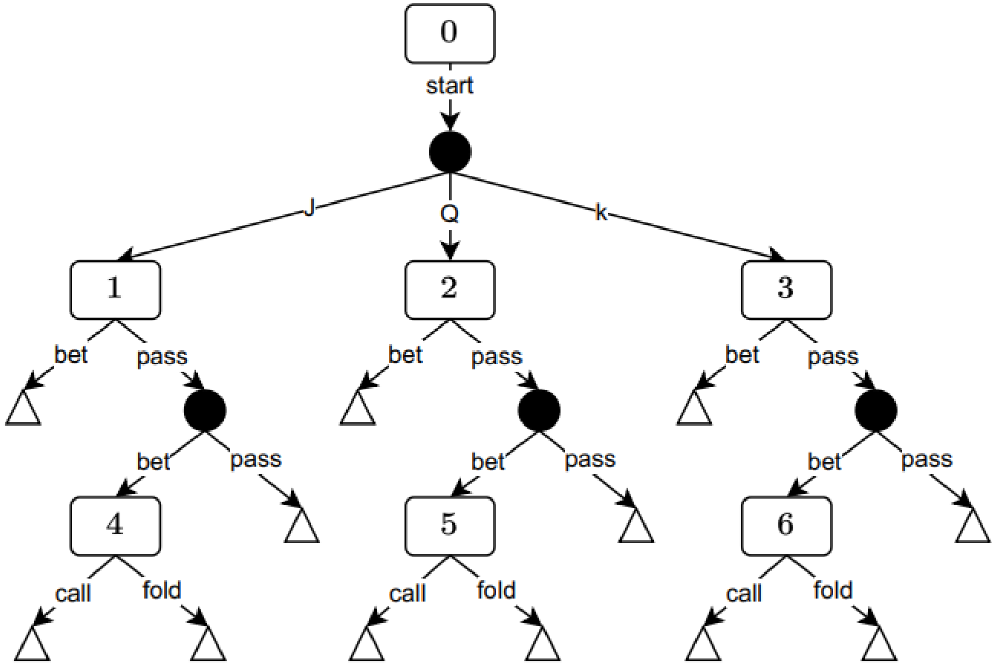
Equivalence Analysis between Counterfactual Regret Minimization and Online Mirror Descent

Weiming Liu[*], Huacong Jiang, Bin Li[*], Houqiang Li

University of Science and Technology of China

Introduction

- Extensive-Form Games (**EFGs**)
  - ⇒ Sequential Decision Process



Decision process in Kuhn poker

Introduction

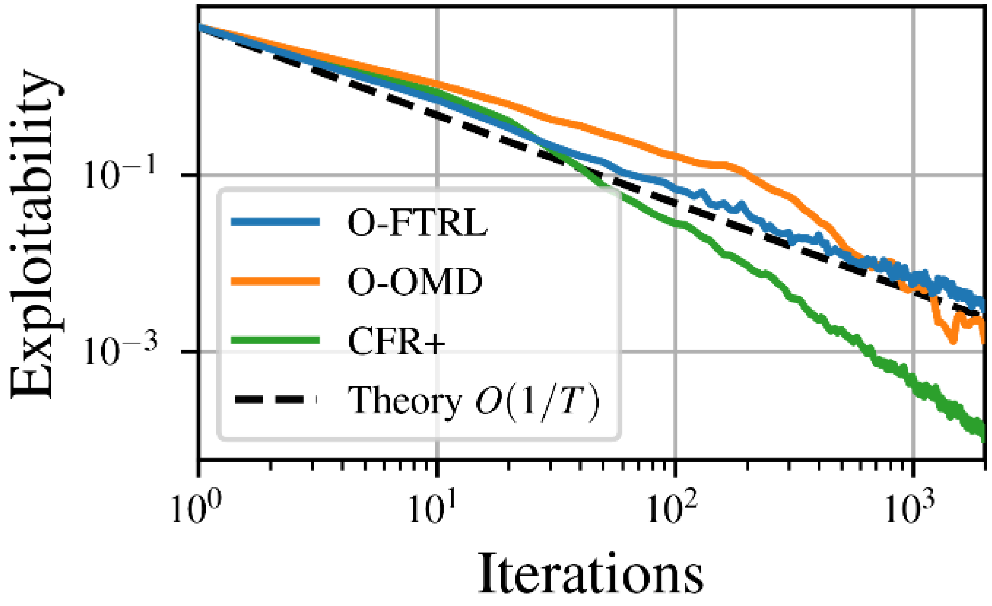
- Extensive-Form Games (**EFGs**)
  - ⇒ Sequential Decision Process
- **Algorithms:**
  - Counterfactual Regret Minimization (**CFR**) (Zinkevich et al., 2007)
  - Follow-the-Regularized-Leader (**FTRL**) (Abernethy et al., 2008)
  - Online Mirror Descent (**OMD**) (Beck & Teboulle, 2003)
  - Excessive Gap Technique (Hoda et al., 2010; Kroer et al., 2020), …
- Target:
  - Converge to a Nash Equilibrium
  - ⇒ Exploitability to zero

