GAME THEORY

László Gulyás
Associate Professor, ELTE, AI Department
✉ lgulyas@inf.elte.hu

Tamás Takács
PhD student, ELTE, AI Department
✉ tamastheactual@inf.elte.hu
🌐 tamastheactual.github.io

Game Theory

Lecture 8

Correlated Equilibrium and Learning in Games

1 Correlated Equilibrium

2 QR Equilibrium

3 Learning Dynamics

Bayesian Nash Equilibrium

Strategy for player i: $\sigma_i : \Theta_i \to \Delta(A_i)$ (map types to action distributions)

Definition: Strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a **Bayesian Nash equilibrium (BNE)** if for all i and all $\theta_i \in \Theta_i$:

$$\sigma_i(\theta_i) \in \arg\max_{a_i \in A_i} E_{\theta_{-i} \sim p(\cdot|\theta_i)}[u_i(a_i, \sigma_{-i}(\theta_{-i}), \theta_i, \theta_{-i})]$$

Intuition: Each type of each player best responds given beliefs about others' types

Special case: When all types are public ($|\Theta_i| = 1\ for\ all\ i$),BNE = NE

Example 1: First-Price Auction

Setup:

- n bidders for single item
- Each bidder i has private valuation $v_i \sim U[0,1]$ (independent)
- Bidders submit sealed bids $b_i \in [0, \infty)$
- Highest bidder wins, pays their bid
- Ties broken uniformly at random

Payoffs:

$$u_i(b_1, \dots, b_n, v_i) = \begin{cases} v_i - b_i & \text{if } b_i > \max_{j \neq i} b_j \\ 0 & \text{otherwise} \end{cases}$$

Question: What is the Bayesian Nash equilibrium?

The Reverse Game Theory Problem

Classical game theory:

- **Given**: Game (players, actions, payoffs)
- **Find**: Equilibria and outcomes

Mechanism design:

- **Given:** Desired outcomes (efficiency, fairness, revenue)
- **Design:** Game rules (who moves when, what they know, how payoffs determined)
- **Find:** Mechanisms where equilibrium achieves desired outcome

"Reverse engineering": Choose rules so that rational play implements social goals

Vickrey-Clarke-Groves Mechanism

Setup: Social choice among alternatives X based on private valuations

Mechanism:

1. Each player i reports valuation function $\hat{v}_i: X \to R$
2. Choose outcome $x^* = \arg\max_{x \in X} \sum_i \hat{v}_i(x)$ (maximize reported total value)
3. Player i pays **Clarke pivot tax**:

$$t_i = \sum_{j \neq i} \hat{v}_j(x^*) - \max_{x \in X} \sum_{j \neq i} \hat{v}_j(x)$$

Intuition: Player pays the **externality** they impose on others

Second-Price vs First-Price

Second Price:

- Equilibrium Strategy: $b^*(v_i) = v_i (truthful)$
- Revenue: $E[\text{2nd highest } v]$
- Properties: DSIC, simple

First Price:

- Equilibrium Strategy: $b^*(v_i) = \frac{n-1}{n} v_i (shading)$
- Revenue: $\frac{n-1}{n+1}$
- Properties: BIC, strategic

Revenue equivalence theorem: Both auctions yield same expected revenue (in symmetric IPV model)!

Shading in first-price exactly compensates for difference in payment rule

Game Theory

Lecture 8

Correlated Equilibrium and Learning in Games

Beyond Independent Randomization

Nash Equilibrium limitation:
Players choose strategies **independently,** cannot coordinate even when beneficial

Example: Traffic Light:
- Two players approach intersection
- Payoffs: both **Go** → accident $(0,0)$
- Both **Stop** → deadlock $(1,1)$
- One **Go** one **Stop** → efficient $(3,2)$ or $(2,3)$

Nash equilibria:
- Both Stop (inefficient)
- **Mixed NE**: each Go with prob $1/2$ → expected welfare 1.5

Correlated solution:
- **Traffic light** signals: "Row Go" or "Column Go" with equal probability
- Expected welfare: 2.5 (much better!)

Correlated Equilibrium

A **correlated equilibrium (CE)** is a distribution $\mu \in \Delta(A)$ such that for every player i and every pair of actions $a_i, a_i' \in A_i$:

$$\sum_{a_{-i}} \mu(a_i, a_{-i})[u_i(a_i, a_{-i}) - u_i(a_i', a_{-i})] \geq 0$$

Notation:

- $\mu(a_i, a_{-i}) =$ probability mediator recommends profile (a_i, a_{-i})
- Sum over a_{-i} gives expectation conditional on i receiving a_i

CE: Mediator Interpretation

Interpretation (mediator form):

1. Mediator samples $a = (a_1, \dots, a_n)$ from μ
2. Privately recommends a_i to player i
3. Each player finds it optimal to **obey** recommendation

Obedience constraint:

- Conditional on receiving recommendation a_i, player i's expected utility from playing a_i is at least as good as deviating to any other action a_i'
- Mediator must be **trusted** and recommendations **private**

Coarse Correlated Equilibrium (CCE)

A **coarse correlated equilibrium (CCE)** is a distribution $\mu \in \Delta(A)$ such that for every player i and every action $a_i' \in A_i$:

$$\sum_{a \in A} \mu(a) u_i(a) \geq \sum_{a \in A} \mu(a) u_i(a_i', a_{-i})$$

Interpretation:

- Players observe μ **before choosing actions**
- No private recommendations
- Can only commit to **fixed deviation** before seeing signal

Difference from CE:

- CCE: deviation chosen **ex-ante** (before signal)
- CE: deviation can depend on **recommendation received** (ex-post)

Weaker concept: Every CE is a CCE, but not vice versa

Inclusion Relations

Theorem: For any game,

$$\mathrm{NE} \subseteq \mathrm{CE} \subseteq \mathrm{CCE}$$

Both inclusions can be strict!

Proof: NE is subset of CE

Proof of NE $\subseteq$ CE:

- Let σ be a Nash equilibrium
- Define $\mu(a) = \prod_i \sigma_i(a_i)$ (product distribution)
- For any i, a_i, a_i':

$$\sum_{a_{-i}} \mu(a_i, a_{-i}) u_i(a_i, a_{-i}) = \sum_{a_{-i}} \sigma_i(a_i) \prod_{j \neq i} \sigma_j(a_j) u_i(a_i, a_{-i}) = \sigma_i(a_i) u_i(\sigma_i, \sigma_{-i})$$

- Since σ is NE, $u_i(\sigma_i, \sigma_{-i}) \geq u_i(a_i', \sigma_{-i})$ for all a_i'
- Multiplying by $\sigma_i(a_i) > 0$ gives CE constraint

Proof: CE is subset of CCE

Proof of CE $\subseteq$ CCE:

- CE constraints are **stronger** (conditional on recommendation)
- CCE constraints average over all recommendations
- Summing CE constraints weighted by $\mu(a_i)$ yields CCE

Intuition:

- CE: Can adapt deviation to recommendation received (ex-post)
- CCE: Must commit to deviation before seeing signal (ex-ante)

Example: Simple Coordination

	L	R
U	$(1,1)$	$(0,0)$
D	$(0,0)$	$(1,1)$

Nash equilibria:
- (U,L) pure NE
- (D,R) pure NE
- Mixed NE: each plays $1/2 - 1/2$ $\rightarrow$ expected payoff $(1/2, 1/2)$

Correlated equilibrium:
- $\mu(U,L) = 1/2, \mu(D,R) = 1/2$
- Obedience constraints:
  - Row gets U: expected payoff 1 (vs 0 from D)
  - Row gets D: expected payoff 1 (vs 0 from U)
  - Column symmetric
- Expected payoff: $(1,1)$ (much better than mixed NE!)

