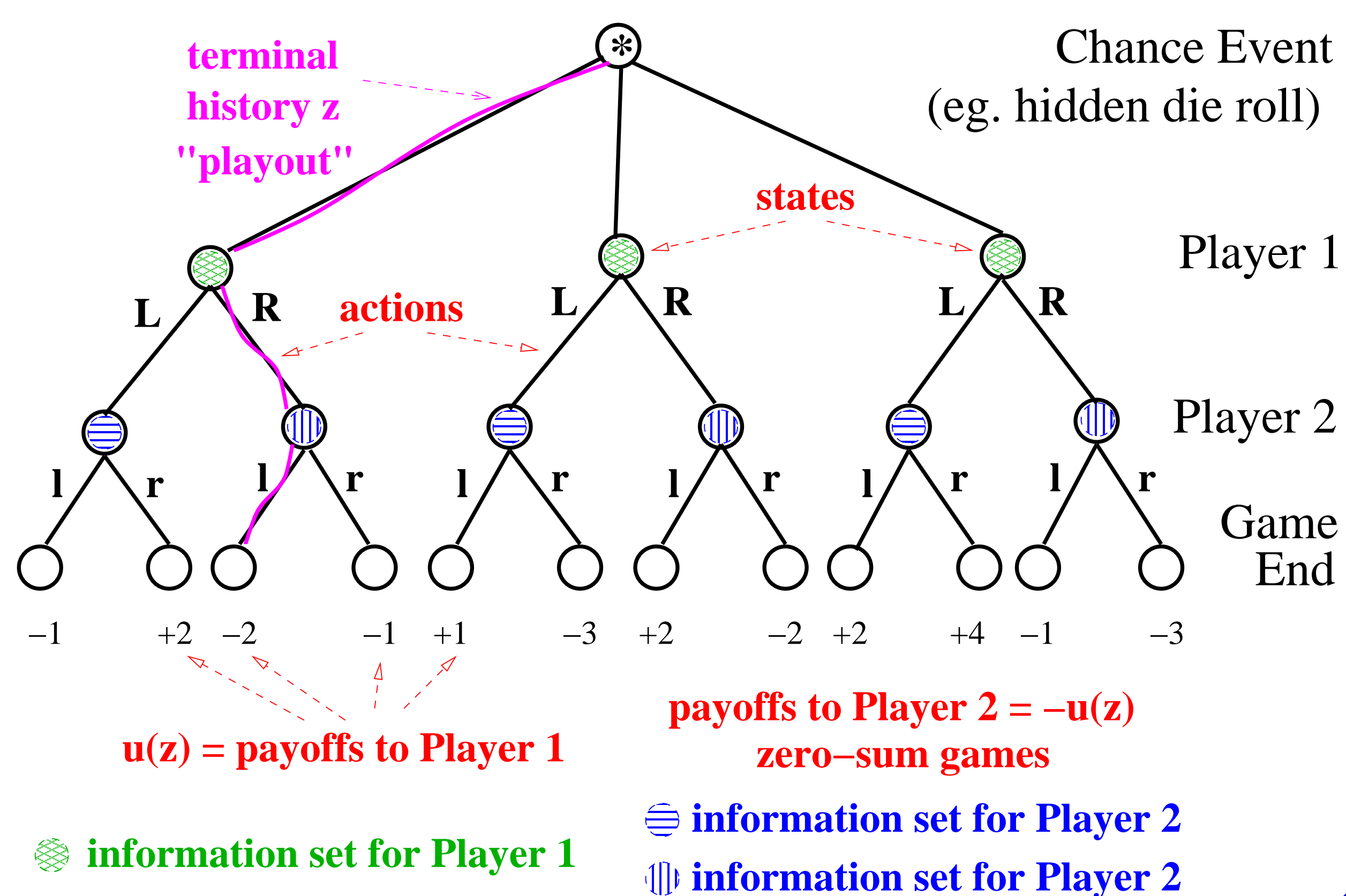


Problem

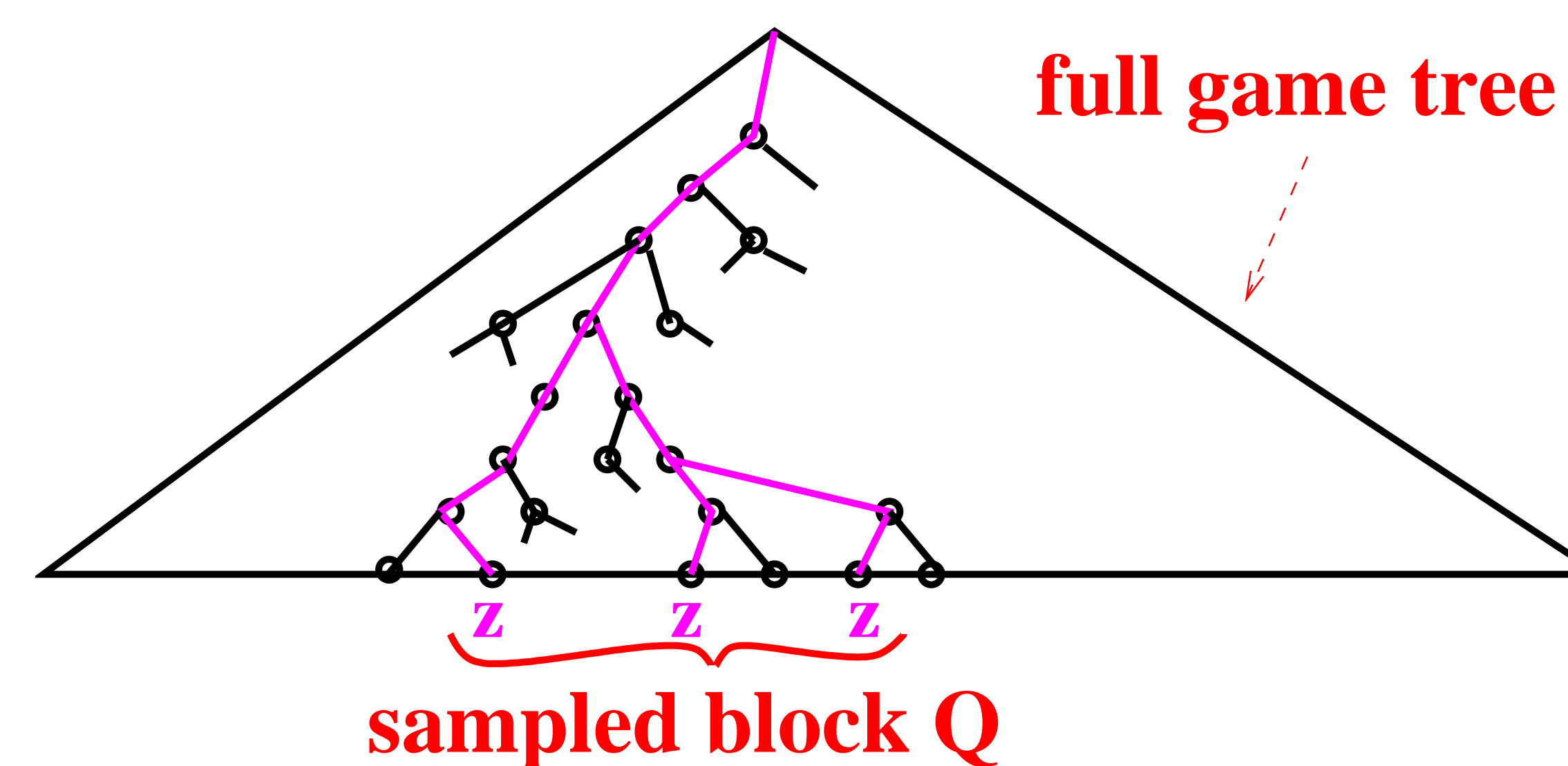
Goal Solve large, zero-sum, imperfect info. games

Model Extensive-form games (EFGs)