Introduction

- Inconsistency between theory and practice
- Theory:
  - CFR+ (Tammelin et al., 2015): $O(1/\sqrt{T})$
  - Optimistic OMD (O-OMD) (Farina et al., 2019): $O(1/T)$

Introduction

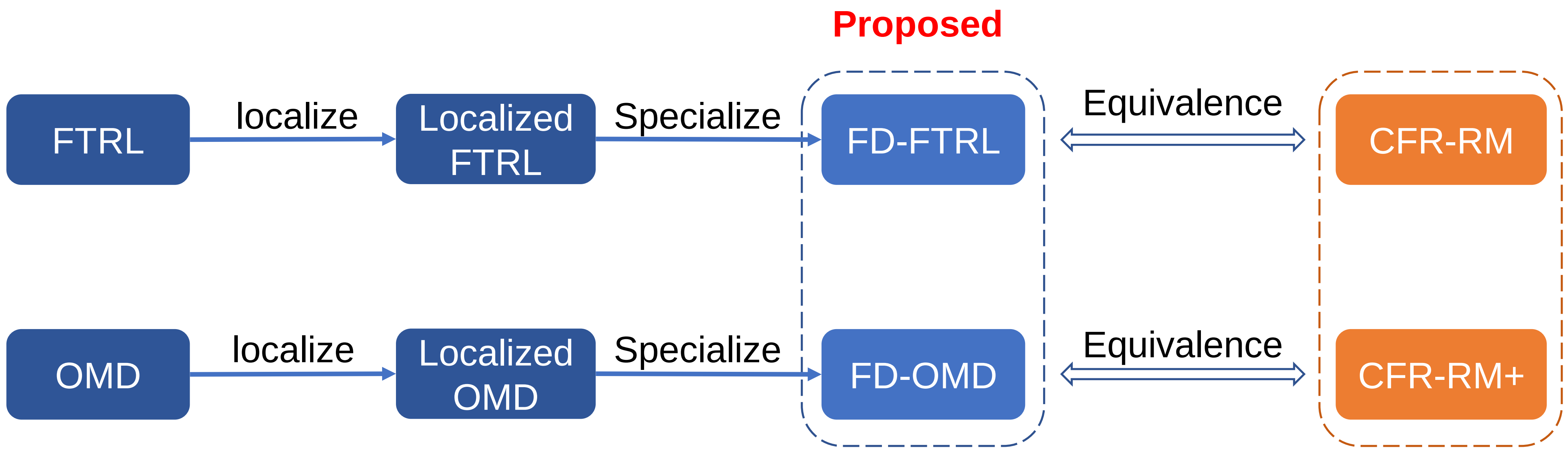
- Inconsistency between theory and practice
- Theory:
  - CFR+ (Tammelin et al., 2015): $O(1/\sqrt{T})$
  - Optimistic OMD (O-OMD) (Farina et al., 2019): $O(1/T)$
- Practice:
  - CFR+ (**green line**) is much **faster**.



