A. Proof of Lemma 3

Lemma 3. *The stability parameter of a performance measure $\Psi(\cdot)$ can be written as $\delta(\epsilon) \leq L_\Psi \cdot \epsilon$ iff its sufficient dual region is bounded in a ball of radius $\Theta\left(L_\Psi\right)$.*

Proof. Let us denote primal variables using the notation $\mathbf{x} = (u, v)$ and dual variables using the notation $\boldsymbol{\theta} = (\alpha, \beta)$. The proof follows from the fact that any value of $\boldsymbol{\theta}$ for which $\Psi^*(\boldsymbol{\theta}) = -\infty$ can be safely excluded from the sufficient dual region.

For proving the result in one direction suppose Ψ is stable with $\delta(\epsilon) = L\epsilon$ for some $L > 0$. Now consider some $\boldsymbol{\theta} \in \mathbb{R}^2$ such that $\|\boldsymbol{\theta}\|_2 \geq L$. Now set $\mathbf{x}_C = -C \cdot \boldsymbol{\theta}$. Then we have

$$
\begin{aligned}
\Psi^*(\boldsymbol{\theta}) &= \inf_{\mathbf{x}} \left\{ \langle \boldsymbol{\theta}, \mathbf{x} \rangle - \Psi(\mathbf{x}) \right\} \\
&\leq \inf_{C>0} \left\{ \langle \boldsymbol{\theta}, \mathbf{x}_C \rangle - \Psi(\mathbf{x}_C) \right\} \\
&= \inf_{C>0} \left\{ -C \left\|\boldsymbol{\theta}\right\|_2^2 - \Psi(\mathbf{x}_C) \right\} \\
&\leq \inf_{C>0} \left\{ -C \left\|\boldsymbol{\theta}\right\|_2^2 - \Psi(\mathbf{0}) + CL \left\|\boldsymbol{\theta}\right\|_\infty \right\} \\
&\leq \inf_{C>0} \left\{ -C \left\|\boldsymbol{\theta}\right\|_2^2 - \Psi(\mathbf{0}) + CL \left\|\boldsymbol{\theta}\right\|_2 \right\} \\
&= \inf_{C>0} \left\{ -C \left\|\boldsymbol{\theta}\right\|_2 \left(\left\|\boldsymbol{\theta}\right\|_2 - L \right) \right\} - \Psi(\mathbf{0}) \\
&\leq \inf_{C>0} \left\{ -C \left\|\boldsymbol{\theta}\right\|_2 - \Psi(\mathbf{0}) \right\} \\
&= -\infty
\end{aligned}
$$

Thus, we can conclude that no dual vector with norm greater than L can be a part of the sufficient dual region. This shows that the sufficient dual region is bounded inside a ball of radius L. For proving the result in the other direction, suppose the dual sufficient region is indeed bounded in a ball of radius R. Consider two points $\mathbf{x}_1, \mathbf{x}_2$ such that

$$
\begin{aligned}
\boldsymbol{\theta}_1^* &= \arg\min_{\boldsymbol{\theta} \in \mathcal{A}_\Psi} \left\{ \langle \boldsymbol{\theta}, \mathbf{x}_1 \rangle - \Psi^*(\boldsymbol{\theta}) \right\} \\
\boldsymbol{\theta}_2^* &= \arg\min_{\boldsymbol{\theta} \in \mathcal{A}_\Psi} \left\{ \langle \boldsymbol{\theta}, \mathbf{x}_2 \rangle - \Psi^*(\boldsymbol{\theta}) \right\}
\end{aligned}
$$

Now define $f(\boldsymbol{\theta}, \mathbf{x}) := \langle \boldsymbol{\theta}, \mathbf{x} \rangle - \Psi^*(\boldsymbol{\theta})$ so that, by the above definition, $f(\boldsymbol{\theta}_1^*, \mathbf{x}_1) = \Psi(\mathbf{x}_1)$ and $f(\boldsymbol{\theta}_2^*, \mathbf{x}_2) = \Psi(\mathbf{x}_2)$. Now we have

$$
\begin{aligned}
\Psi(\mathbf{x}_1) &= f(\boldsymbol{\theta}_1^*, \mathbf{x}_1) \leq f(\boldsymbol{\theta}_2^*, \mathbf{x}_1) \\
&\leq f(\boldsymbol{\theta}_2^*, \mathbf{x}_2) + \left| \langle \boldsymbol{\theta}_2^*, \mathbf{x}_1 - \mathbf{x}_2 \rangle \right| \\
&= \Psi(\mathbf{x}_2) + \left| \langle \boldsymbol{\theta}_2^*, \mathbf{x}_1 - \mathbf{x}_2 \rangle \right| \\
&\leq \Psi(\mathbf{x}_2) + R \left\| \mathbf{x}_1 - \mathbf{x}_2 \right\|_2,
\end{aligned}
$$

where the fourth step follows from the norm bound on $\boldsymbol{\theta}_2^*$. Similarly we have

$$
\Psi(\mathbf{x}_2) \leq \Psi(\mathbf{x}_1) + R \left\| \mathbf{x}_1 - \mathbf{x}_2 \right\|_2
$$

This establishes the result. □

B. Proof of Theorem 4

Theorem 4. *Suppose we are given a stream of random samples $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_T, y_T)$ drawn from a distribution $\mathcal{D}$ over $\mathcal{X} \times \mathcal{Y}$. Let $\Psi(\cdot)$ be a concave, Lipschitz link function. Let Algorithm 1 be executed with a dual feasible set $\mathcal{A} \supseteq \mathcal{A}_\Psi$,*

$\eta_t = 1/\sqrt{t}$ and $\eta'_t = 1/\sqrt{t}$. Then, the average model $\overline{\mathbf{w}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{w}_t$ output by the algorithm satisfies, with probability at least $1-\delta$,

$$\mathcal{P}_\Psi(\overline{\mathbf{w}}) \geq \sup_{\mathbf{w}^*\in\mathcal{W}} \mathcal{P}_\Psi(\mathbf{w}^*) - \delta_\Psi\left(\sqrt{\frac{2B_r^2}{T}\log\frac{1}{\delta}}\right) - \left(L_\Psi^2 + 4B_r^2\right)\frac{1}{2\sqrt{T}} - \left(L_\Psi^2 L_r^2 + R_{\mathcal{W}}^2\right)\frac{1}{2\sqrt{T}} - \sqrt{\frac{2L_\Psi^2 B_r^2}{T}\log\frac{1}{\delta}}.$$

