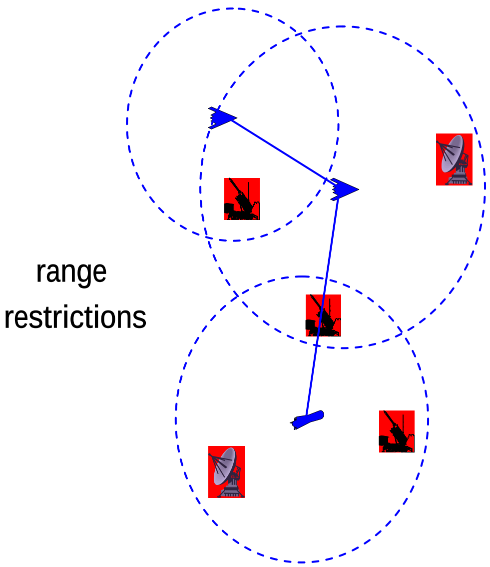
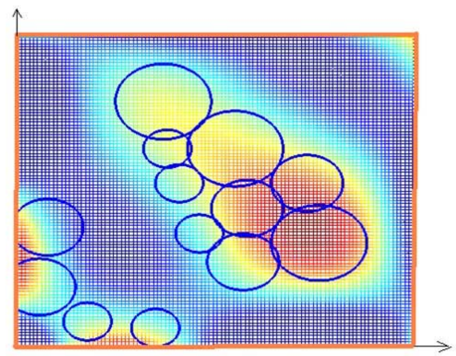
Forecasting in Game Theoretic Learning

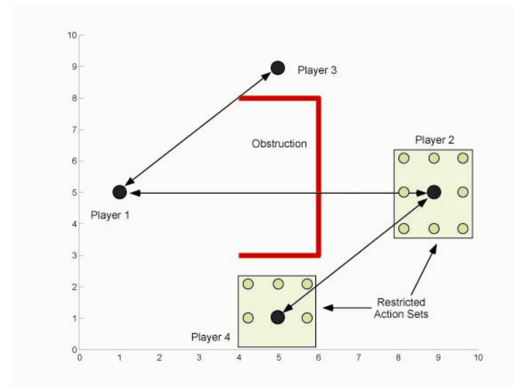
Jeff Shamma
Electrical and Computer Engineering
Georgia Institute of Technology

Joint work with S. Mannor & G. Arslan

Networked Control Systems: Game Theoretic Perspective



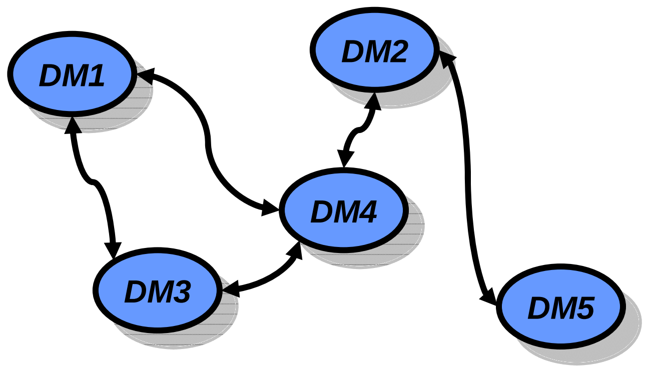




- Examples:
  - Vehicle target assignment
  - Mobile sensor allocation
  - Vehicle rendezvous

- Desirable features:
  - Distributed information & computation
  - Capability of dynamic reconfiguration
  - Circumvention of closed form characterizations
  - Adaptation to actual environment

