & \text{with probability } 0.5.\end{cases}$$

Then

$$\mathbb{E}\,\mathrm{Reg}_1(T)=\mathbb{E}\sum_{t=1}^{T}\left\|\mathcal{P}_\eta^g(\mathbf{x}_t;\nabla f_t(\mathbf{x}_t)\right\|^2\geq O(T).$$

C. Fundamental Properties

Throughout the analysis, we utilize fundamental properties of the proximal mapping operator for L-smooth functions. The descent lemma (see e.g., Lemma 5.7 in (Beck, 2017)) and the sufficient decrease property of the proximal gradient operator (cf. Lemma 10.4 in (Beck, 2017)) are given as follows.

Lemma C.1 (Descent lemma)**.** *Let $f:\mathbb{R}^n\to(-\infty,\infty]$ be an L-smooth function ($L\geq 0$) over a convex set $C\subseteq\mathbb{R}^n$. Then for any $\mathbf{x},\mathbf{y}\in C$, $f(\mathbf{y})\leq f(\mathbf{x})+\langle\nabla f(\mathbf{x}),\mathbf{y}-\mathbf{x}\rangle+\frac{L}{2}\|\mathbf{x}-\mathbf{y}\|^2$.*

Lemma C.2 (Sufficient decrease property)**.** *Let $h:\mathbb{R}^n\to\mathbb{R}\cup\{\infty\}$ be a proper, convex, l.s.c function, and $f:\mathbb{R}^n\to(-\infty,\infty)$ be an L-smooth function ($L\geq 0$) over $\operatorname{dom}h$. Then for any $\mathbf{x}\in\int\operatorname{dom}h$ and $\eta\in(0,L/2)$ it holds for $\mathbf{x}^+=\operatorname{prox}_{\eta h}(\mathbf{x}-\eta\nabla f(\mathbf{x}))$ that*

$$h(\mathbf{x})+f(\mathbf{x})-h(\mathbf{x}^+)-f(\mathbf{x}^+)\geq\eta\left(1-\frac{\eta L}{2}\right)\left\|\frac{1}{\eta}\left(\mathbf{x}^+-\mathbf{x}\right)\right\|^2.$$

We also use a trivial, yet essential, property of the proximal gradient mapping.

Lemma C.3. *For any $\mathbf{x},\mathbf{d}_1,\mathbf{d}_2\in\mathbb{R}^n$ and $\eta>0$ it holds that*

$$\left\|\mathcal{P}_\eta^g(\mathbf{x};\mathbf{d}_1+\mathbf{d}_2)\right\|\leq\left\|\mathcal{P}_\eta^g(\mathbf{x};\mathbf{d}_1)\right\|+\|\mathbf{d}_2\|.$$

Proof. By the triangle inequality and non-expensiveness of the prox operator (cf. (**?**)Theorem 6.42]B17)

$$\begin{aligned}\left\|\mathcal{P}_\eta^g(\mathbf{x};\mathbf{d}_1+\mathbf{d}_2)\right\|-\left\|\mathcal{P}_\eta^g(\mathbf{x};\mathbf{d}_1)\right\|&\leq\left\|\mathcal{P}_\eta^g(\mathbf{x};\mathbf{d}_1+\mathbf{d}_2)-\mathcal{P}_\eta^g(\mathbf{x};\mathbf{d}_1)\right\|\\&\leq\frac{1}{\eta}\left\|(\mathbf{x}-\eta(\mathbf{d}_1+\mathbf{d}_2))-(\mathbf{x}-\eta\mathbf{d}_1)\right\|=\|\mathbf{d}_2\|.\end{aligned}$$

□

D. Proofs of Section 3

Proof of Theorem 3.1. Note that

$$S_t(\mathbf{x})=\frac{1}{w}\sum_{i=t-w+1}^{t}f_i(\mathbf{x})=S_{t-1}(\mathbf{x})+\frac{1}{w}(f_t(\mathbf{x})-f_{t-w}(\mathbf{x})).$$

Setting $h_1=S_{t-1}$, $h_2=\frac{1}{w}(f_t-f_{t-w})$, applying Lemma C.3 and the triangle inequality yields

$$\begin{aligned}\|\mathcal{P}(\mathbf{x}_t;\nabla S_t(\mathbf{x}_t))\|&=\|\mathcal{P}(\mathbf{x}_t;\nabla(h_1+h_2)(\mathbf{x}_t))\|\\&\leq\|\mathcal{P}(\mathbf{x}_t;\nabla S_{t-1}(\mathbf{x}_t))\|+\frac{1}{w}\|\nabla f_t(\mathbf{x}_t)-\nabla f_{t-w}(\mathbf{x}_t)\|.\end{aligned}$$