Introduction

- Questions:
  - How to explain the superior performance of CFR?
  - How to improve the practical performance of FTRL (OMD)?
  - **Are there strong connections between CFR and FTRL (OMD)?**
- Contributions:
  - Prove that CFR-RM (CFR-RM+) is equivalent to a special FTRL (OMD)
  - Propose a special OMD, which is **faster** than CFR+ in some EFGs.

Equivalence Analysis：Overview



Equivalence Analysis: Localization

- Localization (Hoda et al., 2010; Farina et al.,2019b)

Algorithm 1 FTRL

```
for iteration $t = 1$ to $T$ do
    $\boldsymbol{l}^t \leftarrow \text{ObserveLoss}(\boldsymbol{x}^t)$.
    $\boldsymbol{x}^{t+1} \leftarrow \text{argmin}_{\boldsymbol{x} \in \mathcal{X}} \left\{ \langle \boldsymbol{L}^t, \boldsymbol{x} \rangle + q^{0:t}(\boldsymbol{x}) \right\}$.
end for
```

Compute the strategy as a whole

Algorithm 3 Localized (FD-)FTRL

```
for iteration $t = 1$ to $T$ do
    $\boldsymbol{l}^t \leftarrow \text{ObserveLoss}(\boldsymbol{x}^t)$.
    for node $j \in \mathcal{J}$ in bottom-up order do
        $\hat{\boldsymbol{L}}_j'^t[a] \leftarrow \boldsymbol{L}^t[j, a] + \sum_{j' \in C_{j,a}} -\psi_{j'}^{*t}(-\hat{\boldsymbol{L}}_{j'}'^t)$.
        $\hat{\boldsymbol{x}}_j^{t+1} \leftarrow \nabla \psi_j^{*t}(-\hat{\boldsymbol{L}}_j'^t)$.
    end for
    Construct $\boldsymbol{x}^{t+1}$.
end for
```

Recursively compute local decisions

Equivalence Analysis: Localization

- Localization (Hoda et al., 2010; Farina et al.,2019b)
  - ψ is the regularizer, e.g., l_2 norm function

Algorithm 3 Localized (FD-)FTRL

1: **for** iteration $t = 1$ to T **do**
2: $\quad \boldsymbol{l}^t \leftarrow \text{ObserveLoss}(\boldsymbol{x}^t)$.
3: $\quad$ **for** node $j \in \mathcal{J}$ in bottom-up order **do**
4: $\quad\quad \hat{\boldsymbol{L}}_j'^t[a] \leftarrow \boldsymbol{L}^t[j,a] + \sum_{j' \in C_{j,a}} -\psi_{j'}^{*t}(-\hat{\boldsymbol{L}}_{j'}'^t)$.
5: $\quad\quad \hat{\boldsymbol{x}}_j^{t+1} \leftarrow \nabla \psi_j^{*t}(-\hat{\boldsymbol{L}}_j'^t)$.
6: $\quad$ **end for**
7: $\quad$ Construct $\boldsymbol{x}^{t+1}$
8: **end for**

Convex conjugate:

$\psi_j^{*t}(\hat{\boldsymbol{g}}) = \max_{\hat{\boldsymbol{x}}_j \in \Delta^{n_j}} \{\langle \hat{\boldsymbol{g}}, \hat{\boldsymbol{x}}_j \rangle - \psi_j^t(\hat{\boldsymbol{x}}_j)\}$

$\psi_j^t(\hat{\boldsymbol{x}}_j) = \frac{1}{2}\beta_j^t \|\hat{\boldsymbol{x}}_j\|_2^2$

Equivalence Analysis: Localization

- Localization (Hoda et al., 2010; Farina et al.,2019b)
  - ψ is the regularizer, e.g., l_2 norm function
  - **Is there a regularizer such that FTRL can recover CFR-RM?**