Proof. For this proof we shall assume that Ψ is L_Ψ-Lipschitz so that its sufficient dual region can be bounded by an application of Lemma 3. Notice that the updates for (α,β) can be written as follows:

$$(\alpha_{t+1},\beta_{t+1}) \leftarrow \Pi_{\mathcal{A}_\Psi}\left((\alpha_t,\beta_t) - \eta_t\nabla_{(\alpha,\beta)}\ell_t^d(\alpha_t,\beta_t)\right),$$

where

$$\ell_t^d(\alpha,\beta) = \begin{cases} \alpha r^+(\mathbf{w}_t;\mathbf{x}_t,y_t) - \Psi^*(\alpha,\beta) & \text{if } y_t > 0 \\ \beta r^-(\mathbf{w}_t;\mathbf{x}_t,y_t) - \Psi^*(\alpha,\beta) & \text{if } y_t < 0 \end{cases}$$

which can be interpreted as simple gradient descent with ℓ_t. Moreover, since Ψ^* is concave, ℓ_t^d is convex with respect to (α,β) for every t. Note that the terms $r^+(\mathbf{w}_t;\mathbf{x}_t,y_t)$ and $r^-(\mathbf{w}_t;\mathbf{x}_t,y_t)$ do not involve α,β and hence act as arbitrary bounded positive constants for this part of the analysis.

Note that by Lemma 3, we have the radius of $\mathcal{A}_\Psi$ bounded by L_Ψ. Also, since Ψ is a monotone function, by a similar argument, $\Psi^*(\alpha,\beta)$ can be shown to be a $\Psi(B_r,B_r)$-Lipschitz function. For all the performance measures considered, we have $\Psi(B_r,B_r) \leq B_r$. Thus, $\ell_t^d(\alpha,\beta)$ is a $2B_r$-Lipschitz function. Hence, using a standard GIGA-style analysis (Zinkevich, 2003) on the (descent) updates on α_t and β_t in Algorithm 1, we have (for $\eta_t = \frac{1}{\sqrt{t}}$)

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r^+(\mathbf{w}_t;\,\mathbf{x}_t,y_t) \,+\, \beta_t r^-(\mathbf{w}_t;\,\mathbf{x}_t,y_t) \,-\, \Psi^*(\alpha_t,\beta_t)\right] \\
&\quad\leq \inf_{(\alpha,\beta)\in\mathcal{A}}\left\{\frac{1}{T}\sum_{t=1}^{T}\left[\alpha r^+(\mathbf{w}_t;\,\mathbf{x}_t,y_t) \,+\, \beta r^-(\mathbf{w}_t;\,\mathbf{x}_t,y_t) \,-\, \Psi^*(\alpha,\beta)\right]\right\} \,+\, \left(L_\Psi^2 + 4B_r^2\right)\frac{1}{2\sqrt{T}} \\
&\quad= \inf_{(\alpha,\beta)\in\mathcal{A}}\left\{\alpha\frac{1}{T}\sum_{t=1}^{T} r^+(\mathbf{w}_t;\,\mathbf{x}_t,y_t) \,+\, \beta\frac{1}{T}\sum_{t=1}^{T} r^-(\mathbf{w}_t;\,\mathbf{x}_t,y_t) \,-\, \Psi^*(\alpha,\beta)\right\} \,+\, \left(L_\Psi^2 + 4B_r^2\right)\frac{1}{2\sqrt{T}} \\
&\quad= \Psi\left(\frac{1}{T}\sum_{t=1}^{T} r^+(\mathbf{w}_t;\,\mathbf{x}_t,y_t),\, \frac{1}{T}\sum_{t=1}^{T} r^-(\mathbf{w}_t;\,\mathbf{x}_t,y_t)\right) \,+\, \left(L_\Psi^2 + 4B_r^2\right)\frac{1}{2\sqrt{T}},
\end{aligned}$$

where the last step follows from Fenchel conjugacy.

Further, noting that $\mathbb{E}_{\mathbf{x}_t,y_t}\left[\!\left[r^+(\mathbf{w}_t;\,\mathbf{x}_t,y_t)\,\middle|\,\mathbf{x}_{1:t-1},y_{1:t-1}\right]\!\right] = P(\mathbf{w}_t)$, and $\mathbb{E}_{\mathbf{x}_t,y_t}\left[\!\left[r^-(\mathbf{w}_t;\,\mathbf{x}_t,y_t)\,\middle|\,\mathbf{x}_{1:t-1},y_{1:t-1}\right]\!\right] = N(\mathbf{w}_t)$, we use the standard online-batch conversion bounds (Cesa-Bianchi et al., 2001) to the loss functions r^+ and r^- individually to obtain w.h.p.

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} r^+(\mathbf{w}_t;\,\mathbf{x}_t,y_t) &\leq \sum_{t=1}^{T} P(\mathbf{w}_t) + \sqrt{\frac{2B_r^2}{T}\log\frac{1}{\delta}} \\
\frac{1}{T}\sum_{t=1}^{T} r^-(\mathbf{w}_t;\,\mathbf{x}_t,y_t) &\leq \sum_{t=1}^{T} N(\mathbf{w}_t) + \sqrt{\frac{2B_r^2}{T}\log\frac{1}{\delta}}
\end{aligned}$$