By the definition of the method, i.e. $\|\mathcal{P}(\mathbf{x}_t;\nabla S_{t-1}(\mathbf{x}_t))\|\leq\frac{\delta}{w}$, we thus have that

$$\|\mathcal{P}(\mathbf{x}_t;\nabla S_t(\mathbf{x}_t))\|\leq\frac{\delta}{w}+\frac{1}{w}\|\nabla f_t(\mathbf{x}_t)-\nabla f_{t-w}(\mathbf{x}_t)\|,\qquad\forall t\in[T],$$

and consequently, for any $t\in[T]$,

$$\|\mathcal{P}(\mathbf{x}_t;\nabla S_t(\mathbf{x}_t))\|^2\leq\frac{2\delta^2}{w^2}+\frac{2}{w^2}\|\nabla f_t(\mathbf{x}_t)-\nabla f_{t-w}(\mathbf{x}_t)\|^2.$$

Summing over $t=1,\ldots,T$, then results with

$$\operatorname{Reg}_w(T)=\sum_{t=1}^{T}\|\mathcal{P}(\mathbf{x}_t;\nabla S_t(\mathbf{x}_t))\|^2\leq\frac{2}{w^2}\left(T\delta^2+V_w[T]\right).$$

□

To prove that Algorithm 1 executes $O(w^2)$ proximal gradient calls, we require a sufficient decrease property that is given next.

Lemma D.1 (Sufficient decrease property)**.** *Let $t \in [T]$, and let τ_t be the number of times step 3 is executed at the t-th iteration. Then*

$$S_{t,w}(\mathbf{x}_t) + g(\mathbf{x}_t) - S_{t,w}(\mathbf{x}_{t+1}) - g(\mathbf{x}_{t+1}) \geq \tau_t \left(\eta - \frac{\eta^2 L}{2}\right) \frac{\delta^2}{w^2}, \qquad \forall t \in [T].$$

Proof. Denote the sequence generated in the inner loop at time $t \in [T]$ by

$$\mathbf{y}_t^0 = \mathbf{x}_t, \quad \mathbf{y}_t^{k+1} = \operatorname*{arg\,min}_{\mathbf{z}\in\mathbb{R}^n} g(\mathbf{z}) + \langle \nabla S_t(\mathbf{y}_t^k), \mathbf{z} - \mathbf{y}_t^k \rangle + \frac{1}{2\eta}\|\mathbf{z} - \mathbf{y}_t^k\|^2, \qquad k = 0, 1, \ldots, \tau_t - 1,$$

and note that $\mathbf{y}_t^{\tau_t} = \mathbf{x}_{t+1}$. By the sufficient decrease property of the proximal gradient operator (cf. Lemma C.2), and the stopping criteria of the inner loop, we have that for all $k = 0, 1, \ldots, \tau_t - 1$

$$S_t(\mathbf{y}_t^k) + g(\mathbf{y}_t^k) - S_t(\mathbf{y}_t^{k+1}) - g(\mathbf{y}_t^{k+1}) \geq \left(\eta - \frac{\eta^2 L}{2}\right) \left\|\mathcal{P}(\mathbf{y}_t^k; \nabla S_t(\mathbf{y}_t^k))\right\|^2 \geq \left(\eta - \frac{\eta^2 L}{2}\right) \frac{\delta^2}{w^2}. \tag{D.1}$$

Summing (D.1) over $k = 0, 1, \ldots, \tau_t - 1$, then yields

$$\begin{aligned} S_t(\mathbf{x}_t) + g(\mathbf{x}_t) - S_t(\mathbf{x}_{t+1}) - g(\mathbf{x}_{t+1}) &= S_t(\mathbf{y}_t^0) + g(\mathbf{y}_t^0) - S_t(\mathbf{y}_t^{\tau_t}) - g(\mathbf{y}_t^{\tau_t}) \\ &\geq \tau_t \left(\eta - \frac{\eta^2 L}{2}\right) \frac{\delta^2}{w^2} \end{aligned}$$

which completes our proof. □

We will now bound the number of proximal gradient iterations executed by Algorithm 1.

Proof of Theorem 3.2. Recall that $S_0(\mathbf{x}_0) \equiv 0$, and $S_t(\mathbf{x}) = \frac{1}{w}(f_t(\mathbf{x}) - f_{t-w}(\mathbf{x})) + S_{t-1}(\mathbf{x})$. Thus,

$$\begin{aligned} S_T(\mathbf{x}_T) &= \sum_{t=1}^{T} (S_t(\mathbf{x}_t) - S_{t-1}(\mathbf{x}_{t-1})) \\ &= \frac{1}{w}\sum_{t=1}^{T} (f_t(\mathbf{x}_t) - f_{t-w}(\mathbf{x}_t)) + \sum_{t=2}^{T} (S_{t-1}(\mathbf{x}_t) - S_{t-1}(\mathbf{x}_{t-1})) \\ &\leq \frac{2MT}{w} + \sum_{t=2}^{T} (S_{t-1}(\mathbf{x}_t) - S_{t-1}(\mathbf{x}_{t-1})), \end{aligned}$$