Algorithm 3 Localized (FD-)FTRL

for iteration $t = 1$ to T **do**
  $\boldsymbol{l}^t \leftarrow \text{ObserveLoss}(\boldsymbol{x}^t)$.
  for node $j \in \mathcal{J}$ in bottom-up order **do**
    $\hat{\boldsymbol{L}}_j'^t[a] \leftarrow \boldsymbol{L}^t[j, a] + \sum_{j' \in C_{j,a}} -\psi_{j'}^{*t}(-\hat{\boldsymbol{L}}_{j'}'^t)$.
    $\hat{\boldsymbol{x}}_j^{t+1} \leftarrow \nabla \psi_j^{*t}(-\hat{\boldsymbol{L}}_j'^t)$.
  end for
  Construct $\boldsymbol{x}^{t+1}$
end for

Convex conjugate:

$\psi_j^{*t}(\hat{\boldsymbol{g}}) = \max_{\hat{\boldsymbol{x}}_j \in \Delta^{n_j}} \{\langle \hat{\boldsymbol{g}}, \hat{\boldsymbol{x}}_j \rangle - \psi_j^t(\hat{\boldsymbol{x}}_j)\}$

$\psi_j^t(\hat{\boldsymbol{x}}_j) = \frac{1}{2}\beta_j^t \|\hat{\boldsymbol{x}}_j\|_2^2$

Algorithm 2 CFR-RM

for iteration $t = 1$ to T **do**
  $\boldsymbol{l}^t \leftarrow \text{ObserveLoss}(\boldsymbol{x}^t)$.
  for node $j \in \mathcal{J}$ in bottom-up order **do**
    $\hat{\boldsymbol{l}}_j^t[a] \leftarrow \boldsymbol{l}^t[j, a] + \sum_{j' \in C_{j,a}} \langle \hat{\boldsymbol{l}}_{j'}^t, \hat{\boldsymbol{x}}_{j'}^t \rangle$.
    $\hat{\boldsymbol{R}}_j^t \leftarrow \hat{\boldsymbol{R}}_j^{t-1} + \langle \hat{\boldsymbol{l}}_j^t, \hat{\boldsymbol{x}}_j^t \rangle \mathbf{1} - \hat{\boldsymbol{l}}_j^t$, } RM
    $\hat{\boldsymbol{x}}_j^{t+1} \leftarrow [\hat{\boldsymbol{R}}_j^t]^+ / \|[\hat{\boldsymbol{R}}_j^t]^+\|_1$. } RM
  end for
  Construct $\boldsymbol{x}^{t+1}$.
end for

Equivalence Analysis：Specialization

- Future-Dependent FTRL （ FD-FTRL）
  - The regularizer is dependent on the next decision $\hat{\boldsymbol{x}}_j^{t+1}$.
  - **Note**: the next decision does not depend on itself.

Algorithm 3 Localized (FD-)FTRL

for iteration $t = 1$ to T **do**
  $\boldsymbol{l}^t \leftarrow \text{ObserveLoss}(\boldsymbol{x}^t)$.
  for node $j \in \mathcal{J}$ in bottom-up order **do**
    $\hat{\boldsymbol{L}}_j'^t[a] \leftarrow \boldsymbol{L}^t[j, a] + \sum_{j' \in C_{j,a}} -\psi_{j'}^{*t}(-\hat{\boldsymbol{L}}_{j'}'^t)$.
    $\hat{\boldsymbol{x}}_j^{t+1} \leftarrow \nabla \psi_j^{*t}(-\hat{\boldsymbol{L}}_j'^t)$.
  end for
  Construct $\boldsymbol{x}^{t+1}$.
end for

$\psi_j^t(\hat{\boldsymbol{x}}_j) = \frac{1}{2}\beta_j^t \|\hat{\boldsymbol{x}}_j\|_2^2 + \frac{1}{2}\beta_j^t \|\hat{\boldsymbol{x}}_j^{t+1}\|_2^2$

Equivalence Analysis：Theorem

- CFR-RM (CFR-RM+) $\Leftrightarrow$ a special FD-FTRL (FD-OMD)
  - CFR-RM (CFR-RM+) is an **adaptive** FTRL (OMD).
  - FTRL(OMD) with adaptation could be faster.

