A. Proofs

A.1. Optimistic Follow-the-Regularized-Leader

We offer a proof of Theorem 8.

First, we introduce the following argmin-function:

$$\tilde{x} : \boldsymbol{L} \mapsto \underset{\boldsymbol{x}\in\mathcal{X}}{\operatorname{argmin}} \left\{ \langle \boldsymbol{x}, \boldsymbol{L} \rangle + \frac{1}{\eta} R(\boldsymbol{x}) \right\}. \tag{18}$$

Furthermore, let $\boldsymbol{L}^t := \sum_{\tau=1}^{t} \boldsymbol{\ell}^\tau$. With this notation, the decisions produced by OFTRL, as defined in (8), can be expressed as $\boldsymbol{x}^t = \tilde{x}(\boldsymbol{L}^{t-1} + \boldsymbol{m}^t)$.

Continuity of the argmin-function. The first step in the proof is to study the continuity of the argmin-function $\tilde{x}$. Intuitively, the role of the regularizer R is to *smooth out* the linear objective function $\langle \cdot, \boldsymbol{L} \rangle$. So, it seems only reasonable to expect that, the higher the constant that multiplies R, the less the argmin $\tilde{x}(\boldsymbol{L})$ is affected by small changes of $\boldsymbol{L}$. In fact, the following holds:

Lemma 5. *The argmin-function $\tilde{x}$ is η-Lipschitz continuous with respect to the dual norm, that is*

$$\|\tilde{x}(\boldsymbol{L}) - \tilde{x}(\boldsymbol{L}')\| \le \eta \|\boldsymbol{L} - \boldsymbol{L}'\|_*.$$

Proof. The variational inequality for the optimality of $\tilde{x}(\boldsymbol{L})$ implies

$$\left\langle L + \frac{1}{\eta} \nabla R(\tilde{x}(\boldsymbol{L})), \tilde{x}(\boldsymbol{L}') - \tilde{x}(\boldsymbol{L}) \right\rangle \ge 0. \tag{19}$$

Symmetrically for $\tilde{x}(\boldsymbol{L}')$, we find that

$$\left\langle L' + \frac{1}{\eta} R(\tilde{x}(\boldsymbol{L}')), \tilde{x}(\boldsymbol{L}) - \tilde{x}(\boldsymbol{L}') \right\rangle \ge 0. \tag{20}$$

Summing inequalities 19 and 20, we obtain

$$\frac{1}{\eta} \left\langle \nabla R(\tilde{x}(\boldsymbol{L})) - \nabla R(\tilde{x}(\boldsymbol{L}')), \tilde{x}(\boldsymbol{L}) - \tilde{x}(\boldsymbol{L}') \right\rangle \le \left\langle \boldsymbol{L}' - \boldsymbol{L}, \tilde{x}(\boldsymbol{L}) - \tilde{x}(\boldsymbol{L}') \right\rangle.$$

Using strong convexity of $R(\cdot)$ on the left-hand side and the generalized Cauchy-Schwarz inequality on the right-hand side, we obtain

$$\frac{1}{\eta} \|\tilde{x}(\boldsymbol{L}) - \tilde{x}(\boldsymbol{L}')\|^2 \le \|\tilde{x}(\boldsymbol{L}) - \tilde{x}(\boldsymbol{L}')\| \, \|\boldsymbol{L} - \boldsymbol{L}'\|_*,$$

and dividing by $\|\tilde{x}(\boldsymbol{L}) - \tilde{x}(\boldsymbol{L}')\|$ we obtain the Lipschitz continuity of the argmin-function $\tilde{x}$. □

A direct consequence of Lemma 5 is the following corollary, which measures the stability (small step size) of the decisions output by OFTRL:

Corollary 2. *At each time t, the iterates produced by OFTRL satisfy $\|\boldsymbol{x}^t - \boldsymbol{x}^{t-1}\| \le 3\eta\Delta_\ell$.*

Proof.

$$\begin{aligned} \|\boldsymbol{x}^t - \boldsymbol{x}^{t-1}\| &= \left\| \tilde{x}(\boldsymbol{L}^{t-1} + \boldsymbol{m}^t) - \tilde{x}(\boldsymbol{L}^{t-2} + \boldsymbol{m}^{t-1}) \right\| \\ &\le \eta \|\boldsymbol{\ell}^{t-1} + \boldsymbol{m}^t - \boldsymbol{m}^{t-1}\|_* \le 3\eta\Delta_\ell, \end{aligned}$$

where the first inequality holds by Lemma 5 and the second one by definition of Δ_ℓ and the triangle inequality. □

The rest of the proof, specifically the predictivity parameters α and β of OFTRL follow directly from the proof of Theorem 19 in the appendix of Syrgkanis et al. (2015).