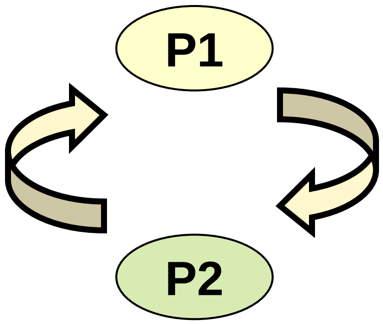
Learning in Games



- Setup (repeated matrix games):
  - Multiple decision makers
  - Evolving strategies
  - *Restricted information*
- Focus: Dynamics ***away*** from equilibrium
- Extensive prior work, e.g.:
  - *Theory of Learning in Games*, Fudenberg & Levine, 1998
  - *Individual Strategy & Social Structure,* Young, 1998
  - *Strategic Learning and Its Limits,* Young, 2004
  - *Population Games and Evolutionary Dynamics*, Sandholm, forthcoming
- Key Challenges:
  - Learning/adaptation in an environment of other learners
  - Descriptive vs prescriptive vs ***hybrid*** agenda

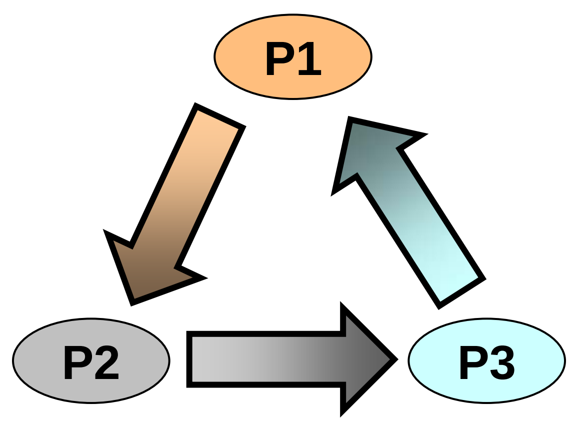
Shapley "Fashion" Game

- 2 players, each with 3 moves: {*Red*, *Green*, *Blue*}
  - Player1: Fashion leader wants to differ from Player2
  - Player2: Fashion follower wants to copy Player1
- Key assumption: Players do ***not*** announce preferences
- Daily routine:
  - Play game
  - Observe actions
  - Update strategies



Jordan "Anti-coordination" Game

- 3 players, each with 2 moves: {*Left*, *Right*}
  - Player1 wants to differ from Player2
  - Player2 wants to differ from Player3
  - Player3 wants to differ from Player1
- Players do ***not*** announce preferences
- Daily routine:
  - Play game
  - Observe actions
  - Update strategies



Best Response to Myopic Forecast

- Constraint: $p_i(k) = F(\text{information up to time } k)$

- Opponent action measurements:
  - Forecast opponent strategy
  - Play best response to forecast
  - Observe opponent actions
  - Revise forecast & repeat

Forecasting

- Finite number of possible outcomes
- Repeated in time
- Objective: Predict *probability* of outcome
- Performance measurements:
  - Model based consistency: Classes of sources
  - Universal consistency: All sources

Example: Smooth Fictitious Play

- Forecast: Empirical frequencies of opponent

$$q_{-i}(k+1) = q_{-i}(k) + \frac{1}{k+1}(a_{-i}(k) - q_{-i}(k))$$

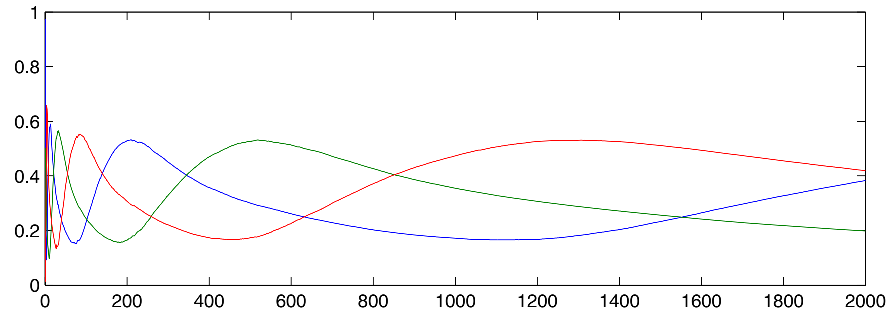
- Play: Smooth best response to forecast