Theorem 3.7. *CFR-RM (CFR-RM+) is equivalent to a special case of FD-FTRL (FD-OMD) with* $\beta_j^t = \|[\hat{\boldsymbol{R}}_j^t]^+\|_1$ *(*$\|\hat{\boldsymbol{Q}}_j^t\|_1$*),* $\forall j \in \mathcal{J}, t \geq 0$.

Application：Adaptive FD-FTRL (FD-OMD)

- Reparametrized FD-FTRL (FD-OMD) with **adaptation**

Algorithm 3 Localized (FD-)FTRL

for iteration $t = 1$ to T **do**
  $\boldsymbol{l}^t \leftarrow \text{ObserveLoss}(\boldsymbol{x}^t)$.
  for node $j \in \mathcal{J}$ in bottom-up order **do**
    $\hat{\boldsymbol{L}}_j^{\prime t}[a] \leftarrow \boldsymbol{L}^t[j,a] + \sum_{j' \in C_{j,a}} -\psi_{j'}^{*t}(-\hat{\boldsymbol{L}}_{j'}^{\prime t})$.
    $\hat{\boldsymbol{x}}_j^{t+1} \leftarrow \nabla\psi_j^{*t}(-\hat{\boldsymbol{L}}_j^{\prime t})$.
  end for
  Construct $\boldsymbol{x}^{t+1}$.
end for

Algorithm 4 Reparameterized FD-FTRL

for iteration $t = 1$ to T **do**
  $\boldsymbol{l}^t \leftarrow \text{ObserveLoss}(\boldsymbol{x}^t)$.
  for node $j \in \mathcal{J}$ in bottom-up order **do**
    $\hat{\boldsymbol{L}}_j^{\prime t}[a] \leftarrow \boldsymbol{L}^t[j,a] + \sum_{j' \in C_{j,a}} \alpha_{j'}^t$.
    $\hat{\boldsymbol{R}}_j^{\prime t} \leftarrow \alpha_j^t \mathbf{1} - \hat{\boldsymbol{L}}_j^{\prime t}$.
    $\hat{\boldsymbol{x}}_j^{t+1} \leftarrow [\hat{\boldsymbol{R}}_j^{\prime t}]^+ / \|[\hat{\boldsymbol{R}}_j^{\prime t}]^+\|_1$.
  end for
  Construct $\boldsymbol{x}^{t+1}$.
end for

Application：Adaptive FD-FTRL (FD-OMD)

- Reparametrized FD-FTRL (FD-OMD) with **adaptation**
- Adapt the parameters like in CFR-RM (CFR-RM+)

Algorithm 3 Localized (FD-)FTRL

for iteration $t = 1$ to T **do**
  $\boldsymbol{l}^t \leftarrow \text{ObserveLoss}(\boldsymbol{x}^t)$.
  for node $j \in \mathcal{J}$ in bottom-up order **do**
    $\hat{\boldsymbol{L}}_j^{\prime t}[a] \leftarrow \boldsymbol{L}^t[j, a] + \sum_{j' \in C_{j,a}} -\psi_{j'}^{*t}(-\hat{\boldsymbol{L}}_{j'}^{\prime t})$.
    $\hat{\boldsymbol{x}}_j^{t+1} \leftarrow \nabla \psi_j^{*t}(-\hat{\boldsymbol{L}}_j^{\prime t})$.
  end for
  Construct $\boldsymbol{x}^{t+1}$.
end for

Algorithm 4 Reparameterized FD-FTRL

for iteration $t = 1$ to T **do**
  $\boldsymbol{l}^t \leftarrow \text{ObserveLoss}(\boldsymbol{x}^t)$.
  for node $j \in \mathcal{J}$ in bottom-up order **do**
    $\hat{\boldsymbol{L}}_j^{\prime t}[a] \leftarrow \boldsymbol{L}^t[j, a] + \sum_{j' \in C_{j,a}} \alpha_{j'}^t$.
    $\hat{\boldsymbol{R}}_j^{\prime t} \leftarrow \alpha_j^t \mathbf{1} - \hat{\boldsymbol{L}}_j^{\prime t}$.
    $\hat{\boldsymbol{x}}_j^{t+1} \leftarrow [\hat{\boldsymbol{R}}_j^{\prime t}]^+ / \|[\hat{\boldsymbol{R}}_j^{\prime t}]^+\|_1$.
  end for
  Construct $\boldsymbol{x}^{t+1}$.
end for

$\|[\alpha_j^t \mathbf{1} - \hat{\boldsymbol{L}}_j^{\prime t}]^+\|_2^2 = \lambda_j^t$.

