Human Boosting Appendix

1 Proof of Theorem 1

By the β-strong-smoothness of $\mathcal{R}_{emp}[f]$,

$$\mathcal{R}_{emp}[f_{t+1}] \leq \mathcal{R}_{emp}[f_t] + \langle \nabla \mathcal{R}_{emp}[f_t], f_{t+1} - f_t \rangle + \beta/2 \|f_{t+1} - f_t\|_2^2.$$

Denoting $\hat{\nabla}_t := \nabla \mathcal{R}_{emp}[f_t]$, and $\nabla_t := \mathbb{E}[\hat{\nabla}_t]$, and noting that $f_{t+1} = f_t - \frac{1}{\beta} \frac{\langle \hat{\nabla}_t, h_t \rangle}{\|h_t\|^2} h_t$, where $h_t \in \arg\max_{h \in H} \langle \hat{\nabla}_t, h \rangle / \|h\|$, we obtain,

$$\mathcal{R}_{emp}[f_{t+1}] \leq \mathcal{R}_{emp}[f_t] - \frac{1}{\beta \|h_t\|^2} \langle \hat{\nabla}_t, h_t \rangle \langle \nabla_t, h_t \rangle + \frac{1}{2\beta \|h_t\|^2} [\langle \hat{\nabla}_t, h_t \rangle]^2.$$

Denote $e_t := \hat{\nabla}_t - \nabla_t$, we obtain,

$$\begin{aligned}
\mathcal{R}_{emp}[f_{t+1}] - \mathcal{R}_{emp}[f_t] &\leq -\frac{1}{2\beta \|h_t\|^2} [\langle \hat{\nabla}_t, h_t \rangle]^2 + \frac{1}{\beta \|h_t\|^2} \langle e_t, h_t \rangle \langle \hat{\nabla}_t, h_t \rangle \\
&\leq -\frac{\gamma}{2\beta} \|\hat{\nabla}_t\|^2 + \|e_t\| \|\hat{\nabla}_t\|,
\end{aligned}$$

where we use the edge condition, $\langle \hat{\nabla}_t, h_t \rangle \geq \gamma \, \|\hat{\nabla}_t\| \|h_t\|$. Using the inequality that $\|\hat{\nabla}_t\|^2 \geq \frac{1}{2} \|\nabla_t\|^2 - e_t^2$, we get

$$\beta(\mathcal{R}_{emp}[f_{t+1}] - \mathcal{R}_{emp}[f_t]) \leq -\frac{\gamma}{4} \|\nabla_t\|^2 + (1+\gamma) \|e_t\|^2 + \|e_t\| \|\nabla_t\|,$$

so that taking expectations on both sides, we get

$$\beta \, \mathbb{E}[\mathcal{R}_{emp}[f_{t+1}] - \mathcal{R}_{emp}[f_t]] \leq -(\frac{\gamma}{4} - (\kappa^2 + \kappa(1 + \gamma/2))) \|\nabla_t\|^2,$$

where we used the variance bound $\|e_t\| \leq \kappa \|\nabla_t\|$ as specified in the theorem. Using the α-strong-convexity of $\mathcal{R}_{emp}[f]$, we get,

$$\mathcal{R}_{emp}[f^*] - \mathcal{R}_{emp}[f_t] \geq -\frac{1}{2\alpha} \|\hat{\nabla}_t\|^2.$$

Substituting the bound for $\|\hat{\nabla}_t\|^2$ in the previous inequality, we get,

$$\begin{aligned}
\mathbb{E}[\mathcal{R}_{emp}[f_{t+1}]] - \mathcal{R}_{emp}[f^*] \leq{} & \mathbb{E}[\mathcal{R}_{emp}[f_t]] - \mathcal{R}_{emp}[f^*] \\
& - 2\alpha/\beta(\frac{\gamma}{4} - (\kappa^2 + \kappa(1 + \gamma/2))) \mathbb{E}[\mathcal{R}_{emp}[f_t]] - \mathcal{R}_{emp}[f^*],
\end{aligned}$$

from which the theorem follows.