where the last inequality follows from our blanket assumptions. Consequently, by Lemma D.1, we have that

$$\begin{aligned} S_T(\mathbf{x}_T) + g(\mathbf{x}_T) - g(\mathbf{x}_1) &\leq \frac{2MT}{w} + \sum_{t=2}^{T} (S_{t-1}(\mathbf{x}_t) + g(\mathbf{x}_t) - S_{t-1}(\mathbf{x}_{t-1}) - g(\mathbf{x}_{t-1})) \\ &\leq \frac{2MT}{w} - \sum_{t=1}^{T-1} \tau_t \left(\eta - \frac{\eta^2 L}{2}\right) \frac{\delta^2}{w^2} \\ &\leq \frac{2MT}{w} - \tau \left(\eta - \frac{\eta^2 L}{2}\right) \frac{\delta^2}{w^2}, \end{aligned}$$

where the last inequality uses $\tau = \sum_{t=1}^{T-1} \tau_t$. On the other hand, by our blanket assumptions,

$$S_T(\mathbf{x}_T) = \frac{1}{w} \sum_{i=T-w+1}^{T} f_i(\mathbf{x}_i) \geq -M.$$

By combining both sides we obtain that

$$-M \leq g(\mathbf{x}_1) - g(\mathbf{x}_T) + \frac{2MT}{w} - \tau \left(\eta - \frac{\eta^2 L}{2}\right) \frac{\delta^2}{w^2},$$

and the desired immediately follows from the nonnegativity of g:

$$\tau \leq \frac{g(\mathbf{x}_1) - g(\mathbf{x}_T) + M + \frac{2MT}{w}}{\left(\eta - \frac{\eta^2 L}{2}\right)\frac{\delta^2}{w^2}} \leq \frac{2Tw(g(\mathbf{x}_1) + 3M)}{(2-\eta L)\,\eta\delta^2}. \qquad \square$$

We conclude with the implication of our guarantees to the stochastic offline setting.

Proof of Corollary 3.1. From the choice of t_*, Jensen's inequality, and Theorem 3.1, we have that

$$\begin{aligned}
\mathbb{E}_{t_*}\left(\|\nabla f(\mathbf{x}_{t_*})\|^2\right) &= \frac{1}{T-w}\sum_{t=w}^{T}\|\mathbb{E}\left(\nabla f_t(\mathbf{x}_t)\right)\|^2 \\
&= \frac{1}{T-w}\sum_{t=w}^{T}\left\|\mathbb{E}\left(\frac{1}{w}\sum_{i=t-w+1}^{t}\nabla f_i(\mathbf{x}_t)\right)\right\|^2 \\
&\leq \frac{1}{T-w}\sum_{t=w}^{T}\mathbb{E}\left(\left\|\frac{1}{w}\sum_{i=t-w+1}^{t}\nabla f_i(\mathbf{x}_t)\right\|^2\right) \\
&\leq \frac{1}{T-w}\mathbb{E}\left(\mathrm{Reg}_w(T)\right) \\
&\leq \frac{2}{(T-w)w^2}\left(T\delta^2 + V_w[T]\right).
\end{aligned}$$

Plugging the parameters' values $T = 2w$, $w = \lceil\sqrt{\frac{2(\delta^2+c)}{\varepsilon}}\rceil$, and $V_w[T] = cT$, we immediately obtain that

$$\mathbb{E}\left(\|\nabla f(\mathbf{x}_{t_*})\|^2\right) \leq \frac{2}{(T-w)w^2}\left(\delta^2 T + V_w[T]\right) \leq \frac{4}{w^2}\left(\delta^2 + c\right) \leq \varepsilon.$$

Once again, by plugging the parameters' values we obtain from (10) in Theorem 3.2 that

$$\tau \leq \frac{2w^2(g(\mathbf{x}_1) + 2M)}{(2-\eta L)\,\eta\delta^2} \propto O(\varepsilon^{-1}).$$

Since for each proximal gradient update the algorithm computes w gradient samples (for each function sampled in the time-window), the SFO complexity is

$$\tau w \propto O(\varepsilon^{-3/2}). \qquad \square$$

E. Proofs of Section 4

Before proceeding to the stochastic analysis, we make some notational conventions for the sake of readability: $S_t \equiv S_{t,w}$, $T(\mathbf{x};\mathbf{d}) \equiv T_\eta^{f,g}(\mathbf{x};\mathbf{d})$, and $\mathcal{P}(\mathbf{x};\mathbf{d}) \equiv \mathcal{P}_\eta^g(\mathbf{x};\mathbf{d})$. Additionally, we set $\mathbf{y}_t^k = \mathbf{y}_t^{\tau_t}$ for all $k \geq \tau_t$; this means that $\mathbf{y}_t^k = \mathbf{y}_t^{k+1}$ if and only if $k \geq \tau_t$.