A.2. Regret Bounds

Lemma 1. *For all $k \in \mathcal{K}$, $R_k^{\triangle,T} = \sum_{j \in \mathcal{C}_k} R_j^{\triangle,T}$.*

Proof. By definition of $R_k^{\triangle,T}$,

$$R_k^{\triangle,T} = \sum_{t=1}^{T} \langle \boldsymbol{\ell}_k^{\triangle,t}, \boldsymbol{x}_k^{\triangle,t} \rangle - \min_{\tilde{\boldsymbol{x}}_k^{\triangle} \in X_k^{\triangle}} \sum_{t=1}^{T} \langle \boldsymbol{\ell}_k^{\triangle,t}, \tilde{\boldsymbol{x}}_k^{\triangle} \rangle.$$

By using (12) and (11), we can break the dot products and the minimization problem into independent parts, one for each $j \in \mathcal{C}_k$:

$$\begin{aligned} R_k^{\triangle,T} &= \sum_{j \in \mathcal{C}_k} \sum_{t=1}^{T} \langle \boldsymbol{\ell}_j^{\triangle,t}, \boldsymbol{x}_j^{\triangle,t} \rangle - \sum_{j \in \mathcal{C}_k} \min_{\tilde{\boldsymbol{x}}_j^{\triangle} \in X_j^{\triangle}} \sum_{t=1}^{T} \langle \boldsymbol{\ell}_j^{\triangle,t}, \tilde{\boldsymbol{x}}_j^{\triangle} \rangle \\ &= \sum_{j \in \mathcal{C}_k} \left(\sum_{t=1}^{T} \langle \boldsymbol{\ell}_j^{\triangle,t}, \boldsymbol{x}_j^{\triangle,t} \rangle - \min_{\tilde{\boldsymbol{x}}_j^{\triangle} \in X_j^{\triangle}} \sum_{t=1}^{T} \langle \boldsymbol{\ell}_j^{\triangle,t}, \tilde{\boldsymbol{x}}_j^{\triangle} \rangle \right) \\ &= \sum_{j \in \mathcal{C}_k} R_j^{\triangle,T}, \end{aligned}$$

as we wanted to show. □

Lemma 2. *For all $j \in \mathcal{J}$, $R_j^{\triangle,T} \le \hat{R}_j^T + \max_{k \in \mathcal{C}_j} R_k^{\triangle,T}$.*

Proof. By definition of $R_j^{\triangle,T}$,

$$R_j^{\triangle,T} = \sum_{t=1}^{T} \langle \boldsymbol{\ell}_j^{\triangle,t}, \boldsymbol{x}_j^{\triangle,t} \rangle - \min_{\tilde{\boldsymbol{x}}_j^{\triangle} \in X_j^{\triangle}} \sum_{t=1}^{T} \langle \boldsymbol{\ell}_j^{\triangle,t}, \tilde{\boldsymbol{x}}_j^{\triangle} \rangle.$$

By combining (13) and (11), we can break the dot products and the minimization problem into independent parts, one for each $k \in \mathcal{C}_j$, as well as a part that depends solely on $\hat{\boldsymbol{x}}_j$:

$$\begin{aligned} R_j^{\triangle,T} = \sum_{t=1}^{T} &\left(\langle [\boldsymbol{\ell}_j^{\triangle,t}]_j, \hat{\boldsymbol{x}}_j^t \rangle + \sum_{\substack{a \in A_j \\ k=\rho(j,a)}} \hat{\boldsymbol{x}}_{ja}^t \langle \boldsymbol{\ell}_k^{\triangle,t}, \boldsymbol{x}_k^{\triangle,t} \rangle \right) \\ &\qquad - \min_{\tilde{\boldsymbol{x}}_j \in \Delta^{n_j}} \left\{ \left(\sum_{t=1}^{T} \langle [\boldsymbol{\ell}_j^{\triangle,t}]_j, \tilde{\boldsymbol{x}}_j \rangle \right) + \sum_{\substack{a \in A_j \\ k=\rho(j,a)}} \tilde{\boldsymbol{x}}_{ja} \left(\min_{\tilde{\boldsymbol{x}}_k^{\triangle} \in X_k^{\triangle}} \sum_{t=1}^{T} \langle \boldsymbol{\ell}_k^{\triangle,t}, \tilde{\boldsymbol{x}}_k^{\triangle} \rangle \right) \right\} \\ = \sum_{t=1}^{T} &\left(\langle [\boldsymbol{\ell}_j^{\triangle,t}]_j, \hat{\boldsymbol{x}}_j^t \rangle + \sum_{\substack{a \in A_j \\ k=\rho(j,a)}} \hat{\boldsymbol{x}}_{ja}^t \langle \boldsymbol{\ell}_k^{\triangle,t}, \boldsymbol{x}_k^{\triangle,t} \rangle \right) \\ &\qquad - \min_{\tilde{\boldsymbol{x}}_j \in \Delta^{n_j}} \left\{ \left(\sum_{t=1}^{T} \langle [\boldsymbol{\ell}_j^{\triangle,t}]_j, \tilde{\boldsymbol{x}}_j \rangle \right) + \sum_{\substack{a \in A_j \\ k=\rho(j,a)}} \tilde{\boldsymbol{x}}_{ja} \left(-R_k^{\triangle,T} + \sum_{t=1}^{T} \langle \boldsymbol{\ell}_k^{\triangle,t}, \boldsymbol{x}_k^{\triangle,t} \rangle \right) \right\} \end{aligned}$$

$$\leq \sum_{t=1}^{T} \left(\langle [\boldsymbol{\ell}_j^{\triangle,t}]_j, \hat{\boldsymbol{x}}_j^t \rangle + \sum_{\substack{a \in A_j \\ k=\rho(j,a)}} \hat{\boldsymbol{x}}_{ja}^t \langle \boldsymbol{\ell}_k^{\triangle,t}, \boldsymbol{x}_k^{\triangle,t} \rangle \right)$$

$$- \min_{\tilde{\boldsymbol{x}}_j \in \Delta^{n_j}} \left\{ \sum_{t=1}^{T} \left(\langle [\boldsymbol{\ell}_j^{\triangle,t}]_j, \tilde{\boldsymbol{x}}_j \rangle + \sum_{\substack{a \in A_j \\ k=\rho(j,a)}} \tilde{\boldsymbol{x}}_{ja} \langle \boldsymbol{\ell}_k^{\triangle,t}, \boldsymbol{x}_k^{\triangle,t} \rangle \right) \right\} + \max_{\tilde{\boldsymbol{x}}_j \in \Delta^{n_j}} \sum_{a \in A_j} \tilde{\boldsymbol{x}}_{ja} R_k^{\triangle,T},$$

where the equality follows by the definition of $R_k^{\triangle,T}$, and the inequality follows from breaking the minimization of a sum into a sum of minimization problems. By identifying the difference between the first two terms as the counterfactual regret $\hat{R}_j^T$ (that is, the regret of $\hat{\mathcal{R}}_j$ up to time T), we obtain

$$R_j^{\triangle,T} \leq \hat{R}_j^T + \max_{\tilde{\boldsymbol{x}}_j \in \Delta^{n_j}} \sum_{k \in \mathcal{C}_j} \tilde{\boldsymbol{x}}_{ja} R_k^{\triangle,T} = \hat{R}_j^T + \max_{k \in \mathcal{C}_j} R_k^{\triangle,T},$$

as we wanted to show. □

A.3. Stable-Predictive Regret Minimizer

We will prove both Lemma 3 and Lemma 4 with respect to the 2-norm. This does not come at the cost of generality, since all norms are equivalent on finite-dimensional vector spaces, that is, for every choice of norm $\|\cdot\|$, there exist constants $m, M > 0$ such that for all $\boldsymbol{x}$, $m\|\boldsymbol{x}\| \leq \|\boldsymbol{x}\|_2 \leq M\|\boldsymbol{x}\|$.