By monotonicity of Ψ, we get

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r^+(\mathbf{w}_t;\,\mathbf{x}_t,y_t) \,+\, \beta_t r^-(\mathbf{w}_t;\,\mathbf{x}_t,y_t) \,-\, \Psi^*(\alpha_t,\beta_t)\right] \\
&\quad\leq \Psi\left(\frac{1}{T}\sum_{t=1}^{T} P(\mathbf{w}_t) + \sqrt{\frac{2B_r^2}{T}\log\frac{1}{\delta}},\, \frac{1}{T}\sum_{t=1}^{T} N(\mathbf{w}_t) + \sqrt{\frac{2B_r^2}{T}\log\frac{1}{\delta}}\right) \,+\, \left(L_\Psi^2 + 4B_r^2\right)\frac{1}{2\sqrt{T}} \\
&\quad\leq \Psi\left(\frac{1}{T}\sum_{t=1}^{T} P(\mathbf{w}_t),\, \frac{1}{T}\sum_{t=1}^{T} N(\mathbf{w}_t)\right) \,+\, \delta_\Psi\left(\sqrt{\frac{2B_r^2}{T}\log\frac{1}{\delta}}\right) \,+\, \left(L_\Psi^2 + 4B_r^2\right)\frac{1}{2\sqrt{T}}
\end{aligned}$$

$$\begin{aligned}
&\leq \Psi\left(\bar{r}^+\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{w}_t\right),\, \bar{r}^-\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{w}_t\right)\right) + \delta_\Psi\left(\sqrt{\frac{2B_r^2}{T}\log\frac{1}{\delta}}\right) + \left(L_\Psi^2 + 4B_r^2\right)\frac{1}{2\sqrt{T}} \\
&= \Psi\big(P(\overline{\mathbf{w}}),\, N(\overline{\mathbf{w}})\big) + \delta_\Psi\left(\sqrt{\frac{2B_r^2}{T}\log\frac{1}{\delta}}\right) + \left(L_\Psi^2 + 4B_r^2\right)\frac{1}{2\sqrt{T}}, \qquad (1)
\end{aligned}$$

where the second inequality follows from stability of Ψ, and the third inequality follows from concavity of $\bar{r}^+$ and $\bar{r}^-$, Jensen's inequality, and stability of Ψ.

Similarly, the update to $\mathbf{w}$ can be written as

$$\mathbf{w}_{t+1} \leftarrow \Pi_{\mathcal{W}}\left(\mathbf{w}_t - \eta_t'\nabla_{\mathbf{w}}\ell_t^p(\mathbf{w}_t)\right),$$

where $\Pi_{\mathcal{W}}$ is the projection operator for the domain $\mathcal{W}$ and

$$\ell_t^p(\mathbf{w}) = \begin{cases} -\alpha_t r^+(\mathbf{w};\mathbf{x}_t,y_t) + \Psi^*(\alpha_t,\beta_t) & \text{if } y_t > 0 \\ -\beta_t r^-(\mathbf{w};\mathbf{x}_t,y_t) + \Psi^*(\alpha_t,\beta_t) & \text{if } y_t < 0 \end{cases}$$

Since r^+, r^- are concave and the term $\Psi^*(\alpha_t,\beta_t)$ does not involve $\mathbf{w}$, ℓ_t^p is convex in $\mathbf{w}$ for all t. Also, we can show that $\ell_t^p(\mathbf{w})$ is an $(L_\Psi \cdot L_r)$-Lipschitz function. Hence, applying a standard GIGA analysis (Zinkevich, 2003) to the (ascent) update on $\mathbf{w}_t$ in Algorithm 1 (with $\eta_t' = \frac{1}{\sqrt{t}}$), we have for any $\mathbf{w}^* \in \mathcal{W}$,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r^+(\mathbf{w}_t;\mathbf{x}_t,y_t) + \beta_t r^-(\mathbf{w}_t;\mathbf{x}_t,y_t) - \Psi^*(\alpha_t,\beta_t)\right] \\
&\quad\geq \frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r^+(\mathbf{w}^*;\mathbf{x}_t,y_t) + \beta_t r^-(\mathbf{w}^*;\mathbf{x}_t,y_t) - \Psi^*(\alpha_t,\beta_t)\right] - \left(L_\Psi^2 L_r^2 + R_{\mathcal{W}}^2\right)\frac{1}{2\sqrt{T}}.
\end{aligned}$$

Again, observing that by linearity of expectation, we have

$$\mathbb{E}_{\mathbf{x}_t,y_t}\left[\!\left[\alpha_t r^+(\mathbf{w}^*;\mathbf{x}_t,y_t) + \beta_t r^-(\mathbf{w}^*;\mathbf{x}_t,y_t)\,\middle|\,\mathbf{x}_{1:t-1},y_{1:t-1}\right]\!\right] = \alpha_t P(\mathbf{w}^*) + \beta_t N(\mathbf{w}^*),$$

which gives us, through an online-batch conversion argument (Cesa-Bianchi et al., 2001) w.h.p,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r^+(\mathbf{w}_t;\mathbf{x}_t,y_t) + \beta_t r_t^-(\mathbf{w}_t;\mathbf{x}_t,y_t) - \Psi^*(\alpha_t,\beta_t)\right] \\
&\geq \frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t P(\mathbf{w}^*) + \beta_t N(\mathbf{w}^*)\right] - \frac{1}{T}\sum_{t=1}^{T}\Psi^*(\alpha_t,\beta_t) - \sqrt{\frac{2L_\Psi^2 B_r^2}{T}\log\frac{1}{\delta}} - \left(L_\Psi^2 L_r^2 + R_{\mathcal{W}}^2\right)\frac{1}{2\sqrt{T}} \\
&\geq \frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t P(\mathbf{w}^*) + \beta_t N(\mathbf{w}^*)\right] - \Psi^*\left(\frac{1}{T}\sum_{t=1}^{T}\alpha_t, \frac{1}{T}\sum_{t=1}^{T}\beta_t\right) - \sqrt{\frac{2L_\Psi^2 B_r^2}{T}\log\frac{1}{\delta}} - \left(L_\Psi^2 L_r^2 + R_{\mathcal{W}}^2\right)\frac{1}{2\sqrt{T}} \\
&= \bar{\alpha}P(\mathbf{w}^*) + \bar{\beta}N(\mathbf{w}^*) - \Psi^*(\bar{\alpha},\bar{\beta}) - \sqrt{\frac{2L_\Psi^2 B_r^2}{T}\log\frac{1}{\delta}} - \left(L_\Psi^2 L_r^2 + R_{\mathcal{W}}^2\right)\frac{1}{2\sqrt{T}} \\
&\geq \inf_{\alpha,\beta}\left\{\alpha P(\mathbf{w}^*) + \beta N(\mathbf{w}^*) - \Psi^*(\alpha,\beta)\right\} - \sqrt{\frac{2L_\Psi^2 B_r^2}{T}\log\frac{1}{\delta}} - \left(L_\Psi^2 L_r^2 + R_{\mathcal{W}}^2\right)\frac{1}{2\sqrt{T}} \\
&= \Psi\big(P(\mathbf{w}^*),\, N(\mathbf{w}^*)\big) - \sqrt{\frac{2L_\Psi^2 B_r^2}{T}\log\frac{1}{\delta}} - \left(L_\Psi^2 L_r^2 + R_{\mathcal{W}}^2\right)\frac{1}{2\sqrt{T}}, \qquad (2)
\end{aligned}$$

where the second step follows from concavity of Ψ and Jensen's inequality, in the third step $\bar{\alpha} = \frac{1}{T}\sum_{t=1}^{T}\alpha_t$ and $\bar{\beta} = \frac{1}{T}\sum_{t=1}^{T}\beta_t$, and the last step follows from Fenchel conjugacy.