The forthcoming analysis of Algorithm 2 requires delicate treatment of what is known, and what is not, at specific moments during the run. To avoid confusion, we state explicitly what is included in the algorithm's natural filtration at time $t \geq 1$ and at each inner iteration $k \geq 1$, thus extending on our original description.

Definition E.1 (Filtration)**.** *For all $t \geq 1$, the filtration $\mathcal{F}_t$ includes all gradient feedback up to, but not including, the execution of step 2 at stage t. In particular, it includes f_t, $\mathbf{x}_t$ and $\tilde{\nabla} S_{t-1}(\mathbf{x}_t)$, but it does not include $\tilde{\nabla} f_t(\mathbf{x}_t)$.*

For all $t \geq 1$ and all $k \geq 1$, the filtration $\mathcal{F}_{t,k}$ includes all gradient feedback up to, but not including, the execution of the k-th iteration of step 5(b) at time t. In particular, it contains $\mathcal{F}_t$, and includes $\mathbf{y}_t^k, G_t^k$, and $\mathbf{y}_t^{k+1}$, but it does not include $\{\tilde{\nabla} f_i(\mathbf{y}_t^{k+1})\}_{i=t-w}^{t}$, G_t^{k+1}.

We will utilize two trivial technical corollaries of Definition 4.1 given next.

Corollary E.1. *Let* $\mathbf{x} \in \mathbb{R}^n$*, then*

$$\mathbb{E}\left(\|\mathcal{S}_\sigma(\mathbf{x};\omega,h)-\nabla h(\mathbf{x})\|\right)^2 \le \mathbb{E}\left(\|\mathcal{S}_\sigma(\mathbf{x};\omega,h)-\nabla h(\mathbf{x})\|^2\right) \le \sigma^2. \tag{E.1}$$

Lemma E.1. *Let* $\mathbf{x} \in \mathbb{R}^n$ *and* $h_i : \mathbb{R}^n \to \mathbb{R}$ *for any* $i = 1, 2, \ldots, w$*. Then*

$$\mathbb{E}\left(\left\|\frac{1}{w}\sum_{i=1}^{w}\mathcal{S}_\sigma(\mathbf{x};\omega,h_i)-\frac{1}{w}\sum_{i=1}^{w}\nabla h_i(\mathbf{x})\right\|^2\right) \le \sigma^2.$$

Proof. Follows from Jensen's inequality. □

The following technical lemma is of key importance in the analysis ahead.

Lemma E.2. *Let* $t \in [T]$ *and* $k \ge 2$*. It holds that*

$$\mathbb{E}\left(\langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \mathbf{y}_t^{k+1} - \mathbf{y}_t^k\rangle | \mathcal{F}_{t,k-1}\right) \ge -\frac{\eta\sigma^2}{w^2}.$$

Proof. Define the full gradient proximal gradient by $\hat{\mathbf{y}}_t^k = T_\eta^g(\mathbf{y}_t^k; \nabla S_{t,w}(\mathbf{y}_t^k))$, and note that

$$\begin{aligned}\langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \mathbf{y}_t^{k+1} - \mathbf{y}_t^k\rangle &= \langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \mathbf{y}_t^{k+1} - \hat{\mathbf{y}}_t^k\rangle + \langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \hat{\mathbf{y}}_t^k - \mathbf{y}_t^k\rangle \\ &\ge -\|G_t^k - \nabla S_t(\mathbf{y}_t^k)\|\|\mathbf{y}_t^{k+1} - \hat{\mathbf{y}}_t^k\| + \langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \hat{\mathbf{y}}_t^k - \mathbf{y}_t^k\rangle,\end{aligned} \tag{E.2}$$

where the last inequality follows from Cauchy-Schwartz inequality. By the nonexpansivity of the prox operator (**?**)Theorem 6.42]B17 we have that

$$\|\mathbf{y}_t^{k+1} - \hat{\mathbf{y}}_t^k\| \le \|\mathbf{y}_t^k - \eta G_t^k - \mathbf{y}_t^k + \eta\nabla S_t(\mathbf{y}_t^k)\| = \eta\|G_t^k - \nabla S_t(\mathbf{y}_t^k)\|,$$

meaning that

$$-\|G_t^k - \nabla S_t(\mathbf{y}_t^k)\|\|\mathbf{y}_t^{k+1} - \hat{\mathbf{y}}_t^k\| \ge -\eta\|G_t^k - \nabla S_t(\mathbf{y}_t^k)\|^2. \tag{E.3}$$

Plugging (E.3) to (E.2) then implies that

$$\langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \mathbf{y}_t^{k+1} - \mathbf{y}_t^k\rangle \ge -\eta\|G_t^k - \nabla S_t(\mathbf{y}_t^k)\|^2 + \langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \hat{\mathbf{y}}_t^k - \mathbf{y}_t^k\rangle. \tag{E.4}$$

Noting that by Definition 4.1

$$\mathbb{E}\left(\langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \hat{\mathbf{y}}_t^k - \mathbf{y}_t^k\rangle | \mathcal{F}_{t,k-1}\right) = 0,$$

we obtain, from taking expectation on (E.4) and using Lemma E.1, that

$$\mathbb{E}\left(\langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \mathbf{y}_t^{k+1} - \mathbf{y}_t^k\rangle | \mathcal{F}_{t,k-1}\right) \ge -\frac{\eta\sigma^2}{w^2}.$$ □

We can now embark on proving our claims stated in Section 4.