Lemma 3. *Let $k \in \mathcal{K}$ be an observation node, and assume that $\mathcal{R}_j^{\triangle}$ is a $(\gamma_j, O(1), O(1))$-stable-predictive regret minimizer over the sequence-form strategy space $X_j^{\triangle}$ for each $j \in \mathcal{C}_k$. Then, $\mathcal{R}_k^{\triangle}$ is a $(\gamma_k, O(1), O(1))$-stable-predictive regret minimizer over the sequence-form strategy space $X_k^{\triangle}$.*

Proof. By hypothesis, for all $j \in \mathcal{C}_k$ we have

$$R_j^{\triangle,T} \leq \frac{O(1)}{\gamma_j} + O(1)\gamma_j \sum_{t=1}^{T} \|\boldsymbol{\ell}_j^{\triangle,t} - \boldsymbol{m}_j^{\triangle,t}\|_2^2 \tag{21}$$

and

$$\|\boldsymbol{x}_j^{\triangle,t} - \boldsymbol{x}_j^{\triangle,t-1}\|_2 \leq \gamma_j, \tag{22}$$

where $\boldsymbol{x}_j^{\triangle,t}$ is the decision output by $\mathcal{R}_j^{\triangle}$ at time t.

Substituting (21) into the regret bound of Lemma 1:

$$\begin{aligned} R_k^{\triangle,T} &\leq O(1) \sum_{j \in \mathcal{C}_k} \frac{1}{\gamma_j} + O(1) \sum_{j \in \mathcal{C}_k} \sum_{t=1}^{T} \gamma_j \|\boldsymbol{\ell}_j^{\triangle,t} - \boldsymbol{m}_j^{\triangle,t}\|_2^2 \\ &\leq O(1) \frac{n_k^{3/2}}{\gamma_k} + O(1) \frac{\gamma_k}{\sqrt{n_k}} \sum_{t=1}^{T} \sum_{j \in \mathcal{C}_k} \|\boldsymbol{\ell}_j^{\triangle,t} - \boldsymbol{m}_j^{\triangle,t}\|_2^2 \\ &= \frac{O(1)}{\gamma_k} + O(1)\gamma_k \sum_{t=1}^{T} \|\boldsymbol{\ell}_k^{\triangle,t} - \boldsymbol{m}_k^{\triangle,t}\|_2^2 \end{aligned} \tag{23}$$

where the second inequality comes from substituting the value $\gamma_j = \gamma_k/\sqrt{n_k}$ as per (14), and the equality comes from the fact that the $\boldsymbol{\ell}_j^{\triangle,t}$ and $\boldsymbol{m}_j^{\triangle,t}$ form a partition of the vectors $\boldsymbol{\ell}_k^{\triangle,t}$ and $\boldsymbol{m}_k^{\triangle,t}$, respectively.

We now analyze the stability properties of $\mathcal{R}_k^{\triangle}$:

$$\|\boldsymbol{x}_k^{\triangle,t} - \boldsymbol{x}_k^{\triangle,t-1}\|_2 = \sqrt{\sum_{j \in \mathcal{C}_k} \|\boldsymbol{x}_j^{\triangle,t} - \boldsymbol{x}_j^{\triangle,t-1}\|_2^2} \leq \sqrt{\sum_{j \in \mathcal{C}_k} \gamma_j^2} = \gamma_k,$$

where the first equality follows from (1), the inequality holds by (22) and the second equality holds by substituting the value $\gamma_j = \gamma_k/\sqrt{n_k}$ as per (14). This shows that $\mathcal{R}_k^{\triangle}$ is γ_k-stable. Combining this with the predictivity bound (23) above, we obtain the claim. □

Lemma 4. *Let $j \in \mathcal{J}$ be a decision node, and assume that $\mathcal{R}_k^{\triangle}$ is a $(\gamma_k, O(1), O(1))$-stable-predictive regret minimizer over the sequence-form strategy space $X_k^{\triangle}$ for each $k \in \mathcal{C}_j$. Suppose further that $\hat{\mathcal{R}}_j$ is a $(\kappa_j, O(1), O(1))$-stable-predictive regret minimizer over the simplex Δ^{n_j}. Then, $\mathcal{R}_j^{\triangle}$ is a $(\gamma_k, O(1), O(1))$-stable-predictive regret minimizer over the sequence-form strategy space $X_j^{\triangle}$.*