Combining Eq. (1) and (2) gives us the desired result. □

C. Proof of Theorem 5

Theorem 5. *Suppose we have the problem setting in Theorem 4 with the $\Psi_{\text{G-mean}}$ performance measure being optimized for. Consider a modification to Algorithm 1 wherein the reward functions are changed to $r_t^+(\cdot) = r^+(\cdot) + \epsilon(t)$, and $r_t^-(\cdot) = r^-(\cdot) + \epsilon(t)$ for $\epsilon(t) = \frac{1}{t^{1/4}}$. Then, the average model $\overline{\mathbf{w}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{w}_t$ output by the algorithm satisfies, with probability at least $1-\delta$,*

$$\mathcal{P}_{\Psi_{\text{G-mean}}}(\overline{\mathbf{w}}) \geq \sup_{\mathbf{w}^* \in \mathcal{W}} \mathcal{P}_{\Psi_{\text{G-mean}}}(\mathbf{w}^*) - \widetilde{\mathcal{O}}\left(\frac{1}{T^{1/4}}\right).$$

Proof. Suppose $\Psi(u+\epsilon, v+\epsilon) \leq \Psi(u,v) + \delta_\Psi(\epsilon)$ as before. Let $r_t^+(\cdot) = r^+(\cdot) + \epsilon(t)$, and $r_t^-(\cdot) = r^-(\cdot) + \epsilon(t)$. Let us make all updates with respect to r_t^+, r_t^-. Let $r(\epsilon)$ be the radius of the sufficient dual domain $\mathcal{A}$ for a given regularization ϵ. Also let $\bar{\epsilon} = \frac{1}{T}\sum_{i=1}^{T} \epsilon(t)$. We will assume throughout that $\epsilon(t) = O(1)$. Then we have:

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r_t^+(\mathbf{w}_t;\, \mathbf{x}_t, y_t) \,+\, \beta_t r_t^-(\mathbf{w}_t;\, \mathbf{x}_t, y_t) \,-\, \Psi^*(\alpha_t, \beta_t)\right] \\
&\quad\leq \inf_{(\alpha,\beta)\in\mathcal{A}} \left\{\frac{1}{T}\sum_{t=1}^{T}\left[\alpha r_t^+(\mathbf{w}_t;\, \mathbf{x}_t, y_t) \,+\, \beta r_t^-(\mathbf{w}_t;\, \mathbf{x}_t, y_t) \,-\, \Psi^*(\alpha, \beta)\right]\right\} \,+\, \mathcal{O}\left(\frac{r(\bar{\epsilon})}{\sqrt{T}}\right) \\
&\quad= \inf_{(\alpha,\beta)\in\mathcal{A}} \left\{\alpha\frac{1}{T}\sum_{t=1}^{T} r^+(\mathbf{w}_t;\, \mathbf{x}_t, y_t) + \bar{\epsilon} \,+\, \beta\frac{1}{T}\sum_{t=1}^{T} r^-(\mathbf{w}_t;\, \mathbf{x}_t, y_t) + \bar{\epsilon} \,-\, \Psi^*(\alpha, \beta)\right\} \,+\, \mathcal{O}\left(\frac{r(\bar{\epsilon})}{\sqrt{T}}\right) \\
&\quad= \Psi\left(\frac{1}{T}\sum_{t=1}^{T} r^+(\mathbf{w}_t;\, \mathbf{x}_t, y_t) + \bar{\epsilon},\ \frac{1}{T}\sum_{t=1}^{T} r^-(\mathbf{w}_t;\, \mathbf{x}_t, y_t) + \bar{\epsilon}\right) \,+\, \mathcal{O}\left(\frac{r(\bar{\epsilon})}{\sqrt{T}}\right) \qquad (3) \\
&\quad= \Psi\left(\frac{1}{T}\sum_{t=1}^{T} r^+(\mathbf{w}_t;\, \mathbf{x}_t, y_t),\ \frac{1}{T}\sum_{t=1}^{T} r^-(\mathbf{w}_t;\, \mathbf{x}_t, y_t)\right) \,+\, \delta_\Psi(\bar{\epsilon}) \,+\, \mathcal{O}\left(\frac{r(\bar{\epsilon})}{\sqrt{T}}\right)
\end{aligned}$$

We can now use online to batch conversion bounds (Cesa-Bianchi et al., 2001), and monotonicity of Ψ to get

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r^+(\mathbf{w}_t;\, \mathbf{x}_t, y_t) \,+\, \beta_t r^-(\mathbf{w}_t;\, \mathbf{x}_t, y_t) \,-\, \Psi^*(\alpha_t, \beta_t)\right] \\
&\quad\leq \Psi\big(P(\overline{\mathbf{w}}),\, N(\overline{\mathbf{w}})\big) \,+\, \delta_\Psi\left(\widetilde{\mathcal{O}}\left(\frac{1}{\sqrt{T}}\right)\right) \,+\, \delta_\Psi(\bar{\epsilon}) \,+\, \mathcal{O}\left(\frac{r(\bar{\epsilon})}{\sqrt{T}}\right), \qquad (4)
\end{aligned}$$