Proof of Theorem 4.1. Recall that $\mathbf{y}_t^1 = \mathbf{x}_t, \mathbf{y}_t^{\tau_t} = \mathbf{x}_{t+1}$, and

$$\mathbf{y}_t^{k+1} = \arg\min_{\mathbf{z}\in\mathbb{R}^n} g(\mathbf{z}) + \langle G_t^k, \mathbf{z} - \mathbf{y}_t^k\rangle + \frac{1}{2\eta}\|\mathbf{z} - \mathbf{y}_t^k\|^2, \qquad k \in [\tau_t - 1].$$

Denote $h_t^k := S_t(\mathbf{y}_t^k) + g(\mathbf{y}_t^k)$. By combining the descent lemma (cf. Lemma C.1), the definition of $\mathbf{y}_t^{k+1}$, and the stopping criteria of the inner loop, we have that for any $k \in [\tau_t - 1]$ (assuming that $\mathcal{F}_t$ is given),

$$\begin{aligned}h_t^k - h_t^{k+1} &\ge \langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \mathbf{y}_t^{k+1} - \mathbf{y}_t^k\rangle + \frac{1}{2}\left(\eta - \eta^2 L\right)\left\|\mathcal{P}(\mathbf{y}_t^k; G_t^k)\right\|^2 \\ &\ge \langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \mathbf{y}_t^{k+1} - \mathbf{y}_t^k\rangle + \frac{1}{2}\left(\eta - \eta^2 L\right)\frac{\delta^2}{w^2}.\end{aligned}$$

Applying expectation to the latter, using the law of total expection (tower rule), and invoking Lemma E.2 and relation (7), we obtain that for any $k \in [\tau_t - 1]$ it holds that

$$\begin{aligned}\mathbb{E}\left(h_t^k - h_t^{k+1}\right) &\geq \mathbb{E}\left(\langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \mathbf{y}_t^{k+1} - \mathbf{y}_t^k\rangle\right) + \frac{1}{2}\left(\eta - \eta^2 L\right)\frac{\delta^2}{w^2} \\ &\geq \frac{2}{w^2}\left(\eta\left(1-\eta L\right)\delta^2 - 2\sigma^2\right) > 0.\end{aligned}$$

Set $\alpha := 2\left(\eta\left(1-\eta L\right)\delta^2 - 2\sigma^2\right)/w^2 > 0$. From the former, by using the law of total expectation, for any $K \geq 1$ we have that

$$\begin{aligned}h_t^1 + M \geq \mathbb{E}\left(h_t^1 - h_t^{K+1}\right) &= \mathbb{E}\left(\sum_{k=1}^{K}(h_t^k - h_t^{k+1})\right) \\ &= \sum_{k=1}^{K}\mathbb{E}\left(h_t^k - h_t^{k+1}\right) \\ &= \sum_{k=1}^{K}\left(\mathbb{E}\left(h_t^k - h_t^{k+1}|\tau_t \geq k+1\right)\mathbb{P}(\tau_t \geq k+1) + 0\cdot\mathbb{P}(\tau_t \leq k)\right) \\ &\geq \alpha\sum_{k=1}^{K}\mathbb{P}(\tau_t > k) \\ &\geq \alpha\sum_{k=1}^{K}\mathbb{P}(\tau_t > K) = \alpha K\mathbb{P}(\tau_t > K).\end{aligned}$$

Consequently, we must have that τ_t is almost surely finite, which in turn implies that τ must be almost surely finite as it is the finite sum of almost surely finite variables. □

Let us now establish the local regret bound stated in Theorem 4.2.