This CE is NOT a product distribution (not NE)

Example: Battle of the Sexes

	B	F
B	$(2, 1)$	$(0, 0)$
F	$(0, 0)$	$(1, 2)$

Nash equilibria:
- (B, B) pure NE → payoff
- pure NE → payoff $(1,2)$
  - Mixed NE → payoff $(2/3, 2/3)$

Correlated equilibrium (fair):
- $\mu(B, B) = \lambda, \mu(F, F) = 1 - \lambda$ for any $\lambda \in [0,1]$
- Setting $\lambda = 1/2$ gives expected payoff $(3/2, 3/2)$
- **Pareto dominates** mixed NE!
- Solves **coordination + fairness** simultaneously

Example: Battle of the Sexes (Cont.)

	B	F
B	$(2, 1)$	$(0, 0)$
F	$(0, 0)$	$(1, 2)$

Obedience constraints:

- Row gets B: payoff 2 (vs 0 from F)
- Row gets F: payoff 1 (vs 0 from B)
- Column symmetric

Example: Matching Pennies (Zero-Sum)

	H	T
H	$(1,-1)$	$(-1,1)$
T	$(-1,1)$	$(1,-1)$

Nash equilibrium:

- Unique mixed NE: each plays $1/2 - 1/2$
- Value: 0 for both players

Correlated equilibrium:

- Any μ satisfying CE constraints
- But in zero-sum: **CE cannot improve over NE value!**

Example: Matching Pennies (Zero-Sum) (Cont.)

	H	T
H	$(1, -1)$	$(-1, 1)$
T	$(-1, 1)$	$(1, -1)$

Theorem (zero-sum invariance):

- In two-player zero-sum games, all CE (and CCE) yield the same **value** as the unique mixed Nash equilibrium
- Correlation changes distribution but not value

Why?

- Player 1's gain $=$ Player 2's loss
- No common interest to exploit
- Correlation only "redistributes" across outcomes, not total value

Example: Prisoner's Dilemma (Dominance)

	C	D
C	$(3,3)$	$(0,4)$
D	$(4,0)$	$(1,1)$

Unique Nash equilibrium: $(D, D) \to$ payoff $(1,1)$
Correlated equilibrium:

- Can \mu put weight on (C, C) to achieve $(3,3)$?
- **No!** D strictly dominates C for both players
- CE constraints require:

$$\sum_{a_{-i}} \mu(C, a_{-i}), [u_i(C, a_{-i}) - u_i(D, a_{-i})] \geq 0$$

- But $u_i(C, a_{-i}) < u_i(D, a_{-i})$ for all a_{-i}(dominance!)
- So $\mu(C, a_{-i}) = 0$ for all a_{-i}

CE cannot overcome **dominance** (credibility issue)

Example: Chicken with Catastrophe Avoidance

	Swerve	Straight
Swerve	$(0, 0)$	$(-1, 1)$
Straight	$(1, -1)$	$(-10, -10)$

Nash equilibria:

- (S, St) and (St, S) pure NE
- Mixed NE with positive prob on (St, St) → catastrophe!

Correlated equilibrium (avoid catastrophe):

- $\mu(S, St) = 1/2,\ \mu(St, S) = 1/2,\ \mu(St, St) = 0$
- Obedience constraints satisfied (same as BoS structure)
- **Expected payoff:** $(0, 0)$ vs mixed NE with catastrophe risk

CE via Linear Programming

Formulation:

$$\sum_{a \in A} \mu(a) = 1, \quad \mu(a) \geq 0 \quad \forall a s.t. \sum_{a_{-i}} \mu(a_i, a_{-i})[u_i(a_i, a_{-i}) - u_i(a_i', a_{-i})]$$

$$\geq 0 \quad \forall i, a_i, a_i' \ max_\mu \sum_{a \in A} \mu(a) W(a) \quad \{(or\ any\ linear\ objective)\}$$

Variables: $\mu(a)$ for each action profile $a \ \in A$

Constraints: Obedience for each player, action pair

CE: Computational Properties

Number of constraints:

- n players, $|A_i|$ actions each
- $n \cdot |A_i| \cdot (|A_i| - 1)$ obedience constraints
- Polynomial in game representation

Complexity comparison:

CE via LP: Polynomial time
NE: PPAD-complete (hard!)
CCE: Even fewer constraints (ex-ante deviations only)

Advantage: Can optimize social welfare subject to incentive compatibility

CCE via Linear Programming

Formulation:

$$\text{s.t.} \quad \sum_{a\in A} \mu(a) u_i(a) \geq \sum_{a\in A} \mu(a) u_i(a_i', a_{-i}) \quad \forall i, a_i' \max_{\mu} \sum_{a\in A} \mu(a) W(a)$$

$$\mu(a) \geq 0 \quad \forall a \sum_{a\in A} \mu(a) = 1$$

Number of constraints:

- $n \cdot |A_i|$ constraints (one per player per deviation)
- **Fewer constraints than CE!**
- Even easier to compute

CCE via Linear Programming (Cont.)

CCE polytope:

- CCE set forms a convex polytope
- NE are vertices of NE polytope
- CE polytope contained in CCE polytope

Geometric interpretation:

- NE → vertices of product simplex intersection
- CE → polytope in probability simplex
- CCE → larger polytope

CE Polytope Geometry

Example: 2×2 game

	L	R
U	(a, e)	(b, f)
D	(c, g)	(d, h)

Variables: $x_{UL}, x_{UR}, x_{DL}, x_{DR}$ (probabilities)

Plus: $x_{UL} + x_{UR} + x_{DL} + x_{DR} = 1$, all ≥ 0

Solution space: Polytope in 4D probability simplex

CE Constraints for 2×2 Game

CE constraints for row player:

$$\begin{aligned}\text{Rec } U\text{:}\quad & x_{UL}(a-c) + x_{UR}(b-d) \geq 0,\\ \text{Rec } D\text{:}\quad & x_{DL}(c-a) + x_{DR}(d-b) \geq 0.\end{aligned}$$

CE constraints for column player:

$$\begin{aligned}\text{Rec } L\text{:}\quad & x_{UL}(e-f) + x_{DL}(g-h) \geq 0,\\ \text{Rec } R\text{:}\quad & x_{UR}(f-e) + x_{DR}(h-g) \geq 0.\end{aligned}$$

When Does Correlation Help?