For the primal updates, we get, for any $\mathbf{w}^* \in \mathcal{W}$,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r_t^+(\mathbf{w}_t; \mathbf{x}_t, y_t) \,+\, \beta_t r_t^-(\mathbf{w}_t; \mathbf{x}_t, y_t) \,-\, \Psi^*(\alpha_t, \beta_t)\right] \\
&\quad\geq \frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r_t^+(\mathbf{w}^*; \mathbf{x}_t, y_t) \,+\, \beta_t r_t^-(\mathbf{w}^*; \mathbf{x}_t, y_t) \,-\, \Psi^*(\alpha_t, \beta_t)\right] \,-\, \widetilde{\mathcal{O}}\left(\frac{r(\bar{\epsilon})}{\sqrt{T}}\right) \\
&\quad= \frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r^+(\mathbf{w}^*; \mathbf{x}_t, y_t) \,+\, \beta_t r^-(\mathbf{w}^*; \mathbf{x}_t, y_t) \,-\, \Psi^*(\alpha_t, \beta_t)\right] + \frac{1}{T}\sum_{t=1}^{T}\epsilon(t)(\alpha_t + \beta_t) \,-\, \widetilde{\mathcal{O}}\left(\frac{r(\bar{\epsilon})}{\sqrt{T}}\right) \\
&\quad\geq \frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r^+(\mathbf{w}^*; \mathbf{x}_t, y_t) \,+\, \beta_t r^-(\mathbf{w}^*; \mathbf{x}_t, y_t) \,-\, \Psi^*(\alpha_t, \beta_t)\right] \,-\, \widetilde{\mathcal{O}}\left(\frac{r(\bar{\epsilon})}{\sqrt{T}}\right),
\end{aligned}$$

since $\epsilon(t), \alpha_t, \beta_t \geq 0$. Again using an online-batch conversion argument (Cesa-Bianchi et al., 2001) we get w.h.p,

$$\frac{1}{T}\sum_{t=1}^{T}\left[\alpha_t r^+(\mathbf{w}_t; \mathbf{x}_t, y_t) \,+\, \beta_t r_t^-(\mathbf{w}_t; \mathbf{x}_t, y_t) \,-\, \Psi^*(\alpha_t, \beta_t)\right] \geq \Psi\big(P(\mathbf{w}^*),\, N(\mathbf{w}^*)\big) \,-\, \widetilde{\mathcal{O}}\left(\frac{r(\bar{\epsilon})}{\sqrt{T}}\right). \qquad (5)$$

Combining Eq. (4) and (5) gives us

$$\Psi\big(P(\overline{\mathbf{w}}),\ N(\overline{\mathbf{w}})\big) \geq \Psi\big(P(\mathbf{w}^*),\ N(\mathbf{w}^*)\big)\ -\ \widetilde{\mathcal{O}}\left(\frac{r(\bar{\epsilon})}{\sqrt{T}}\right) - \delta_\Psi\left(\widetilde{\mathcal{O}}\left(\frac{1}{\sqrt{T}}\right)\right)\ -\ \delta_\Psi(\bar{\epsilon})$$

For G-mean, $\delta_\Psi(x) = \sqrt{x}$, and by an application of Lemma 3,we have $r(\epsilon) = O(1/\sqrt{\epsilon})$. Thus we have

$$\Psi\big(P(\overline{\mathbf{w}}),\ N(\overline{\mathbf{w}})\big) \geq \Psi\big(P(\mathbf{w}^*),\ N(\mathbf{w}^*)\big)\ -\ \widetilde{\mathcal{O}}\left(\frac{1}{\sqrt{T\bar{\epsilon}}}\right) - \widetilde{\mathcal{O}}\left(\frac{1}{\sqrt[4]{T}}\right)\ -\ \sqrt{\bar{\epsilon}}$$

For $\bar{\epsilon} = \mathcal{O}\left(\frac{1}{\sqrt[4]{T}}\right)$, we get

$$\Psi\big(P(\overline{\mathbf{w}}),\ N(\overline{\mathbf{w}})\big) \geq \Psi\big(P(\mathbf{w}^*),\ N(\mathbf{w}^*)\big)\ -\ \widetilde{\mathcal{O}}\left(\frac{1}{\sqrt[4]{T}}\right)$$

This can be achieved with $\epsilon(t) = \frac{1}{\sqrt[4]{t}}$. □

D. Proof of Theorem 8

Theorem 8. *Let Algorithm 2 be executed with a performance measure $\mathcal{P}_{(\mathbf{a},\mathbf{b})}$ and reward functions that offer values in the range $[0, m)$. Let $\mathcal{P}^* := \sup_{\mathbf{w}\in\mathcal{W}} \mathcal{P}_{(\mathbf{a},\mathbf{b})}(\mathbf{w})$. Also let $\Delta_t = \mathcal{P}^* - \mathcal{P}_{(\mathbf{a},\mathbf{b})}(\mathbf{w}_t)$ be the excess error for the model $\mathbf{w}_t$ generated at time t. Then there exists a value $\eta(m) < 1$ such that for $\Delta_t \leq \Delta_0 \cdot \eta(m)^t$.*

Proof. In order to be generic in its treatment, the proof will require the following regularity conditions on the performance measure

1. $b_0 \neq 0$
2. $\alpha - \mathcal{P}(\mathbf{w}) \cdot \gamma \geq 0$ for all $\mathbf{w} \in \mathcal{W}$
3. $\beta - \mathcal{P}(\mathbf{w}) \cdot \delta \geq 0$ for all $\mathbf{w} \in \mathcal{W}$
4. $-1 < f \leq \gamma \cdot P(\mathbf{w}) + \delta \cdot N(\mathbf{w}) \leq g$ for all $\mathbf{w} \in \mathcal{W}$

Define $e_t := V(\mathbf{w}_{t+1}, v_t) - v_t$. Then we can state the following lemmata which together yield the convergence bound proof.