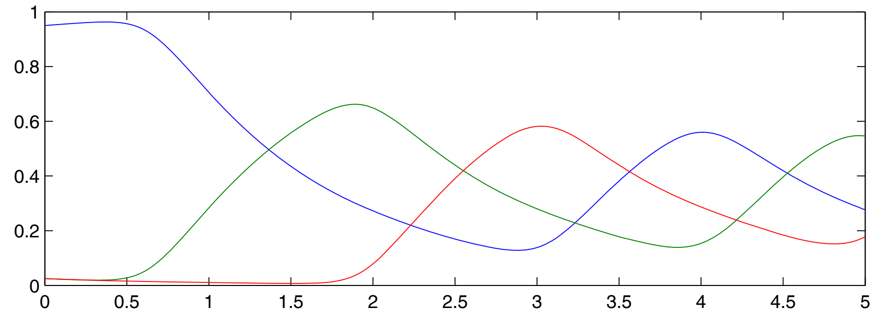
$$a_i(k) = \mathrm{rand}[\beta(q_{-i}(k))]$$
$$\beta(v) = \arg\max_{s\in\Delta} s^T M v + \tau\mathcal{H}(s)$$

- Presumption: Stationary opponent
- Memory requirement: #opponent actions

Convergence Properties for FP

- Convergent cases:
  - zero-sum games (1951)
  - 2x2 games (1961)
  - identical interest “team” games (1996)
  - potential games (2002)
  - 2xN games (2003)
- Counterexamples:
  - Shapley fashion game (1964)
  - Jordan anticoordination game (1993)
  - Foster/Young merry-go-round game (1998)



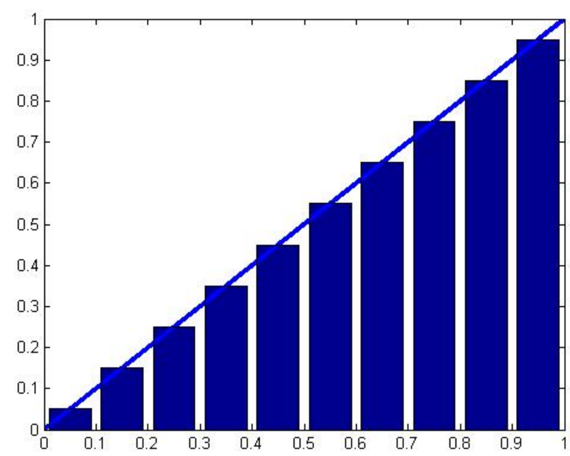


Calibrated Forecasts

- Finite collection of forecasts $f(k) \in \{f_1, f_2, \ldots, f_N\}$
- Calibration condition (asymptotically)

$$\frac{1}{K}\sum_{k=1}^{K} I_{[f(k)=f_i]}\,(a(k) - f_i) \leq \varepsilon$$

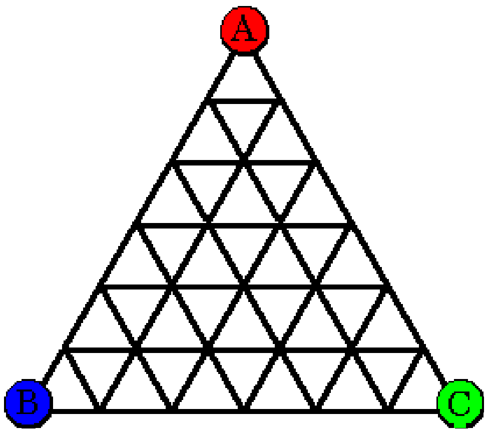
- Implication: Consistency for all *persistent* forecasts



Calibration & Learning

- Calibrated forecast of opponent (FV97):
  - Universal property against arbitrary opponent
  - Self play: Strategies converge to set of *correlated equilibria*
- Calibrated forecast of *joint* behavior (KF04):
  - Universal property against arbitrary opponent
  - Self play: Strategies converges to *convex hull of Nash equilibria*
- Memory requirement of existing algorithms: #forecasts