Correlation improves outcomes when:

- **Multiple equilibria:** Coordination on "better" equilibrium (BoS, traffic light)
- **Asymmetric equilibria:** Fair randomization between asymmetric NE
- **Common interest:** Exploit shared incentives to avoid bad outcomes (Chicken)

Correlation does NOT help when:

- **Dominant strategies:** Dominance eliminates correlation gains (Prisoner's Dilemma)
- **Zero-sum games:** Minimax value unchanged by correlation (Matching Pennies)
- **Pure conflict:** No common interest to exploit

Implementing Correlation Devices

Methods:

1. **Public randomness:**
   - Dice, coin flips, market signals
   - All observe same realization
   - Cheap talk protocols (if credible)

2. **Trusted mediator:**
   - Samples from μ and gives **private recommendations**
   - Requires trust and privacy
   - Examples: referees, arbitrators, platform algorithms

3. **Cryptographic protocols:**
   - Secure multi-party computation
   - Commit-and-reveal schemes
   - No trusted third party needed

Game Theory

Lecture 8

Correlated Equilibrium and Learning in Games

1 Correlated Equilibrium

2 QR Equilibrium

3 Learning Dynamics

Why QRE? (Bounded Rationality Meets Equilibrium)

Empirical observation:

- Lab experiments often deviate from Nash predictions
- But play is **payoff-sensitive** (not random)
- Higher-payoff actions chosen more often (but not exclusively)

Nash limitation:

- Assumes perfect rationality: play **only** best responses
- Real players make mistakes, explore, have computational constraints

Quantal Response Equilibrium (McKelvey & Palfrey):

- Replace hard best response with **smooth**, stochastic choice rule
- Better actions chosen with **higher probability** (not certainty)
- Fixed points of smooth responses are **equilibria**

Logit Response: Foundation

Random-utility model (Gumbel noise):

- Player i perceives utility: $u_i(a_i; \sigma_{-i}) + \varepsilon_{a_i}$
- Noise $\varepsilon_{a_i} \sim$ type-I extreme value (i.i.d.), scaled by $1/\lambda$
- Player maximizes perceived utility

Choice probabilities are logit:

$$\sigma_i(a_i) = \frac{\exp\{\lambda \cdot u_i(a_i; \sigma_{-i})\}}{\sum_{a_i' \in A_i} \exp\{\lambda \cdot u_i(a_i'; \sigma_{-i})\}}$$

Intuition: Higher-payoff actions get higher probability, not certainty

Logit Response: Rationality Parameter

Parameter $\lambda \geq 0$ ("precision" or "rationality"):

- $\lambda = 0$: Uniform randomization (complete noise)
- $\lambda \to \infty$: Mass on **payoff maximizers** (approaching Nash)
- Intermediate λ: Smooth tradeoff between exploration and exploitation

Interpretation:

- Low λ: High noise, nearly uniform play
- High λ: Low noise, nearly best-respond

Behavioral meaning: λ measures cognitive precision or rationality level

Logit QRE: Definition

A **logit QRE** at precision λ is a profile $\sigma = (\sigma_i)_{i \in N}$ such that:

$$\sigma_i(a_i) = \frac{\exp\{\lambda \cdot u_i(a_i; \sigma_{-i})\}}{\sum_{a_i' \in A_i} \exp\{\lambda \cdot u_i(a_i'; \sigma_{-i})\}} \quad \forall i, a_i$$

Interpretation: σ is a **fixed point** of smooth best-response mapping SBR_λ

Each player chooses stochastically, with probabilities proportional to exponential of expected payoffs

QRE: Existence and Properties

Existence:

- SBR_λ is continuous on compact convex domain
- **Brouwer fixed-point theorem** $\Rightarrow$ QRE exists for every $\lambda \geq 0$

Connection to Nash:

- As $\lambda \to \infty$, QRE accumulation points are **Nash equilibria**
- QRE provides **refinement path** from uniform to Nash

Uniqueness:

- Not guaranteed in general (multiple QRE possible for same λ)
- Often unique for small/large λ

Example: Battle of the Sexes QRE Path

Game:

	B	F
B	$(2, 1)$	$(0, 0)$
F	$(0, 0)$	$(1, 2)$

QRE equations:

- Row plays B with probability p, Column plays B with probability q
- Row's logit:

$$p = \frac{1}{1 + \exp\{-\lambda[(2-0)q + (0-1)(1-q)]\}} = \frac{1}{1 + \exp\{-\lambda(3q-1)\}}$$

- Column's logit:

$$q = \frac{1}{1 + \exp\{-\lambda[(1-0)p + (0-2)(1-p)]\}} = \frac{1}{1 + \exp\{-\lambda(3p-2)\}}$$

Example: Battle of the Sexes QRE Path (Cont.)

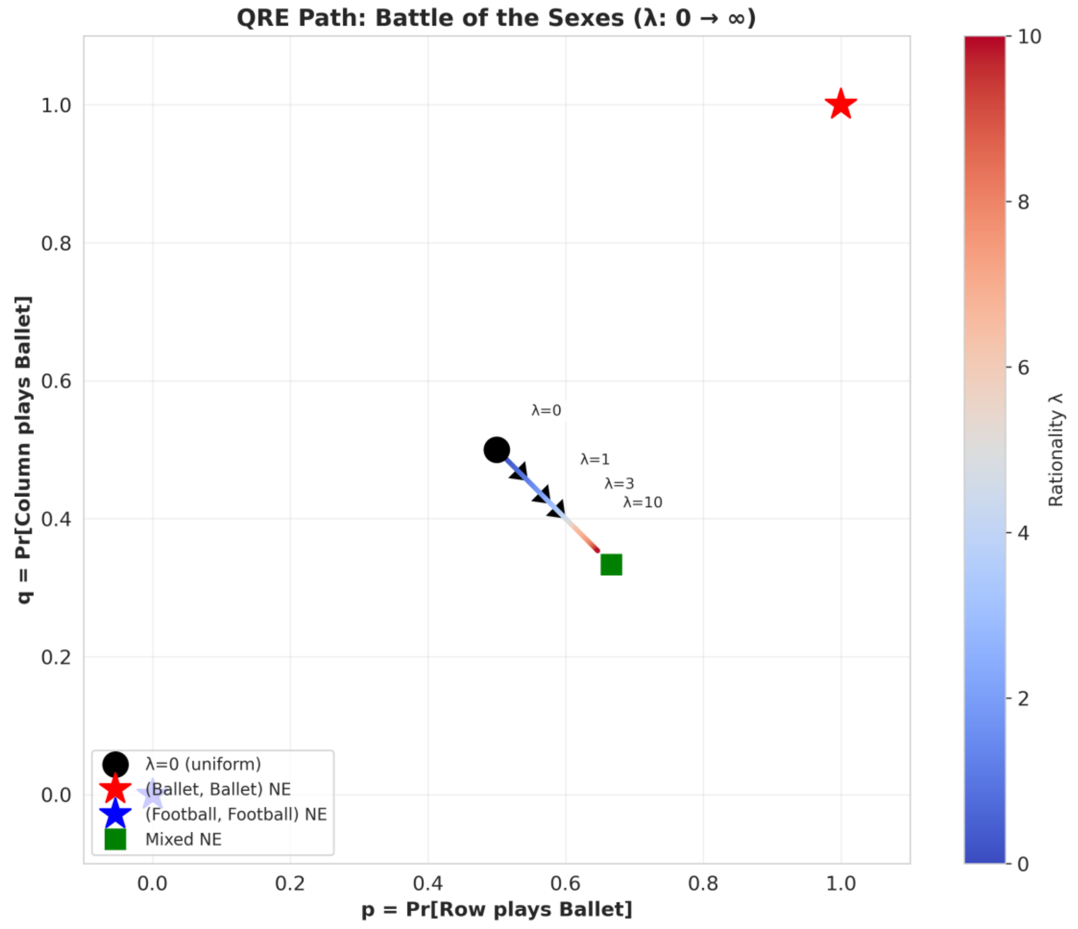
Game:

	B	F
B	$(2, 1)$	$(0, 0)$
F	$(0, 0)$	$(1, 2)$

QRE path behavior:

- At $\lambda = 0$: $(p, q) = (0.5, 0.5)$ (uniform play)
- As $\lambda \to \infty$: Path **bifurcates** toward pure NE $(1,1)$ or $(0,0)$
- Symmetric intermediate values along path

QRE Path in BoS



Interpretation:

- **Horizontal axis:** λ (rationality parameter)
- **Vertical axis:** probability of playing B
- **Solid curve:** QRE path from uniform to NE
- **Bifurcation point:** Where path splits toward two pure NE
- **Mixed NE:** Unstable point on path (not attractive)

QRE Interpretation

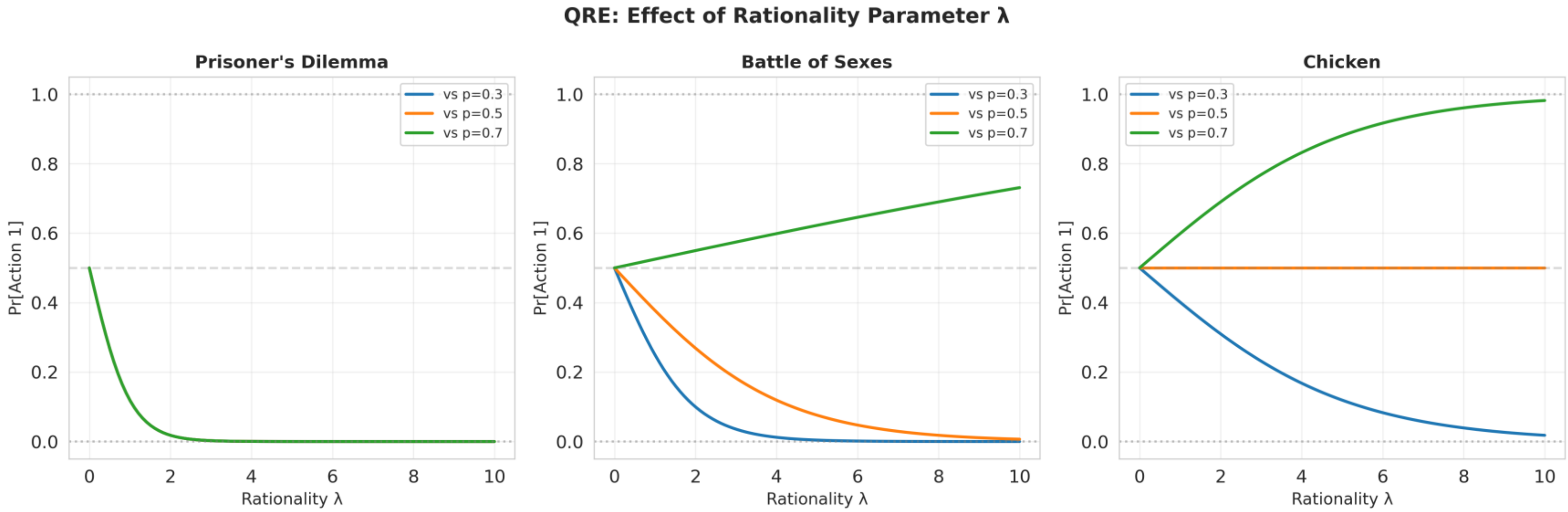
Descriptive power:
- QRE fits experimental data **better than NE** in many settings

Prescriptive use:
- Models bounded rationality, computational constraints, exploration
- Bounded rationality as explicit feature (not "mistake")

Predictive use:
- For intermediate λ, QRE predicts **which actions are more likely**
- Gives **probabilities** over all actions

Normative caution:
- QRE is descriptive, not normative
- Does not tell players what to do

QRE Interpretation



Explanation:
Low λ: Nearly uniform (high noise)
Medium λ: Significant payoff-sensitivity but not extreme
High λ: Concentrates on best responses (Nash)

Practical values:

- Lab experiments: $\lambda \in [0.5,5]$ typical
- Context-dependent (game complexity, stakes, time pressure)

QRE Applications

Experimental economics:

- Fitting lab data with rationality parameter λ
- Estimating λ from observed play
- Predicting out-of-sample behavior

Behavioral game theory:

- Modeling mistakes, trembles, noise in equilibrium
- Alternative to perfect rationality assumption
- Bounded rationality as first-class concept

QRE in Design and Learning

Mechanism design:

- Robust mechanisms that work under QRE play (not just NE)
- Worst-case guarantees with bounded rationality
- Example: Auctions with noisy bidders

Learning dynamics:

- QRE connects to **logit learning** dynamics
- Online learning with entropy regularization
- Smoothness parameter relates to learning rate

QRE is single-parameter model; richer models exist (level-k, cognitive hierarchy)

Game Theory

Lecture 8

Correlated Equilibrium and Learning in Games

1 Correlated Equilibrium

2 QR Equilibrium

3 Learning Dynamics

Why Learning Dynamics?

Static equilibrium concepts (NE, CE, QRE):

- Assume players already know the game and opponents' strategies
- Coordination on equilibrium: how does it emerge?

Learning models:

- How players **adapt** over time through repeated play
- Discovering equilibria via experience
- Dynamic adjustment processes

Classic Learning Dynamics

Three classic dynamics:

1. **Fictitious Play (FP):** Best-respond to empirical frequency
2. **Replicator Dynamics (RD):** Evolutionary selection (above-average payoffs grow)
3. **No-regret learning:** Minimize regret over time

Connection to equilibria:

- FP converges in some games (zero-sum, potential)
- RD stationary points include NE
- No-regret → CCE; no swap-regret → CE

External Regret and CCE

External regret (for player i after T rounds):

$$R_T^{\text{ext}}(a_i') = \sum_{t=1}^{T} [u_i(a_i', a_{-i}^t) - u_i(a_i^t, a_{-i}^t)]$$

Interpretation:

- Regret from not always playing a_i' instead of actual actions a_i^t
- Measures "I wish I had played a_i' every round"

No external regret:

$$\frac{1}{T} \max_{a_i' \in A_i} R_T^{\text{ext}}(a_i') \to 0 \quad \text{as } T \to \infty$$

External Regret and CCE (Cont.)

Theorem (Hannan, 1957):

If all players have no external regret, then empirical distribution

$$\widehat{\mu_T} = \frac{1}{T}\sum_{t=1}^{T} \delta_{a^t}$$

converges to the **CCE** set.

Proof sketch:

- Average external regret inequalities: $\frac{1}{T} R_T^{\text{ext}}(a_i') \leq 0$ for all a_i'
- These are exactly **CCE constraints** on $\widehat{\mu_T}$!

Swap Regret and CE

Swap regret (for player i and deviation map $\phi_i: A_i \to A_i$):

$$R_T^{\text{swap}}(\phi_i) = \sum_{t=1}^{T} [u_i(\phi_i(a_i^t), a_{-i}^t) - u_i(a_i^t, a_{-i}^t)]$$

Interpretation:

- Regret from not swapping each action a_i^t to $\phi_i(a_i^t)$
- "I wish I had used swap function ϕ_i"
- No swap regret:

$$\frac{1}{T} \max_{\phi_i} R_T^{\text{swap}}(\phi_i) \to 0 \quad \text{as } T \to \infty$$

Theorem (Foster & Vohra, 1997):

If all players have no swap regret, then $\widehat{\mu_T}$ converges to **CE** set.