Lemma 10. $\frac{e_t}{1+f} \geq \mathcal{P}^* - v_t$

Proof. Assume that for some $\mathbf{w}^*$, $\mathcal{P}(\mathbf{w}^*) = v_t + e_t + e'$ where $e' > 0$. Then we have

$$\begin{aligned} V(\mathbf{w}^*, v_t) &= \left(\frac{e_t}{1+f} + e'\right)(1 + \gamma \cdot P(\mathbf{w}^*) + \delta \cdot N(\mathbf{w}^*)) - e_t \\ &\geq \left(\frac{e_t}{1+f} + e'\right)(1 + f) - e_t \\ &= e'(1+f) > 0, \end{aligned}$$

which contradicts the fact that no classifier can achieve a valuation greater than $v_t + e_t$ at level v_t, thus proving the desired result. □

Lemma 11. *For any $\mathbf{w}$ that achieves $V(\mathbf{w}, v) = v + e$ such that $e \geq 0$, we have*

$$\mathcal{P}(\mathbf{w}) \geq v + \frac{e}{g+1}$$

Proof. Let $v' = v + \frac{e}{g+1}$. We will show that $V(\mathbf{w}, v') \geq v'$ which will establish the result by pseudo-linearity. We have

$$\begin{aligned} V(\mathbf{w}, v') - v' &= c + (\alpha - v'\gamma)\cdot P(\mathbf{w}) + (\beta - v'\delta)\cdot N(\mathbf{w}) - v' \\ &= c + (\alpha - v\gamma)\cdot P(\mathbf{w}) + (\beta - v\delta)\cdot N(\mathbf{w}) - v' - \frac{e}{g+1}(\gamma\cdot P(\mathbf{w}) + \delta\cdot N(\mathbf{w})) \\ &= v + e - v' - \frac{e}{g+1}(\gamma\cdot P(\mathbf{w}) + \delta\cdot N(\mathbf{w})) \\ &\geq v + e - v' - \frac{ge}{g+1} = 0, \end{aligned}$$

where we have used the bounds on $\gamma\cdot P(\mathbf{w}) + \delta\cdot N(\mathbf{w})$ and the fact that $1 + g > 0$. □

Given the above results we can establish the convergence bound. More specifically, we can show the following: let $\Delta_t = \mathcal{P}^* - \mathcal{P}(\mathbf{w}_t)$. Then we have

$$\Delta_{t+1} \leq \frac{g-f}{g+1}\cdot\Delta_t$$

To see this, consider the following

$$\begin{aligned} \Delta_{t+1} &= \mathcal{P}^* - \mathcal{P}(\mathbf{w}_{t+1}) \leq \mathcal{P}^* - \left(v_t + \frac{e_t}{g+1}\right) \leq \mathcal{P}^* - \left(v_t + \frac{(1+f)(\mathcal{P}^* - v_t)}{g+1}\right) \\ &= \mathcal{P}^* - \left(\mathcal{P}(\mathbf{w}_t) + \frac{(1+f)(\mathcal{P}^* - \mathcal{P}(\mathbf{w}_t))}{g+1}\right) = \Delta_t - \frac{1+f}{g+1}\cdot\Delta_t = \frac{g-f}{g+1}\cdot\Delta_t, \end{aligned}$$

which proves the result. Notice that Table 2 gives the rates of convergence for the different performance measures by calculating bounds on the value of $\frac{g-f}{g+1}$ for those performance measures. □

E. An analysis of the AMP Algorithm under Inexact Maximizations

For this and the next section, we will, for the sake of simplicity, we will focus only on the F-measure for $\beta = 1$ and $p = 1/2$ so that $\theta = 1$. For this setting, the F-measure looks like the following: $F(P, N) = \frac{2P}{2+P-N}$, and the valuation function looks like $V(\mathbf{w}, v) = (1 - v/2)\cdot P(\mathbf{w}) + v/2\cdot N(\mathbf{w})$. We shall denote the performance measure as $F(\mathbf{w})$, and its optimal value as F^*. We will assume that the reward functions give bounded rewards in the range $[0, m)$.

So far we assumed that Step 4 in the Algorithm **AMP** gave us $\mathbf{w}_{t+1}$ such that

$$V(\mathbf{w}_{t+1}, v_t) = \max_{\mathbf{w}\in\mathcal{W}} V(\mathbf{w}, v_t)$$

Now we will only assume that $\mathbf{w}_{t+1}$ satisfies

$$V(\mathbf{w}_{t+1}, v_t) = \max_{\mathbf{w}\in\mathcal{W}} V(\mathbf{w}, v_t) - \epsilon_t$$

We also assume that the level v_t is only approximated in Step 5 of **AMP**, i.e. using Lemma 7 we have

$$v_t = F(\mathbf{w}_t) + \delta_t$$

where δ_t is a signed real number.

Given these approximations, we can prove the following results

Lemma 12. *The following hold for the setting described above*

1. *If* $\delta_t \leq 0$ *then* $e_t \geq 0$
2. *If* $\delta_t > 0$ *then* $e_t \geq -\delta_t\left(1 + \frac{m}{2}\right)$
3. *If* $F^* < v_t$ *(which can happen only if* $\delta_t > 0$*), then* $e_t < 0$
4. *If* $e_t < 0$ *then* $F^* < v_t$

5. *We have*
   (a) *If* $e_t \geq 0$, *then* $e_t \geq \left(\frac{2-m}{2}\right)(F^* - v_t)$.
   (b) *If* $e_t < 0$, *then* $e_t \geq \left(\frac{2+m}{2}\right)(F^* - v_t)$.
6. *If* $V(\mathbf{w}, v) = v + e$, *then*
   (a) *If* $e \geq 0$ *then* $F(\mathbf{w}) \geq v + \frac{2e}{2+m}$
   (b) *If* $e < 0$ *then* $F(\mathbf{w}) \geq v + \frac{2e}{2-m}$

Proof. We give the proof in parts

1. If $\delta_t \leq 0$ then this means that there exists a $\mathbf{w}$ such that $F(\mathbf{w}) \geq v_t$. The result then follows from pseudo linearity.
2. $v_t = F(\mathbf{w}_t) + \delta_t$ gives us, by pseudo linearity of F-measure,

$$(1 - v_t/2) \cdot P(\mathbf{w}_t) + v_t/2 \cdot N(\mathbf{w}_t) = v_t - \delta_t \left(1 + \frac{P(\mathbf{w}_t) - N(\mathbf{w}_t)}{2}\right) \geq v_t - \delta_t \left(1 + \frac{m}{2}\right).$$