Monte Carlo CFR

1. Sample a block of terminal histories $Q \subseteq Z$.
2. Update regrets only on sampled histories.



$$\tilde{v}(\sigma, I) = \sum_{z \in Q \cap Z_I} \frac{\pi_{-i}^\sigma(z[I]) u_i(z) \pi^\sigma(z[I], z)}{\Pr(z)}$$

$$E[\tilde{v}(\sigma, I)] = v(\sigma, I) \Rightarrow \text{regret updates match CFR}$$

External Sampling

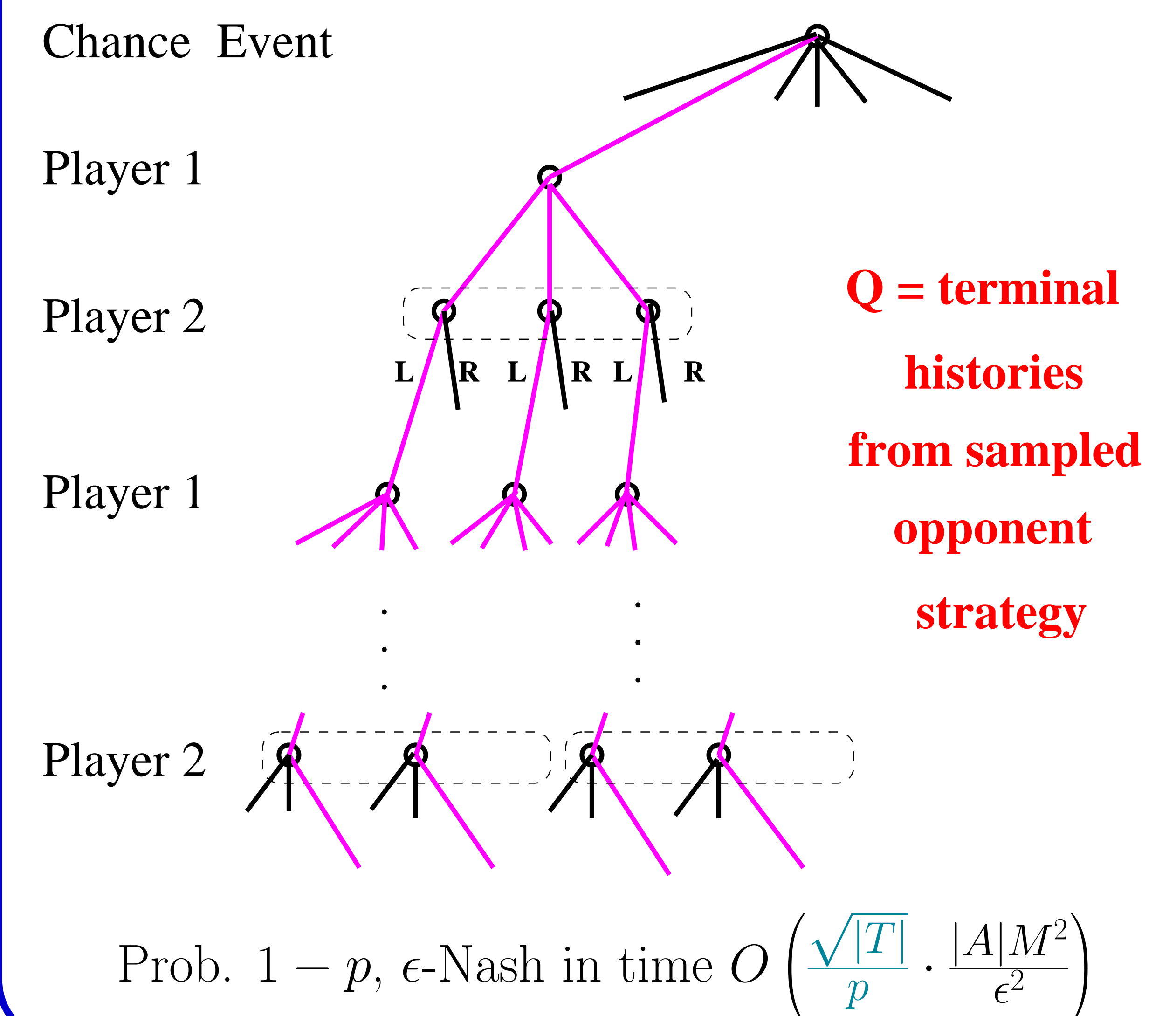
Chance Event