Proof. By hypothesis, for all $k \in \mathcal{C}_j$ we have

$$R_k^{\triangle,T} \le \frac{O(1)}{\gamma_k} + O(1)\gamma_k \sum_{t=1}^{T} \|\boldsymbol{\ell}_k^{\triangle,t} - \boldsymbol{m}_k^{\triangle,t}\|_2^2 \tag{24}$$

and

$$\|\boldsymbol{x}_k^{\triangle,t} - \boldsymbol{x}_k^{\triangle,t-1}\|_2 \le \gamma_k. \tag{25}$$

We substitute (24) into the regret bound of Lemma 2. The key observation is that the loss vector—and their predictions—entering the subtree rooted at k ($k \in \mathcal{C}_j$) are simply forwarded from j; with this, we obtain:

$$R_{\triangle_j}^T \le \hat{R}_j^T + \frac{O(1)}{\gamma_k} + O(1)\gamma_k \sum_{t=1}^{T} \|\boldsymbol{\ell}_j^{\triangle,t} - \boldsymbol{m}_j^{\triangle,t}\|_2^2. \tag{26}$$

On the other hand, by hypothesis $\hat{\mathcal{R}}_j$ is a $(\kappa_j, O(1), O(1))$-stable-predictive regret minimizer. Hence,

$$\begin{aligned} \hat{R}_j^T &\le \frac{O(1)}{\kappa_j} + O(1)\kappa_j \sum_{t=1}^{T} \|\hat{\boldsymbol{\ell}}_j^t - \hat{\boldsymbol{m}}_j^t\|_2^2 \\ &= \frac{O(1)}{\gamma_j} + O(1)\gamma_j \sum_{t=1}^{T} \|\boldsymbol{\ell}_j^{\triangle,t} - \boldsymbol{m}_j^{\triangle,t}\|_2^2, \end{aligned} \tag{27}$$

where the equality comes from the definition of κ_j (Equation (15)) and the fact that

$$\begin{aligned} \|\hat{\boldsymbol{\ell}}_j^t - \hat{\boldsymbol{m}}_j^t\|_2^2 &\le \sum_{k \in \mathcal{C}_j} \|\boldsymbol{x}_k^{\triangle,t}\|_2^2 \cdot \|\boldsymbol{\ell}_k^{\triangle,t} - \boldsymbol{m}_k^{\triangle,t}\|_2^2 \\ &\le \|\boldsymbol{\ell}_j^{\triangle,t} - \boldsymbol{m}_j^{\triangle,t}\|_2^2 \sum_{k \in \mathcal{C}_j} B_k^2 \\ &= O(1)\|\boldsymbol{\ell}_j^{\triangle,t} - \boldsymbol{m}_j^{\triangle,t}\|_2^2. \end{aligned}$$

By substituting (27) into (26) and noting that $\gamma_k = O(1)\gamma_j$, we obtain

$$R_j^{\triangle,T} \le \frac{O(1)}{\gamma_j} + O(1)\gamma_j \sum_{t=1}^{T} \|\boldsymbol{\ell}_j^{\triangle,t} - \boldsymbol{m}_j^{\triangle,t}\|_2^2,$$

which establishes the predictivity of $\mathcal{R}_j^{\triangle}$.

To conclude the proof, we show that $\mathcal{R}_j^{\triangle}$ has stability parameter γ_j. To this end, note that by (2)

$$\begin{aligned} \|\boldsymbol{x}_j^{\triangle,t} - \boldsymbol{x}_j^{\triangle,t-1}\|_2^2 &= \left\| \left(\sum_{a \in A_j} \hat{\boldsymbol{x}}_{ja}^t \boldsymbol{x}_{ja}^{\triangle,t} \right) - \left(\sum_{a \in A_j} \hat{\boldsymbol{x}}_{ja}^{t-1} \boldsymbol{x}_{ja}^{\triangle,t-1} \right) \right\|_2^2 + \|\hat{\boldsymbol{x}}_j^t - \hat{\boldsymbol{x}}_j^{t-1}\|_2^2 \\ &\le \|\hat{\boldsymbol{x}}_j^t - \hat{\boldsymbol{x}}_j^{t-1}\|_2^2 \left(1 + 2 \sum_{k \in \mathcal{C}_j} \|\boldsymbol{x}_k^{\triangle,t}\|_2^2 \right) + 2 \sum_{k \in \mathcal{C}_k} \|\boldsymbol{x}_k^{\triangle,t} - \boldsymbol{x}_k^{\triangle,t-1}\|_2^2 \\ &\le 2n_j B_j^2 \|\hat{\boldsymbol{x}}_j^t - \hat{\boldsymbol{x}}_j^{t-1}\|_2^2 + 2 \sum_{k \in \mathcal{C}_k} \|\boldsymbol{x}_k^{\triangle,t} - \boldsymbol{x}_k^{\triangle,t-1}\|_2^2, \end{aligned}$$