Hedge Algorithm (Multiplicative Weights)

Goal: Minimize external regret → converge to CCE
Algorithm for player i with actions $A_i = \{1, \dots, m\}$:

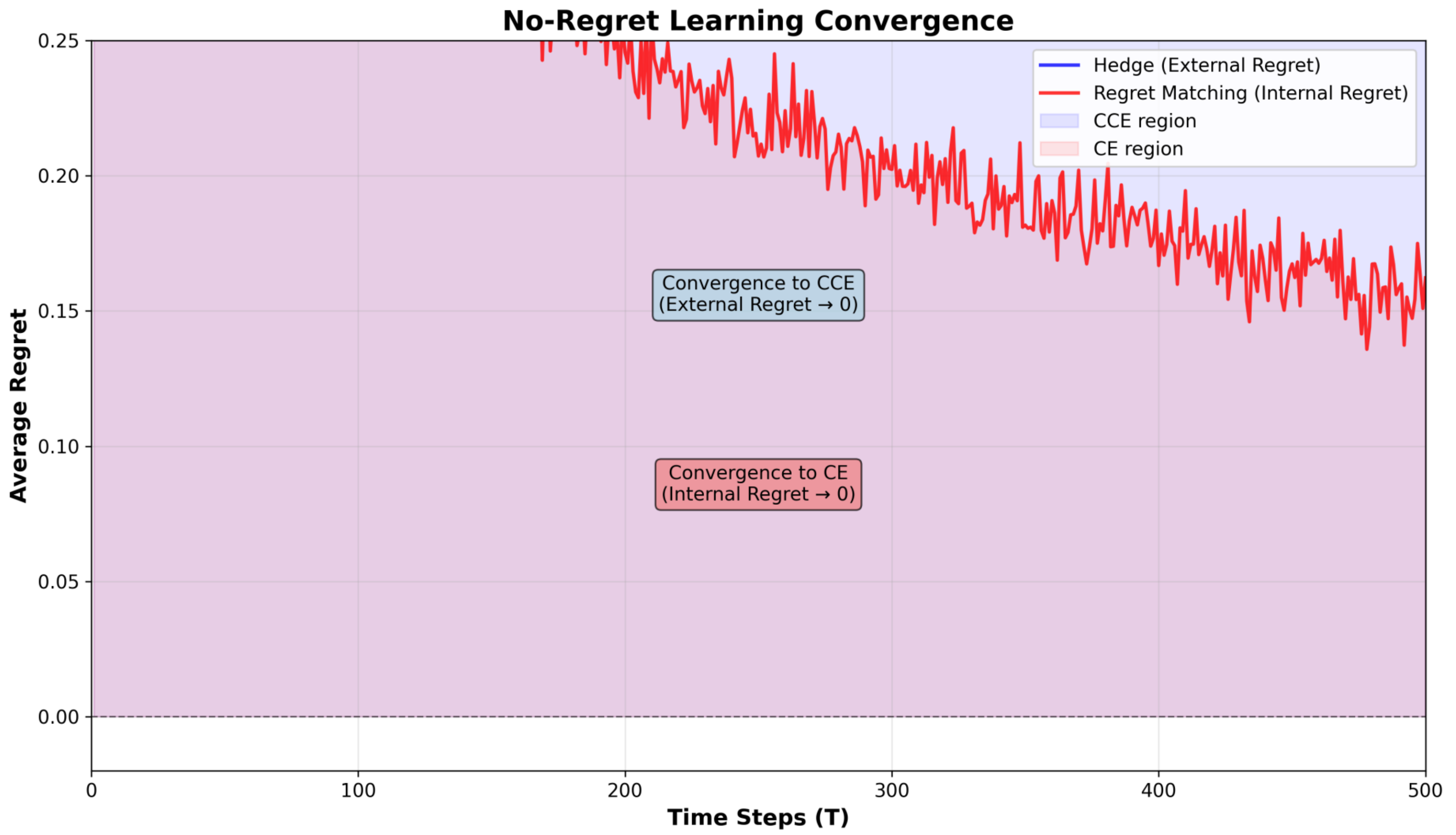
1. **Initialize**: $w_1(a) = 1$ for all $a \in A_i$
2. **At round t:**
   - Play a_i^t sampled with probability $p_t(a) \propto w_t(a)$
   - Observe utility $u_i(a, a_{-i}^t)$ for each action $a \in A_i$
   - **Update**: $w_{t+1}(a) = w_t(a) \cdot \exp\left(\eta \cdot u_i(a, a_{-i}^t)\right)$

Learning rate: $\eta = \sqrt{\frac{2 \log m}{T}}$ (if T known) or adaptive
Guarantee:

$$R_T^{\text{ext}} \leq \sqrt{2T \log m}$$

Average regret: $\frac{1}{T} R_T^{\text{ext}} = O\left(\frac{1}{\sqrt{T}}\right) \to 0$
Result: If all players run Hedge, empirical play converges to CCE

Regret Convergence



Regret Convergence (Cont.)

Plot shows:

- Horizontal axis: Number of rounds T
- Vertical axis: Average external regret $\frac{1}{T} R_T^{\text{ext}}$
- **Curve**: $O\left(1/\sqrt{T}\right)$ decay
- **Shaded region**: Confidence interval

Observation:

- Regret decreases as $1/\sqrt{T}$
- Convergence guaranteed for any game
- Practical: works well even for moderate T (hundreds of rounds)

Regret Matching (Hart & Mas-Colell, 2000)

1. **Track regrets:** $R_t(a \to a') =$ cumulative regret for not switching from a to a' when playing a

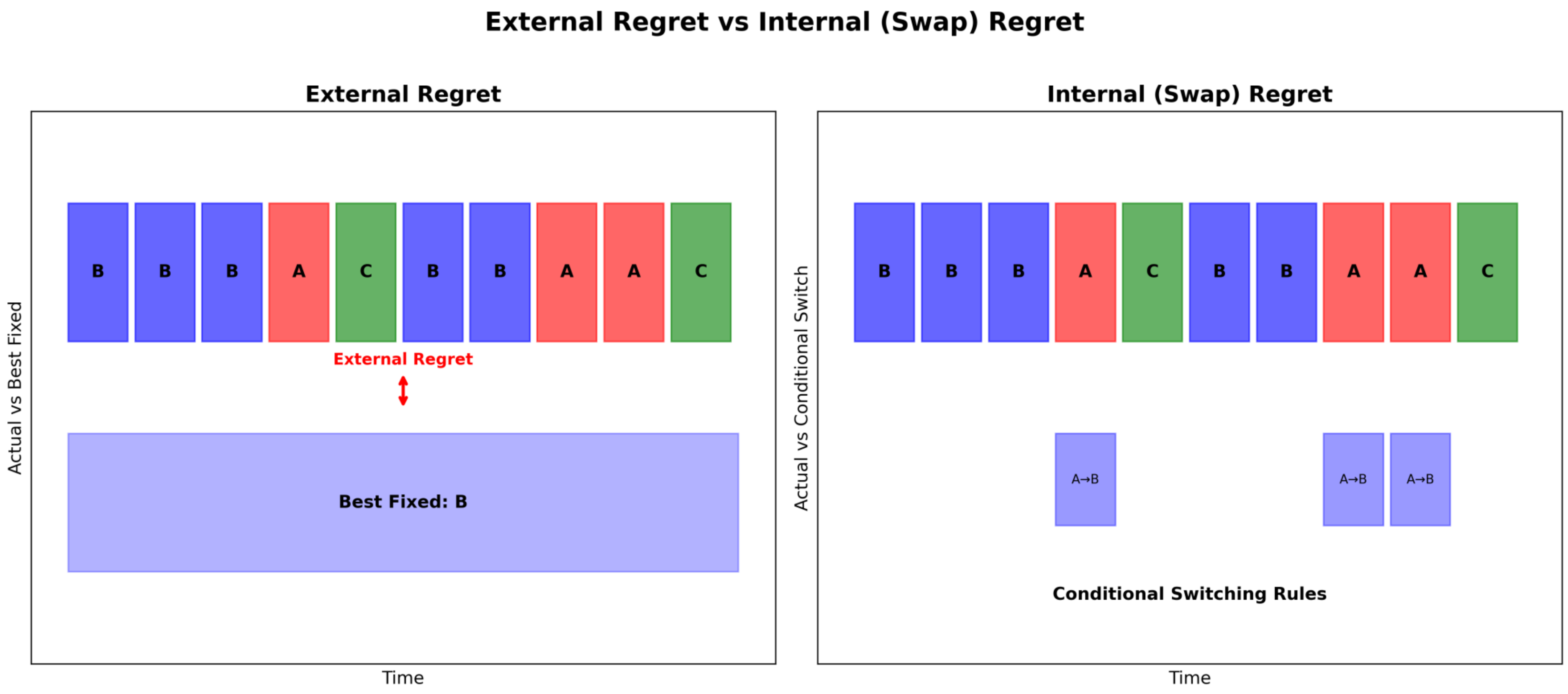
2. **At round t:**
   - For each action $a' \in A_i$, compute:
     $R_t^+(a') = \sum_{a \in A_i} [R_t(a \to a')]_+$
     where $[x]_+ = \max(0, x)$
   - Play a' with probability $\propto R_t^+(a')$(uniform if all zero)