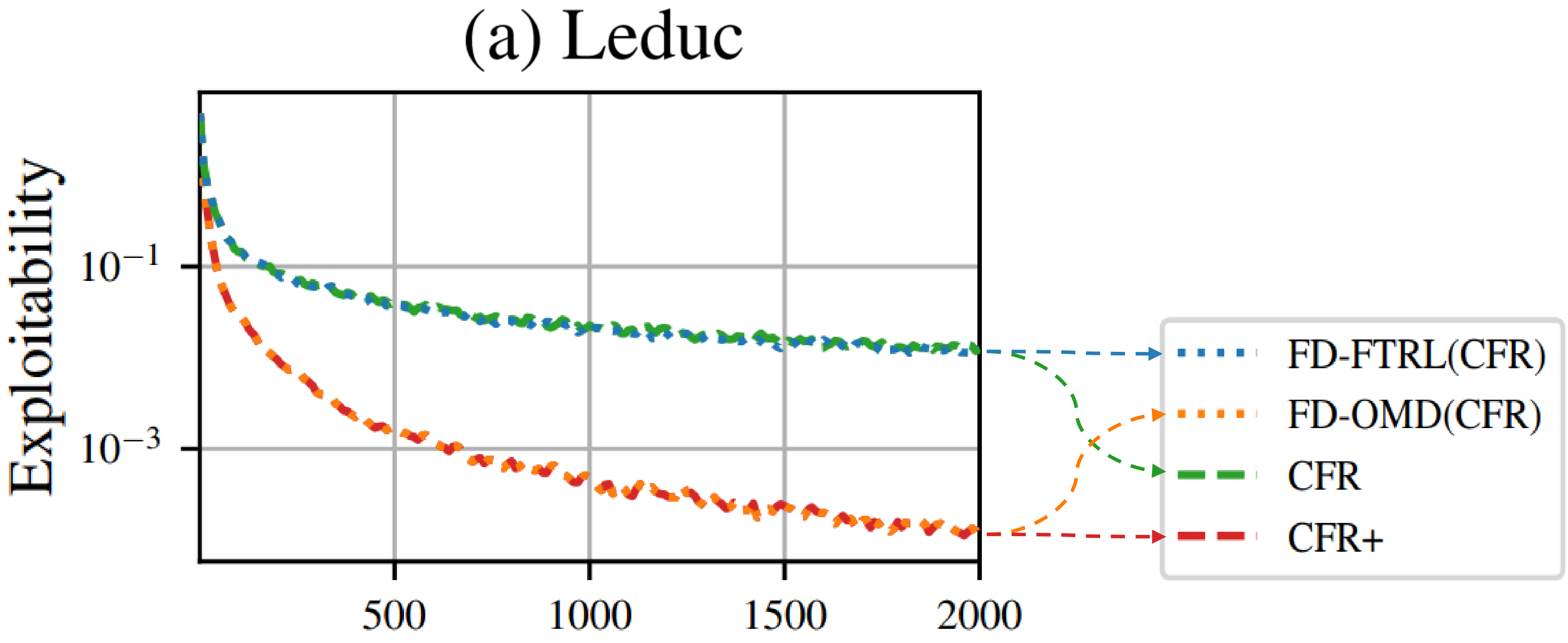
Application：Convergence

- Let λ_j^t be lower bounded by the cumulative counterfactual regret
- FD-FTRL (FD-OMD) can converge at a rate of $O(1/\sqrt{T})$

Corollary 3.9. *If* $\eta \sum_{k=1}^{t} \|\hat{\boldsymbol{r}}_j^k\|_2^2 \leq \lambda_j^t$ *and* $\lambda_j^{t-1} \leq \lambda_j^t$, $\forall j \in \mathcal{J}$, $t > 0$, *then, the total regret of FD-FTRL(R) (FD-OMD(R)) after* T *iterations is* $R^T \leq \sum_{j \in \mathcal{J}} \left(\sqrt{n_j} + \frac{1}{\eta} \right) \sqrt{\lambda_j^T}$.