   The bound on e_t now follows from its definition.
3. Suppose $e_t \geq 0$ then by pseudo linearity of F-measure, we have, for some $\mathbf{w}$, $V(\mathbf{w}, v_t) \geq v_t$ which means $F(\mathbf{w}) \geq v_t$ which contradicts the assumption.
4. Suppose there exists $\mathbf{w}^*$ with $F(\mathbf{w}^*) = v_t + e'$ with $e' \geq 0$ then we have

$$(1 - v_t/2) \cdot P(\mathbf{w}^*) + v_t/2 \cdot N(\mathbf{w}^*) = v_t + e' \left(1 + \frac{P(\mathbf{w}^*) - N(\mathbf{w}^*)}{2}\right) \geq 0,$$

   which contradicts the fact that $e_t < 0$.
5. Part (a) is simply Lemma 10. For part (b), we will prove that $F^* \leq v_t + \frac{2e_t}{2+m}$. Since $\frac{2}{2+m} > 0$, the result will follow. Assume the contrapositive that some $\mathbf{w}^*$ achieves $F(\mathbf{w}^*) = v_t + \frac{2e_t}{2+m} + e'$ for some $e' > 0$. Using the pseudo linearity of F-measure (and using the shorthand $v' = v_t + \frac{2e_t}{2+m} + e'$), this can be expressed as

$$(1 - v'/2) \cdot P(\mathbf{w}^*) + v'/2 \cdot N(\mathbf{w}^*) = v'$$

   where for some $e' > 0$. Then we have

$$\begin{aligned}(1 - v_t/2) \cdot P(\mathbf{w}^*) + v_t/2 \cdot N(\mathbf{w}^*) - v_t - e_t &= v' - v_t - e_t + \frac{1}{2}\left(\frac{2e_t}{2+m} + e'\right)(P(\mathbf{w}^*) - N(\mathbf{w}^*)) \\ &= \frac{2e_t}{2+m} + e' - e_t + \frac{1}{2}\left(\frac{2e_t}{2+m} + e'\right)(P(\mathbf{w}^*) - N(\mathbf{w}^*)) \\ &\geq \frac{2e_t}{2+m} + e' - e_t + \frac{m}{2}\left(\frac{2e_t}{2+m} + e'\right) \\ &= e'\left(1 + \frac{m}{2}\right) + e_t\left(\frac{2}{2+m} - 1 + \frac{m}{2+m}\right) \\ &= e'\left(1 + \frac{m}{2}\right) > 0,\end{aligned}$$

   where we have assumed that e' is chosen small enough so that $\frac{2e_t}{2+m} + e' < 0$ still and used the fact that $P(\mathbf{w}^*) - N(\mathbf{w}^*) \leq m$.
6. Part (a) is simply Lemma 11. To prove part (b), we let $v' = v + \frac{2e}{2-m}$, then we have

$$(1 - \frac{v'}{2}) \cdot P(\mathbf{w}) + \frac{v'}{2} \cdot N(\mathbf{w}) - v' \quad = \quad (1 - \frac{v}{2}) \cdot P(\mathbf{w}) + \frac{v}{2} \cdot N(\mathbf{w}) - v' + \frac{e}{2-m}(N(\mathbf{w}) - P(\mathbf{w}))$$

$$
\begin{aligned}
&\geq (1-\frac{v}{2})\cdot P(\mathbf{w})+\frac{v}{2}\cdot N(\mathbf{w})-v'+\frac{me}{2-m}\\
&= v+e-\left(v+\frac{2e}{2-m}\right)+\frac{me}{2-m}\\
&= e\left(1-\frac{2}{2-m}+\frac{m}{2-m}\right)\\
&= 0,
\end{aligned}
$$

where the second inequality follows since $N(\mathbf{w})-P(\mathbf{w})\leq m$ and $e<0$ by using the bounds on the reward functions. This proves the result.

□

E.1. Convergence analysis

We have the following cases with us

1. Case 1 ($\delta_t\leq 0$): In this case we are setting v_t to a value less than the F-measure of the current classifier. This should hurt performance - we know that $v_t=F(\mathbf{w}_t)+\delta_t$ which gives us, on applying part (a) of the previous lemma using $F^*-v_t=\Delta_t-\delta_t$, the following
$$e_t\geq\frac{2-m}{2}(\Delta_t-\delta_t).$$
Note that we are guaranteed that $e_t\geq 0$ in this case. Now since the maximization in step 4 is also carried our approximately, we have $V(\mathbf{w}_{t+1},v_t)=v_t+e_t-\epsilon_t$. Now we have two sub cases
   (a) Case 1.1 ($\epsilon_t\leq e_t$): In this case we can apply part 6(a) of the previous lemma to get the following result
   $$\Delta_{t+1}\leq\frac{2m}{2+m}\Delta_t-\frac{2m}{2+m}\delta_t+\frac{2\epsilon_t}{2m}$$
   (b) Case 1.2 ($\epsilon_t>e_t$): In this case we are actually making negative progress in the maximization step (since we have $V(\mathbf{w}_{t+1},v_t)\leq v_t$) and we can only invoke Lemma 5.6(b) to get
   $$\Delta_{t+1}\leq\frac{2\epsilon_t}{2-m}$$
   Note that the above result should not be interpreted as a one shot step to a very good classifier. The above result holds along with the condition that $\epsilon_t>e_t$. Thus the performance of the classifier is lower bounded by e_t which depends on how far the current classifier is from the best.

2. Case 2 ($\delta_t>0$): In this case we are setting v_t to the value higher than the F-measure of the current classifier. This can mislead the classifier and results in the following two sub-cases
   (a) Case 2.1 ($F^*\geq v_t$): In this case we are still setting v_t to a legitimate value, i.e. one that is a valid F-measure for some classifier in the hypothesis class. This can only benefit the next optimization stage (in fact if we set $v_t=F^*$, then we would obtain the best classifier in this very iteration!). In this case $e_t\geq 0$ and we can use the analyses of Cases 1.1 and 1.2.
   (b) Case 2.2 ($F^*<v_t$): In this case we are setting v_t to an illegal value, one that is an unachievable value of F-measure. Consequently, using part 3 of the previous lemma, $e_t<0$ and using part(b) of the previous lemma we get
   $$e_t\geq\frac{2+m}{2}(\Delta_t-\delta_t),$$
   which, upon applying part 6(b) of the previous lemma (since $e_t-\epsilon_t\leq e_t<0$) will give us
   $$
   \begin{aligned}
   \Delta_{t+1}&\leq\frac{2m}{2-m}(\delta_t-\Delta_t)+\frac{2\epsilon_t}{2-m}\\
   &\leq\frac{2m}{2-m}\delta_t+\frac{2\epsilon_t}{2-m}
   \end{aligned}
   $$