Grows exponentially in #outcomes



Trade-Offs

- Empirical frequencies
  - Low memory requirement
  - Calibrated vs stationary opponents
- Calibrated forecasts
  - Exponential memory requirements
  - Universally calibrated
- *Tracking Forecasts*
  - Memory requirements of empirical frequencies
  - Calibrated for *classes* of opponent models

Tracking Forecast

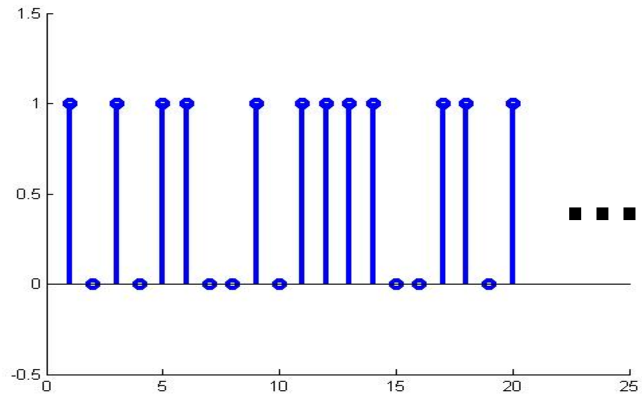
$$f(k+1) = f(k) + \left(\frac{1}{k+1}\right)^{\rho} (a(k) - f(k))$$

$$0 < \rho < 1$$

- Empirical frequencies: $\rho = 1$
- Smaller $\rho \Rightarrow$ Heavier weight on recent outcomes
- Same memory requirements as empirical frequencies
- No discretization
- Weakly calibrated for
  - Stationary opponent
  - *And* broader class of opponents

Analysis (via Stochastic Approximation)

- *Theorem (Mannor, JSS, Arslan 2007):* Tracking forecast is calibrated for
  - Binary sequences ($0 < \rho < 1$)



  - "Relatively slow" sequences ($1/2 < \rho < 1$) (e.g., FP, external regret matching)

$$
\begin{aligned}
X(k+1) &= X(k) + \frac{1}{k+1}\left(F(X(k), e(k), f(k)) + M(k)\right)\\
p(k) &= h(X(k), e(k), f(k))\\
a(k) &= \mathtt{rand}[p(k)]
\end{aligned}
$$

Implication for Learning in Games

- Opponent: Smooth fictitious play (“slow”)
- Strategy: Best response to *tracking* forecast
- Consequence: Play best response to opponent’s *current* strategy