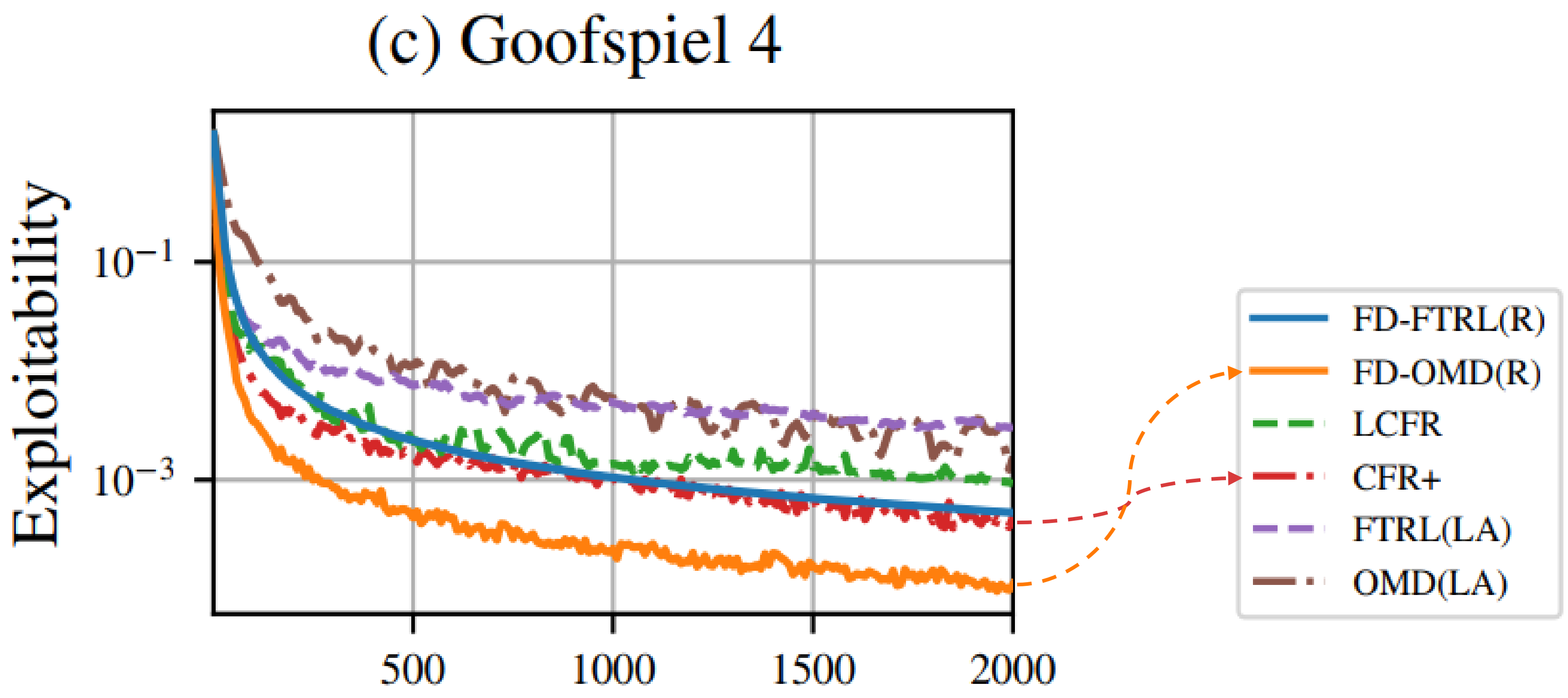
Results：Empirical equivalences

- FD-FTRL (FD-OMD) can recover CFR (CFR+)