3. Update regrets after observing utilities

Guarantee:

- Swap regret vanishes: $\frac{1}{T} \max_{\phi_i} R_T^{\text{swap}}(\phi_i) \to 0$
- Empirical play approaches **CE**

Regret Types Comparison



Regret Types Comparison (Cont.)

Diagram shows:

- **External regret:** Fixed action a' counterfactual
- **Swap regret:** Swap function ϕ counterfactual (stronger)
- **Inclusions:** No swap regret $\subseteq$ No external regret

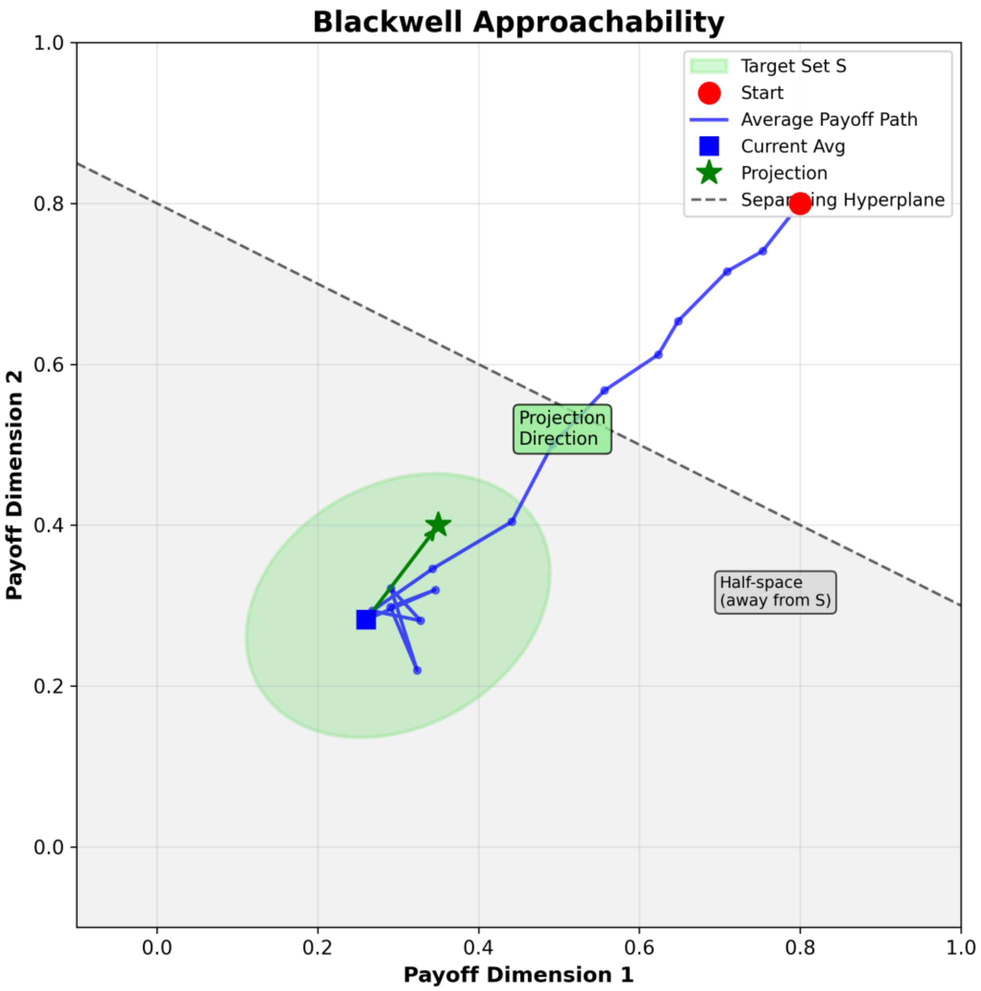
Blackwell Approachability

Blackwell's theorem (1956):
A convex set $S \subseteq R^d$ is **approachable** if player can ensure average vector payoff converges to S regardless of opponent's play

Connection to CE:
- Define **regret vector**: components = regrets for each possible swap
- **Target set**: Non-positive orthant (all regrets ≤ 0)
- Approachability ⇔ achieving no swap regret

Blackwell's algorithm:
1. Compute current average payoff vector $\overline{v_t}$
2. Project $\overline{v_t}$ to nearest point in $S \rightarrow getv^*$
3. Play action that moves expected vector toward v^*
4. Repeat → converge to S

Blackwell Approachability



Blackwell Approachability (Cont.)

Diagram shows:

Target set S: Convex set in 2D (regret space)
Current average $\overline{v_t}$: Outside S
Projection v^*: Closest point in S
Halfspace: Region to play toward (moves toward v^*)

Algorithm:

- Play mixed action whose expectation lies in the halfspace separating $\bar{v}_t from S$
- Guarantees squared distance to S decreases in expectation

Equilibrium Inefficiency

Social welfare: $W(a) = \sum_{i \in N} u_i(a)$

Definitions:

$\text{OPT} = \max_{a \in A} W(a)$ — socially optimal welfare
$textEQ$ — welfare at (worst-case) equilibrium

Price of Anarchy (PoA):

$$\text{PoA} = \frac{\text{OPT}}{\text{EQ}}$$

(For utilities; for costs, use inverse: $\text{PoA} = \frac{\text{EQ}}{\text{OPT}}$)

PoA: Interpretation

Interpretation:

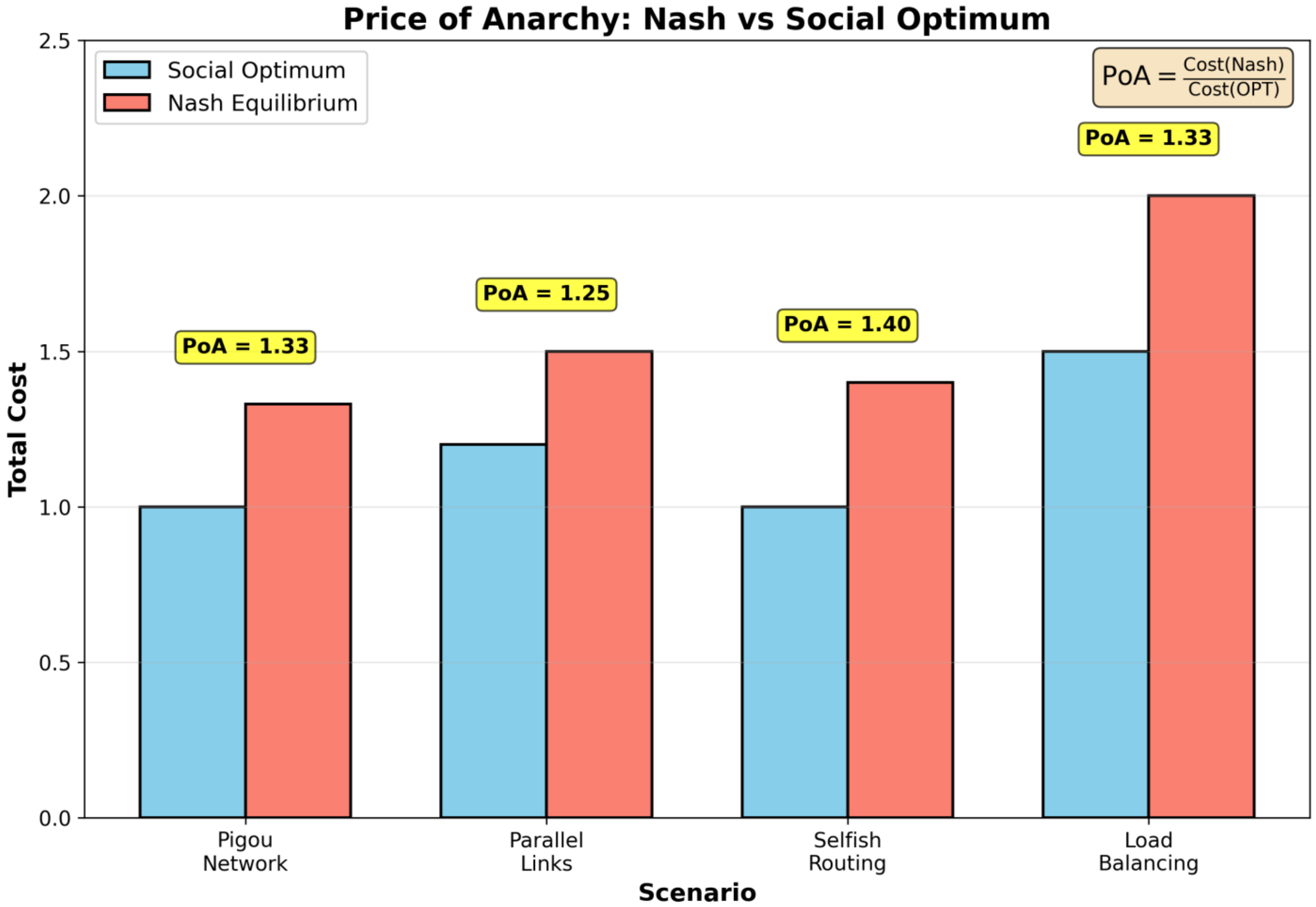
- $\text{PoA} = 1 \rightarrow$ equilibrium is socially optimal
- $\text{PoA} = 2 \rightarrow$ equilibrium achieves half of optimal welfare
- Measures **worst-case inefficiency** of decentralized play

Note: Basic PoA definition introduced in previous lecture. Here we extend to correlated equilibria and develop the smoothness framework.

Robust definition:

- Can define PoA for **NE, CE, or CCE**
- Smoothness framework gives **unified bounds**

Price of Anarchy Examples



Smoothness Framework (Roughgarden, 2009)

Definition:

A game is $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ −**smooth** if for any profiles a, a^*:

$$\sum_{i \in N} u_i(a_i^*, a_{-i}) \geq \lambda \cdot W(a^*) - \mu \cdot W(a)$$

Interpretation:

- Consider **unilateral deviations** from any profile $\boldsymbol{a}$ to optimal profile a^*
- If each player i deviates to a_i^*, total utility gain bounds welfare gap
- λ measures "improvement potential," μ measures "penalty"

Smoothness: Main Theorem

Theorem:

If a game is $(\lambda, \mu) -$ smooth, then:

$$\mathrm{PoA} \leq \frac{1+\mu}{\lambda}$$

Key properties:

1. Bound holds for **NE, CE, and CCE** simultaneously
2. Robust to **learning dynamics** (no-regret players)
3. **Compositional:** Smoothness preserved in concurrent games
4. Does not require **computing equilibrium**!

Example: Pigou Network

Setup: Two parallel routes, unit flow

- **Route 1:** Variable latency $\ell_1(x) = x$
- **Route 2:** Constant latency $\ell_2(x) = 1$

Nash equilibrium (Wardrop):

- Users split flow until latencies equalize
- With flow $r = 1$: All on Route 2 $\rightarrow$ cost 1

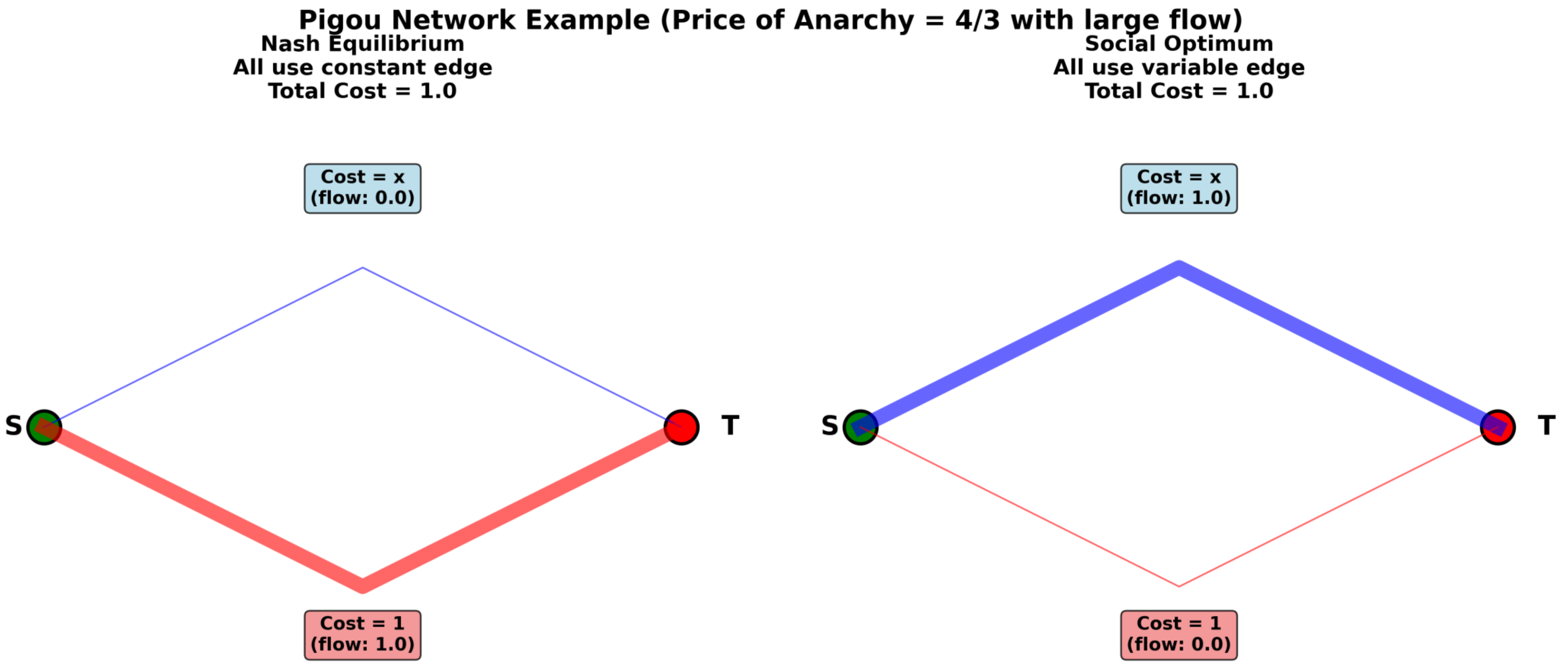
Social optimum:

- Minimize total latency: $\int_0^{x_1} x \, , dx + x_2 \cdot 1$
- Optimal: all on Route 1 $\rightarrow$ cost 1

PoA for $r = 1$: PoA = 1 (same!)