Proof of Theorem 4.2. Recall that

$$\mathrm{Reg}_w(T) = \sum_{t=1}^{T}\left\|\mathcal{P}(\mathbf{x}_t; \nabla S_t(\mathbf{x}_t))\right\|^2 = \sum_{t=1}^{T}\frac{1}{\eta^2}\left\|\mathbf{x}_t - T(\mathbf{x}_t; \nabla S_t(\mathbf{x}_t))\right\|^2. \tag{E.5}$$

By simple algebra,

$$\left\|\mathbf{x}_t - T(\mathbf{x}_t; \nabla S_t(\mathbf{x}_t))\right\|^2 \leq 2\left\|\mathbf{x}_t - T(\mathbf{x}_t; \tilde{\nabla} S_t(\mathbf{x}_t))\right\|^2 + 2\left\|T(\mathbf{x}_t; \tilde{\nabla} S_t(\mathbf{x}_t)) - T(\mathbf{x}_t; \nabla S_t(\mathbf{x}_t))\right\|^2. \tag{E.6}$$

Using the nonexpansivity of the prox operator (**?**)Theorem 6.42]B17 we have that

$$\begin{aligned}\left\|T(\mathbf{x}_t; \tilde{\nabla} S_t(\mathbf{x}_t)) - T(\mathbf{x}_t; \nabla S_t(\mathbf{x}_t))\right\|^2 &\leq \left\|\mathbf{x}_t - \eta\tilde{\nabla} S_t(\mathbf{x}_t) - \mathbf{x}_t + \eta\nabla S_t(\mathbf{x}_t)\right\|^2 \\ &= \eta^2\left\|\tilde{\nabla} S_t(\mathbf{x}_t) - \nabla S_t(\mathbf{x}_t)\right\|^2.\end{aligned}$$

Subsequently, using the law of total expectation and Lemma E.1, we obtain the relation

$$\begin{aligned}\mathbb{E}\left(\left\|T(\mathbf{x}_t; \tilde{\nabla} S_t(\mathbf{x}_t)) - T(\mathbf{x}_t; \nabla S_t(\mathbf{x}_t))\right\|^2\right) &= \mathbb{E}\left[\mathbb{E}\left(\left\|T(\mathbf{x}_t; \tilde{\nabla} S_t(\mathbf{x}_t)) - T(\mathbf{x}_t; \nabla S_t(\mathbf{x}_t))\right\|^2 |\mathcal{F}_t\right)\right] \\ &\leq \eta^2\mathbb{E}\left[\mathbb{E}\left(\left\|\tilde{\nabla} S_t(\mathbf{x}_t) - \nabla S_t(\mathbf{x}_t)\right\|^2 |\mathcal{F}_t\right)\right] \leq \frac{\eta^2\sigma^2}{w^2}.\end{aligned}$$

Then, plugging the latter to the expected value of (E.6) yields

$$\mathbb{E}\left(\|\mathbf{x}_t - T(\mathbf{x}_t; \nabla S_t(\mathbf{x}_t))\|^2\right) \leq 2\eta^2 \mathbb{E}\left(\left\|\mathcal{P}(\mathbf{x}_t; \tilde{\nabla} S_t(\mathbf{x}_t))\right\|^2\right) + \frac{2\eta^2\sigma^2}{w^2}.$$

Thus,

$$\mathbb{E}\left(\mathrm{Reg}_w(T)\right) \leq 2\sum_{t=1}^{T}\left[\mathbb{E}\left(\left\|\mathcal{P}(\mathbf{x}_t; \tilde{\nabla} S_{t,w}(\mathbf{x}_t))\right\|^2\right) + \frac{\sigma^2}{w^2}\right]. \tag{E.7}$$

Setting $G_1 = \tilde{\nabla} S_{t-1}(\mathbf{x}_t)$, $G_2 = \frac{1}{w}(\tilde{\nabla} f_t(\mathbf{x}_t) - \tilde{\nabla} f_{t-w}(\mathbf{x}_t))$, and applying Lemma C.3 yields

$$\begin{aligned}\left\|\mathcal{P}(\mathbf{x}_t; \tilde{\nabla} S_t(\mathbf{x}_t))\right\| = \|\mathcal{P}(\mathbf{x}_t; G_1 + G_2)\| &\leq \left\|\mathcal{P}(\mathbf{x}_t; \tilde{\nabla} S_{t-1}(\mathbf{x}_t))\right\| + \frac{1}{w}\left\|\tilde{\nabla} f_t(\mathbf{x}_t) - \tilde{\nabla} f_{t-w}(\mathbf{x}_t)\right\| \\ &\leq \frac{\delta}{w} + \frac{1}{w}\left\|\tilde{\nabla} f_t(\mathbf{x}_t) - \tilde{\nabla} f_{t-w}(\mathbf{x}_t)\right\|, \end{aligned} \tag{E.8}$$

where the last inequality follows from the termination rule of the inner loop. Therefore,

$$\left\|\mathcal{P}(\mathbf{x}_t; \tilde{\nabla} S_t(\mathbf{x}_t))\right\|^2 \leq \frac{2}{w^2}\left(\delta^2 + \left\|\tilde{\nabla} f_t(\mathbf{x}_t) - \tilde{\nabla} f_{t-w}(\mathbf{x}_t)\right\|^2\right).$$