Player 1

Player 2

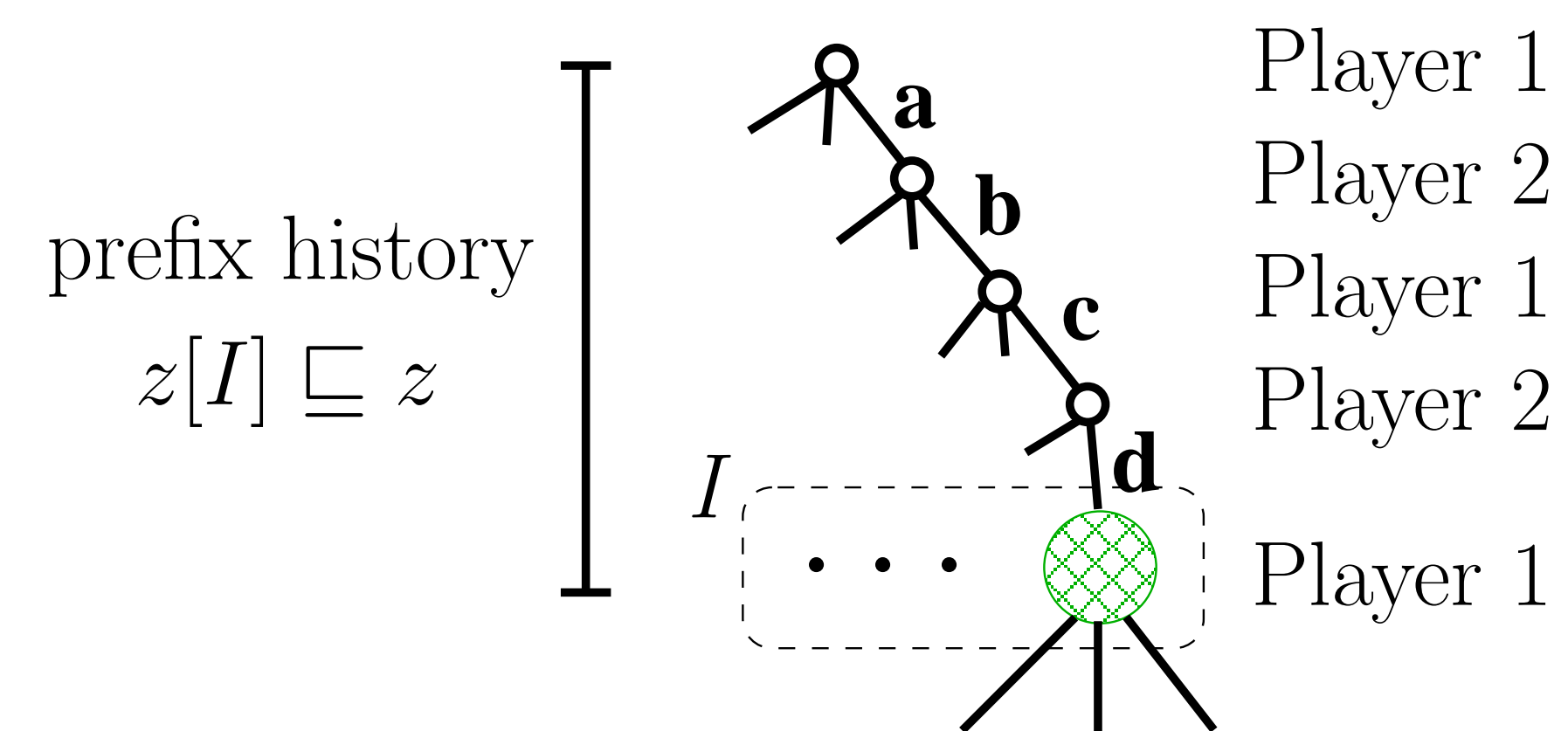
Player 1

Player 2



Counterfactual Regret Minimization

— Zinkevich et. al. 2007



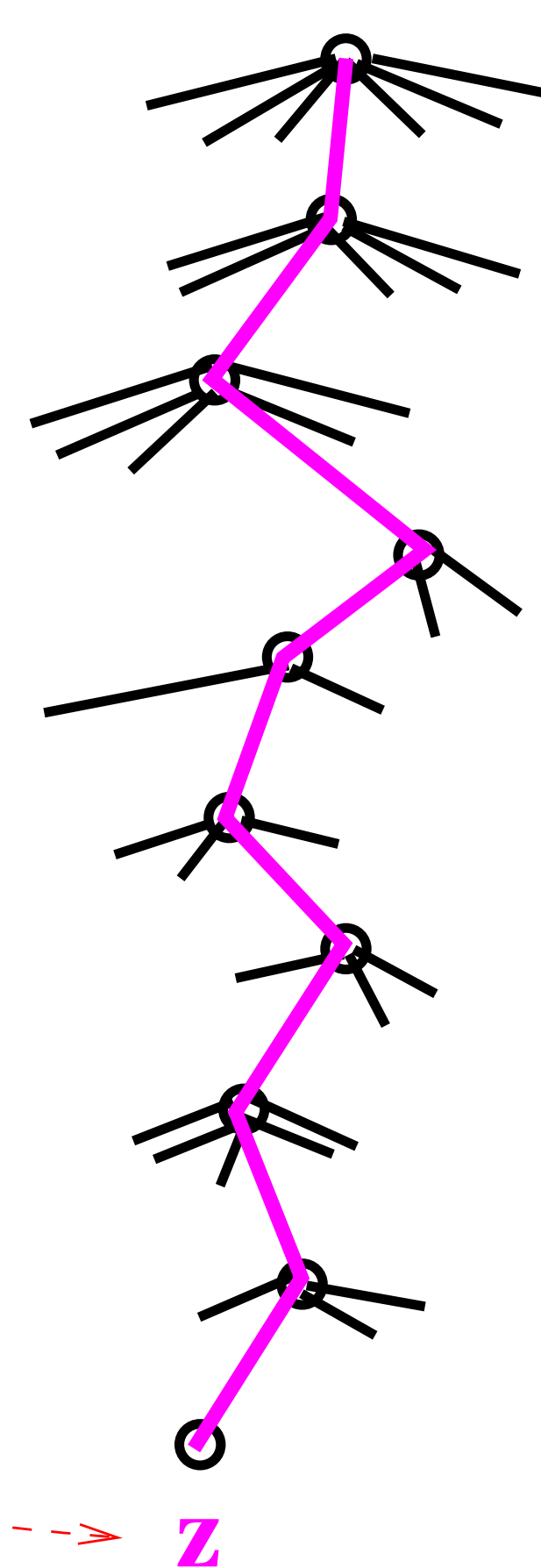
$$\pi_{-i}^\sigma(z[I]) = \sigma_{-i}(b) \cdot \sigma_{-i}(d) = \text{counterfactual weight}$$

1. Minimize weighted swap regret for (I, a) given σ .
2. Update σ using regret-matching.
 - \Rightarrow Minimizes overall regret.
 - $\Rightarrow \bar{\sigma} \rightarrow \epsilon$ -Nash in time $O(|T||A|M^2/\epsilon^2)$.

Iterations require a full tree traversal!

Outcome Sampling

$Q =$
a single
terminal
history



Sampling profile σ'

$$\Pr(z) = \pi^{\sigma'}(z) \geq \delta$$

Prob. $1 - p$, ϵ -Nash in time

$$O\left(\frac{T_{\text{height}}}{p\delta^2} \cdot \frac{|A|M^2}{\epsilon^2}\right)$$

Empirical Results