But with $r > 1$: PoA grows (can be > 1)

Pigou Network



Pigou Network (Cont.)

Diagram shows:

- Two parallel edges from source s to target t
- **Top edge:** Latency $\ell(x) = x$(congestion-sensitive)
- **Bottom edge:** Latency $\ell(x) = 1$(constant)
- **Flow:** Mass r to route

Nash flow:

- Equilibrium splits flow to equalize latencies
- With affine latencies: $\ell_1(x_1) = \ell_2(x_2)$

Pigou Network with General Flow

With flow mass $r > 1$:

Nash equilibrium:

- Users split until $x_1 = 1$ (latency 1 on both routes)
- Remaining flow $r - 1$ on Route 2
- **Total cost:** $\int_0^1 x\,,dx + (r-1) \cdot 1 = 0.5 + r - 1 = r - 0.5$

Social optimum:

- Minimize: $\int_0^{x_1} x\,,dx + x_2 \cdot 1 = \frac{x_1^2}{2} + x_2$
- Subject to: $x_1 + x_2 = r$
- Optimal: $x_1 = 1, x_2 = r - 1$ (same as NE for this specific game!)

Actually PoA = 1 for Pigou!

General affine latencies: PoA $\leq 4/3$ (tight bound, achieved by other networks)

Pigou Network with General Flow

With flow mass $r > 1$:

Nash equilibrium:

- Users split until $x_1 = 1$ (latency 1 on both routes)
- Remaining flow $r - 1$ on Route 2
- **Total cost:** $\int_0^1 x \, , dx + (r-1) \cdot 1 = 0.5 + r - 1 = r - 0.5$

Social optimum:

- Minimize: $\int_0^{x_1} x \, , dx + x_2 \cdot 1 = \frac{x_1^2}{2} + x_2$
- Subject to: $x_1 + x_2 = r$
- Optimal: $x_1 = 1, x_2 = r - 1$ (same as NE for this specific game!)

Actually PoA $=$ 1 for Pigou!

General affine latencies: PoA $\leq 4/3$ (tight bound, achieved by other networks)

Congestion Games and Smoothness

Congestion game model:

- Players choose **resources** (routes, servers, etc.)
- Latency/cost $\ell_e(x_e)$ depends on congestion x_e
- Common in: traffic networks, resource allocation, load balancing

Affine latencies: $\ell_e(x) = a_e x + b_e$

Smoothness result (Roughgarden & Tardos, 2002):

- Congestion games with affine latencies are $(1,1)$ −smooth
- **PoA bound**: $\mathrm{PoA} \leq \frac{1+1}{1} = 2$ for costs

Robustness:

- Same bound holds for **NE, CE, CCE**
- Applies under **learning dynamics** (no-regret play)
- Does **not require** computing equilibrium!

Learning Convergence and Advantages

Learning convergence:

- No-regret algorithms (Hedge, Regret Matching) converge in **polynomial rounds**
- Explicit bounds: $O(\sqrt{T})$ average regret
- Practical: works well even for moderate T (hundreds of rounds)

Advantage of CE/CCE:

- Computationally much easier than NE (polynomial vs PPAD-complete)
- Suitable for large-scale games, online learning settings
- Learning dynamics **automatically** converge (no equilibrium computation needed!)

Limitations of Equilibrium Concepts

Multiplicity:

- Many equilibria may coexist (which one?)
- Selection problem: focal points, conventions, history
- CE: even more equilibria (entire polytope!)

Non-credibility:

- Off-equilibrium threats may not be credible
- Refinements needed: SPE, PBE (extensive form)
- CE mediator must be **trusted**

Limitations of Equilibrium Concepts

Information requirements:
- Common knowledge assumptions strong
- Bounded rationality not modeled in NE/CE
- QRE addresses this but adds parameter λ

Dynamics and convergence:
- Some learning dynamics cycle (e.g., fictitious play in general games)
- Convergence not guaranteed in all games
- Speed of convergence matters for practical use

Coordination on equilibrium:
- How do players coordinate on specific equilibrium?
- Communication, cheap talk, pre-play negotiation
- CE: mediator solves coordination but requires trust

Exercise 1

For Battle of the Sexes with mediator strategy $\mu(B, B) = \lambda, \mu(F, F) = 1 - \lambda$, verify all four CE obedience constraints explicitly for any $\lambda \in [0,1]$.

Exercise 2

Prove that in any two-player zero-sum game, all CE yield the same **value** as the unique mixed Nash equilibrium value.

Exercise 3

Implement Hedge algorithm for Matching Pennies with learning rate $\eta = \sqrt{\frac{2\log 2}{T}}$:

- Run for $T = 1000$ rounds against a fixed opponent strategy (e.g., 60% Heads).
- Plot cumulative regret and average regret over time. Verify $O(1/\sqrt{T})$ convergence.

Exercise 4

For Battle of the Sexes, solve the QRE equations numerically for $\lambda \in \{0.5, 1, 2, 5, 10\}$:

- Plot the QRE path in (p, q) space.
- Identify the bifurcation point where path splits toward pure NE.

Exercise 5

For a 2-route congestion game with latencies $\ell_1(x) = x$ and $\ell_2(x) = 1$, and flow mass $r = 2$:

1. Compute Wardrop equilibrium (Nash flow)
2. Compute social optimum
3. Calculate PoA
4. Verify the $(1,1)$ −smoothness inequality

Summary

- **Correlated Equilibrium (CE):**
  - Mediator recommends actions, players obey
- **Quantal Response Equilibrium (QRE):**
  - Bounded rationality: smooth best-response with noise
- **Learning Dynamics:**
  - External regret → CCE convergence
  - Swap regret → CE convergence
  - No-regret algorithms: Hedge (CCE), Regret Matching (CE)
- **Price of Anarchy (PoA):**
  - Measures worst-case equilibrium inefficiency
  - Applications: congestion games, network routing

Course Textbooks

- Bonanno, G. (2024). ***Game Theory (3rd ed.)***. University of California, Davis. Received from: GT Book
- Axelrod, R. (1984). ***The Evolution of Cooperation***. Basic Books. Received from: Axelrod Article
- Nisan, N., Roughgarden, T., Tardos, É., & Vazirani, V. V. (2007). ***Algorithmic Game Theory***. Cambridge University Press. Received from: AGT Book
- Myerson, R. B. (1991). ***Game Theory: Analysis of Conflict.*** Harvard University Press. Received from: GT Book 2F.
- Christianos et al., ***Multi-Agent Reinforcement Learning: Foundations and Modern Approaches***, 2023. Received from: MARL Book.pdf
- Shoham, Y., & Leyton-Brown, K. (2008). ***Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations***. Cambridge University Press. Received from: MARL Book.pdf
- `***nashpy***` documentation (readthedocs). Link: NashPy Docs

That's All for Today!