Using the triangle inequality and the relation $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$, yields that

$$\begin{aligned}&\left\|\tilde{\nabla} f_t(\mathbf{x}_t) - \tilde{\nabla} f_{t-w}(\mathbf{x}_t)\right\|^2 \leq \\ &3\left\|\tilde{\nabla} f_t(\mathbf{x}_t) - \nabla f_t(\mathbf{x}_t)\right\|^2 + 3\left\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-w}(\mathbf{x}_t)\right\|^2 + 3\left\|\nabla f_{t-w}(\mathbf{x}_t) - \tilde{\nabla} f_{t-w}(\mathbf{x}_t)\right\|^2.\end{aligned}$$

Applying expectation, from the law of total expectation together with Definition 4.1, we obtain that

$$\begin{aligned}\mathbb{E}\left[\left\|\tilde{\nabla} f_t(\mathbf{x}_t) - \nabla f_t(\mathbf{x}_t)\right\|^2\right] &= \mathbb{E}\left[\mathbb{E}\left(\left\|\tilde{\nabla} f_t(\mathbf{x}_t) - \nabla f_t(\mathbf{x}_t)\right\|^2 | \mathcal{F}_t\right)\right] &\leq \frac{\sigma^2}{w^2}, \\ \mathbb{E}\left[\left\|\nabla f_{t-w}(\mathbf{x}_t) - \tilde{\nabla} f_{t-w}(\mathbf{x}_t)\right\|^2\right] &= \mathbb{E}\left[\mathbb{E}\left(\left\|\nabla f_{t-w}(\mathbf{x}_t) - \tilde{\nabla} f_{t-w}(\mathbf{x}_t)\right\|^2 | \mathcal{F}_{t-w}, \mathbf{x}_t\right)\right] &\leq \frac{\sigma^2}{w^2}.\end{aligned}$$

Thus, $\mathbb{E}\left(\left\|\tilde{\nabla} f_t(\mathbf{x}_t) - \tilde{\nabla} f_{t-w}(\mathbf{x}_t)\right\|^2\right) \leq \frac{6\sigma^2}{w^2} + 3\mathbb{E}\left(\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-w}(\mathbf{x}_t)\|^2\right)$, and consequently

$$\begin{aligned}\mathbb{E}\left(\left\|\mathcal{P}(\mathbf{x}_t; \tilde{\nabla} S_t(\mathbf{x}_t))\right\|^2\right) &\leq \frac{2}{w^2}\left(\delta^2 + \mathbb{E}\left(\left\|\tilde{\nabla} f_t(\mathbf{x}_t) - \tilde{\nabla} f_{t-w}(\mathbf{x}_t)\right\|^2\right)\right) \\ &\leq \frac{2}{w^2}\left(\delta^2 + \frac{6\sigma^2}{w^2} + 3\mathbb{E}\left(\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-w}(\mathbf{x}_t)\|^2\right)\right).\end{aligned}$$

Summing over $t \in [T]$ and plugging $V_w[T]$ defined in (6) then yields

$$\sum_{t=1}^{T}\mathbb{E}\left(\left\|\mathcal{P}(\mathbf{x}_t; \tilde{\nabla} S_t(\mathbf{x}_t))\right\|^2\right) \leq 2\left(\delta^2 + \frac{6\sigma^2}{w^2}\right)\left(\frac{T}{w^2}\right) + \frac{6}{w^2}V_w[T].$$

Finally, plugging the latter into (E.7), and recalling that $w \geq 1$, results with the desired bound. □

Finally, we prove the bound on the number of SFO calls, as stated by Theorem 4.3.

Proof of Theorem 4.3. Denote $h_t^k := S_t(\mathbf{y}_t^k) + g(\mathbf{y}_t^k)$. By combining the descent lemma (cf. Lemma C.1), the definition of the sequence $\{\mathbf{y}_t^k\}_{k\geq 1}$, Young's inequality, and the stopping criteria of the inner loop, we have that for any $K \geq 1$

(assuming that $\mathcal{F}_t$ is given)

$$h_t^1 - h_t^{K+1} = \sum_{k=1}^{K}(h_t^k - h_t^{k+1}) \geq \sum_{k=1}^{\min\{K,\tau_t\}} \left(\langle G_t^k - \nabla S_t(\mathbf{y}_t^k), \mathbf{y}_t^{k+1} - \mathbf{y}_t^k \rangle + \frac{1-\eta L}{2\eta} \left\| \mathbf{y}_t^{k+1} - \mathbf{y}_t^k \right\|^2 \right)$$
$$\geq \frac{1}{2} \sum_{k=1}^{\min\{K,\tau_t\}} \left(- \left\| G_t^k - \nabla S_t(\mathbf{y}_t^k) \right\|^2 - \left\| \mathbf{y}_t^{k+1} - \mathbf{y}_t^k \right\|^2 + \frac{1-\eta L}{\eta} \left\| \mathbf{y}_t^{k+1} - \mathbf{y}_t^k \right\|^2 \right).$$