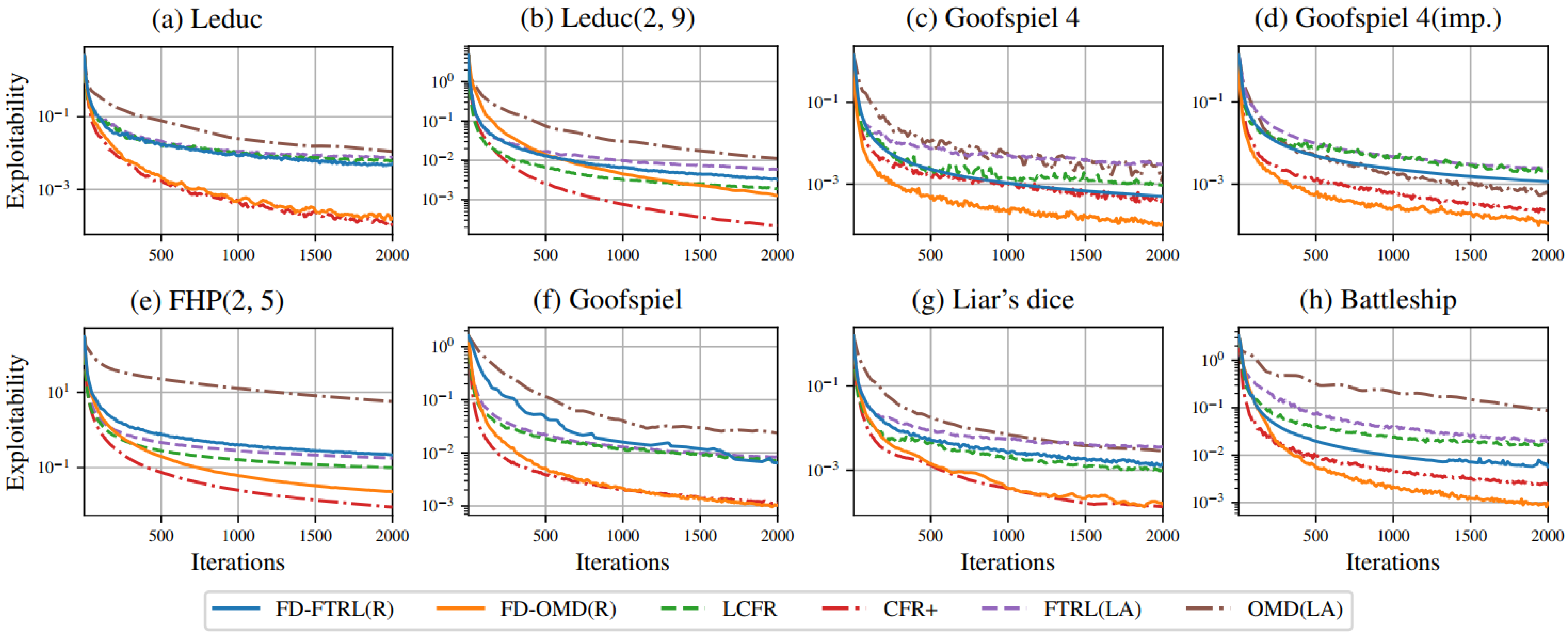
Results：Adaptive FD-OMD

- FD-OMD (Orange) is faster than CFR+ (Red)



Results： Adaptive FD-OMD

- FD-OMD (Orange) is no worse than CFR+ (Red) in 6/8 games.



Conclusions

- CFR-RM (CFR-RM+) is equivalent to a special FTRL (OMD)
  - Partially explain the superior performance of CFRs
- Propose FD-FTRL and FD-OMD
  - Faster than the non-adaptive FTRL and OMD in EFGs
  - Contrary to previous findings, OMD can be faster than CFR+

Thank you!

Find more details in the paper:

Liu, W., Jiang, H., Li, B., and Li, H. Equivalence Analysis between Counterfactual Regret Minimization and Online Mirror Descent. In *International Conference on Machine Learning*, 2022.