We can combine the cases together as follows

$$
\begin{aligned}
\Delta_{t+1} &\leq \max\left\{\mathbf{1}\left\{\delta \leq 0\right\} \cdot \left\{\frac{2m}{2+m}\Delta_t - \frac{2m}{2+m}\delta_t + \frac{2\epsilon_t}{2+m}\right\}, \mathbf{1}\left\{\epsilon_t > e_t\right\} \cdot \frac{2\epsilon_t}{2-m}, \mathbf{1}\left\{\delta > 0\right\} \cdot \left\{\frac{2m}{2-m}\delta_t + \frac{2\epsilon_t}{2-m}\right\}\right\} \\
&\leq \max\left\{\frac{2m}{2+m}\Delta_t + \frac{2m}{2+m}\left|\delta_t\right| + \frac{2\epsilon_t}{2+m}, \mathbf{1}\left\{\epsilon_t > e_t\right\} \cdot \frac{2\epsilon_t}{2-m}, \frac{2m}{2-m}\left|\delta_t\right| + \frac{2\epsilon_t}{2-m}\right\} \\
&\leq \frac{2m}{2+m}\Delta_t + \frac{2m}{2-m}\left|\delta_t\right| + \frac{2\epsilon_t}{2-m}
\end{aligned}
$$

If we let $\eta = \frac{2m}{2+m}$, $\eta' = \frac{2m}{2-m}$, and $\xi_t = |\delta_t| + \epsilon_t/m$, then this gives us

$$\Delta_{t+1} \leq \eta\Delta_t + \eta'\xi_t,$$

which gives us

$$\Delta_T \leq \eta^T\Delta_0 + \frac{\eta'}{\eta} \cdot \sum_{i=0}^{T-1} \eta^{T-i}\xi_i$$

This concludes our analysis.

F. Proof of Theorem 9

Theorem 9. *Let Algorithm 3 be executed with a performance measure $\mathcal{P}_{(\mathbf{a},\mathbf{b})}$ and reward functions with range $[0, m)$. Let $\eta = \eta(m)$ be the rate of convergence guaranteed for $\mathcal{P}_{(\mathbf{a},\mathbf{b})}$ by the **AMP** algorithm. Set the epoch lengths to $s_e, s'_e = \widetilde{\mathcal{O}}\left(\frac{1}{\eta^{2e}}\right)$. Then after $e = \log_{\frac{1}{\eta}}\left(\frac{1}{\epsilon}\log^2\frac{1}{\epsilon}\right)$ epochs, we can ensure with probability at least $1-\delta$ that $\mathcal{P}^* - \mathcal{P}_{(\mathbf{a},\mathbf{b})}(\mathbf{w}_e) \leq \epsilon$. Moreover the number of samples consumed till this point is at most $\widetilde{\mathcal{O}}\left(\frac{1}{\epsilon^2}\right)$.*

Proof. Using Hoeffding's inequality, standard regret and online-to-batch guarantees (Cesa-Bianchi et al., 2001; Zinkevich, 2003), we can ensure that, if the stream lengths for the Model optimization stage and Challenge level estimation stage procedures are s_e and s'_e respectively, then for some fixed $c > 0$ that is independent of the stream length, we have

$$|\delta_t| \leq c \cdot \sqrt{\frac{\log\frac{1}{\delta}}{s'_e}}, |\epsilon_t| \leq c\sqrt{\frac{\log\frac{1}{\delta}}{s_e}}$$

Let $T = \log_{\frac{1}{\eta}}\left(\frac{1}{\epsilon}\log^2\frac{1}{\epsilon}\right)$ and $s_e = \left(\frac{2c}{m}\right)^2\left(\frac{1}{\eta}\right)^{2e}\log\frac{T}{\delta}$ and $s'_e = 4c^2\left(\frac{1}{\eta}\right)^{2e}\log\frac{T}{\delta}$ - this gives us, for each e, with probability at least $1-\delta/T$,

$$\xi_e \leq \eta^e$$

Thus, using a union bound, with probability at least $1-\delta$, we have, by the discussion in the previous section,

$$
\begin{aligned}
\Delta_T &\leq \eta^T\Delta_0 + \frac{\eta'}{\eta}\sum_{i=0}^{T-1}\eta^{T-i}\xi_i \leq \eta^T\Delta_0 + \frac{\eta'}{\eta}T\eta^T \\
&\leq \epsilon\Delta_0\log^{-2}\frac{1}{\epsilon} + \frac{\eta'}{\eta}\log_{\frac{1}{\eta}}\left(\frac{1}{\epsilon}\log^2\frac{1}{\epsilon}\right)\epsilon\log^{-2}\frac{1}{\epsilon} \\
&\leq \epsilon\left(\Delta_0 + \frac{\eta'}{\eta\log\frac{1}{\eta}}\right),
\end{aligned}
$$

where the last step follows from the fact that for any $\epsilon < 1/e^2$, we have

$$\log\left(\frac{1}{\epsilon}\log^2\frac{1}{\epsilon}\right) \leq \log^2\frac{1}{\epsilon}$$

Let $d = \left(\Delta_0 + \frac{\eta'}{\eta \log \frac{1}{\eta}}\right)$ so that we can later set $\epsilon' = \epsilon/d$, and $s = 4c^2\left(1 + \frac{1}{m^2}\right)$ so that $s_e + s'_e = s\left(\frac{1}{\eta}\right)^{2e} \log \frac{T}{\delta}$. The total number of samples required can then be calculated as

$$\sum_{e=1}^{T} s_e + s'_e = s \log \frac{T}{\delta} \sum_{e=1}^{T} \left(\frac{1}{\eta}\right)^{2e} = s \log \frac{T}{\delta} \frac{1}{1-\eta^2} \left(\frac{1}{\eta^2}\right)^{T} \leq s \log \frac{T}{\delta} \frac{1}{1-\eta^2} \frac{1}{\epsilon^2} \log^4 \frac{1}{\epsilon}$$

This gives the number of samples required as

$$\mathcal{O}\left(\frac{1}{\epsilon^2} \log^4 \frac{1}{\epsilon} \left(\log\log \frac{1}{\epsilon} + \log \frac{1}{\delta}\right)\right),$$

to get an ϵ-accurate solution with confidence $1 - \delta$. □