Hence, by Assumption 1 and the stopping condition of the inner loop, we obtain

$$h_t^1 - h_t^{K+1} \geq \frac{1}{2w^2} \sum_{k=1}^{\min\{K,\tau_t\}} \left(-\sigma^2 + (1 - \eta(L+1))\eta\delta^2 \right) = \frac{(1-\eta(L+1))\eta\delta^2 - \sigma^2}{2w^2} \min\{K, \tau_t\} > 0.$$

Recall that $S_{0,w}(\mathbf{x}_0) \equiv 0$, and $S_t(\mathbf{x}) = \frac{1}{w}(f_t(\mathbf{x}) - f_{t-w}(\mathbf{x})) + S_{t-1}(\mathbf{x})$. Using the previous derivations for $t-1$ (setting $K = \tau_{t-1}$ and noting that $h_{t-1}^{\tau_{t-1}+1} = h_{t-1}^{\tau_{t-1}}$), we have that

$$S_{t-1}(\mathbf{x}_t) + g(\mathbf{x}_t) - S_{t-1}(\mathbf{x}_{t-1}) - g(\mathbf{x}_{t-1}) = h_{t-1}^{\tau_{t-1}} - h_{t-1}^1 \leq -\tau_{t-1} \frac{(1-\eta(L+1))\eta\delta^2 - \sigma^2}{2w^2}. \tag{E.9}$$

Thus, since

$$S_T(\mathbf{x}_T) = \sum_{t=1}^{T}(S_t(\mathbf{x}_t) - S_{t-1}(\mathbf{x}_{t-1})) = \sum_{t=1}^{T} \left(\frac{1}{w}(f_t(\mathbf{x}_t) - f_{t-w}(\mathbf{x}_t)) + S_{t-1}(\mathbf{x}_t) - S_{t-1}(\mathbf{x}_{t-1}) \right)$$
$$= \frac{2MT}{w} + \sum_{t=2}^{T} \left(S_{t-1}(\mathbf{x}_t) - S_{t-1}(\mathbf{x}_{t-1}) \right),$$

we have from our blanket assumptions and relation (E.9), that

$$S_T(\mathbf{x}_T) \leq g(\mathbf{x}_1) - g(\mathbf{x}_T) + \frac{2MT}{w} - \tau \frac{(1-\eta(L+1))\eta\delta^2 - \sigma^2}{2w^2}.$$

On the other hand, again by our blanket assumptions, $S_T(\mathbf{x}_T) = \frac{1}{w}\sum_{i=T-w+1}^{T} f_i(\mathbf{x}_i) \geq -M$. By combining both sides, we obtain that

$$-M \leq g(\mathbf{x}_1) - g(\mathbf{x}_T) + \frac{2MT}{w} - \tau \frac{(1-\eta(L+1))\eta\delta^2 - \sigma^2}{2w^2},$$

and the bound on τ immediately follows due to the nonnegativity of g. Finally, the desired bound on the SFO oracle calls follows from the fact that the inner loop makes $O(w)$ SFO calls per loop. □

E.1. Implications to Offline Stochastic Optimization

Next we establish our derivations in the offline scenario described in Section 4.2.

Proof of Theorem 4.4. Note that $f_t \equiv f$ for any $t \in [T]$ implies that $\nabla S_{t,w}(\mathbf{x}) \equiv \nabla f(\mathbf{x})$. From Theorem 4.2 and the choice of t_* we have that

$$\mathbb{E}\left(\left\| \mathcal{P}(\mathbf{x}_{t_*}; \nabla f(\mathbf{x}_{t_*})) \right\|^2 \right) = \frac{1}{T-w} \mathbb{E}\left(\sum_{t=w}^{T} \left\| \mathcal{P}(\mathbf{x}_t; \nabla f(\mathbf{x}_t)) \right\|^2 \right)$$
$$\leq \frac{1}{T-w} \mathbb{E}\left(\mathrm{Reg}_w(T) \right)$$
$$\leq \frac{2}{(T-w)w^2} \left(\left(\delta^2 + 7\sigma^2\right) T + 6V_w[T] \right).$$

□

Proof of Corollary 4.1. From Theorem 4.4 we immediately obtain that

$$\frac{2}{(T-w)w^2}\left(\left(\delta^2+7\sigma^2\right)T+6V_w[T]\right)=\frac{4w}{w^3}\left(\delta^2+7\sigma^2+c\right)\leq\varepsilon.$$

The bound $O(M\sigma\varepsilon^{-3/2})$ is obtained by plugging the assumed values of w, T, and δ^2, to (10) in Theorem 4.3:

$$w\tau\leq\frac{2\eta w^3(g(\mathbf{x}_1)+3M)}{(1-\eta(L+1))\delta^2-\eta\sigma^2}=\frac{2w^3(g(\mathbf{x}_1)+3M)}{\sigma^2}\propto O(M\sigma\varepsilon^{-3/2}),$$

where we used the fact that w is $O(\sigma/\sqrt{\varepsilon})$. □