where we have used the Cauchy-Schwarz inequality and the definition of B_j (Equation 16). By using the stability of $\hat{\mathcal{R}}_j$, that is $\|\hat{\boldsymbol{x}}_j^t - \hat{\boldsymbol{x}}_j^{t-1}\|_2^2 \le \kappa_j^2 = \gamma_j^2/(4n_j B_j^2)$, as well as the hypothesis (25) and (14):

$$\|\boldsymbol{x}_j^{\triangle,t} - \boldsymbol{x}_j^{\triangle,t-1}\|_2 \le \frac{\gamma_j^2}{2} + 2 \sum_{k \in \mathcal{C}_j} \left(\frac{\gamma_j}{2\sqrt{n_j}} \right)^2 = \frac{\gamma_j^2}{2} + 2n_j \left(\frac{\gamma_j}{2\sqrt{n_j}} \right)^2 = \gamma_j^2.$$

Hence, $\mathcal{R}_j^{\triangle}$ has stability parameter γ_j as we wanted to show. $\square$

B. Experiments

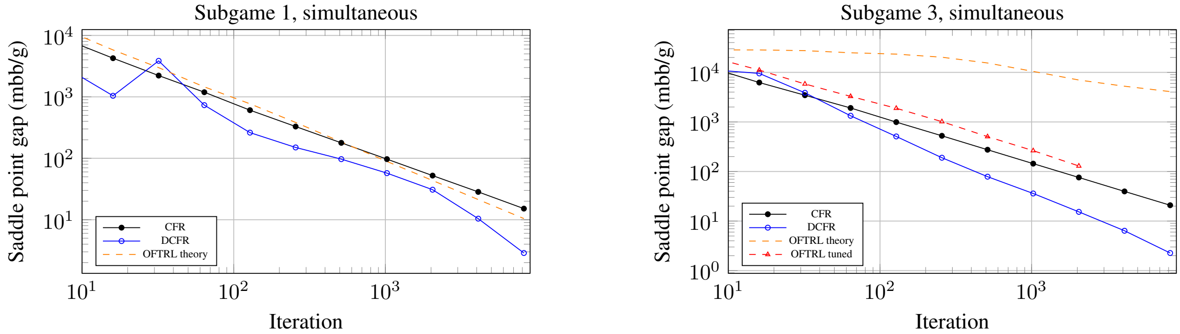

Figure 4. Convergence rate with iterations on the x-axis, and the exploitability in mbb. All algorithms use simultaneous updates.

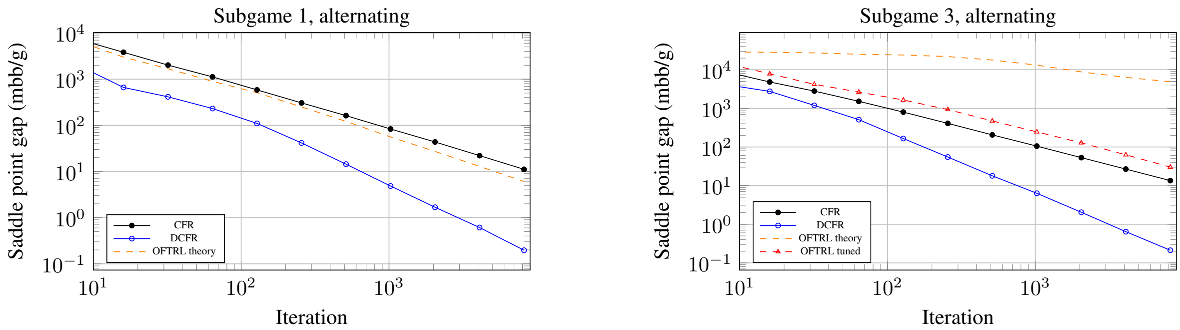

Figure 5. Convergence rate with iterations on the x-axis, and the exploitability in mbb. All algorithms use alternating updates.