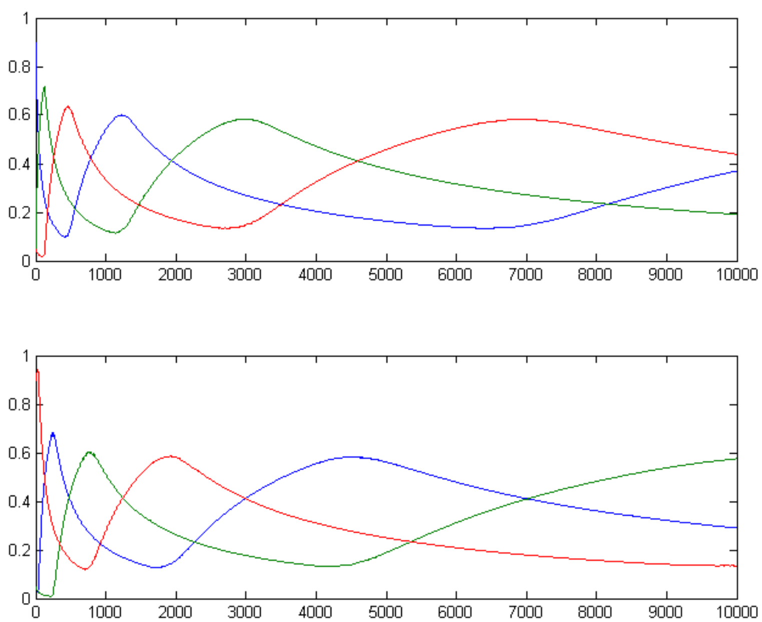
$$f_2 \to \beta_2(q_1)$$

$$\Rightarrow$$

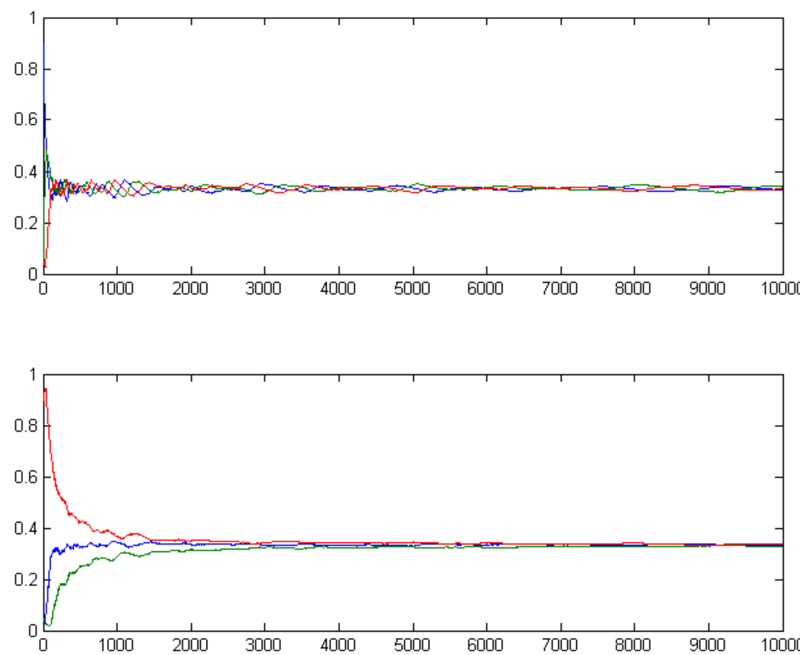
$$\underbrace{\beta_1(f_2)}_{p_1} \to \beta_1(\underbrace{\beta_2(q_1)}_{p_2})$$

Simulations: TF vs FP

$$\begin{pmatrix} 0,0 & 1,0 & 0,1 \\ 0,1 & 0,0 & 1,0 \\ 1,0 & 0,1 & 0,0 \end{pmatrix}$$

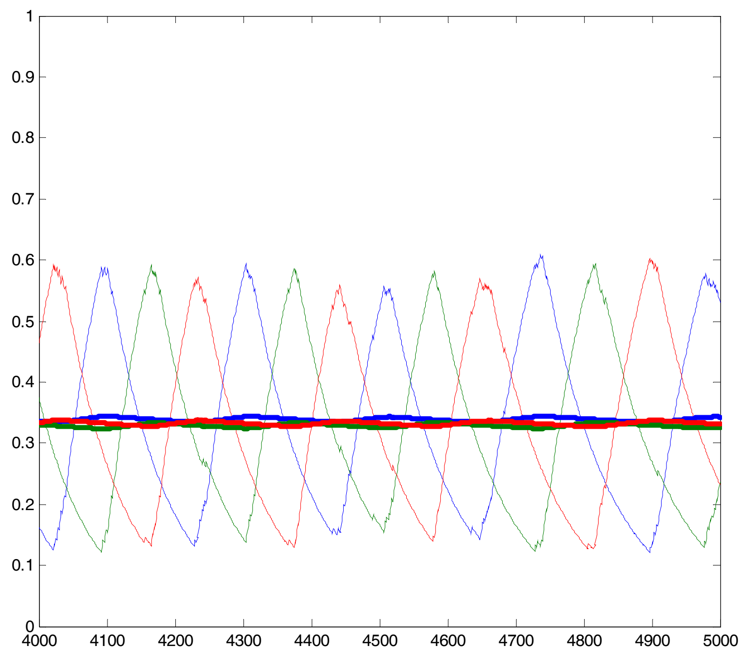


FP vs FP. Avg rewards = (1/2,1/2)



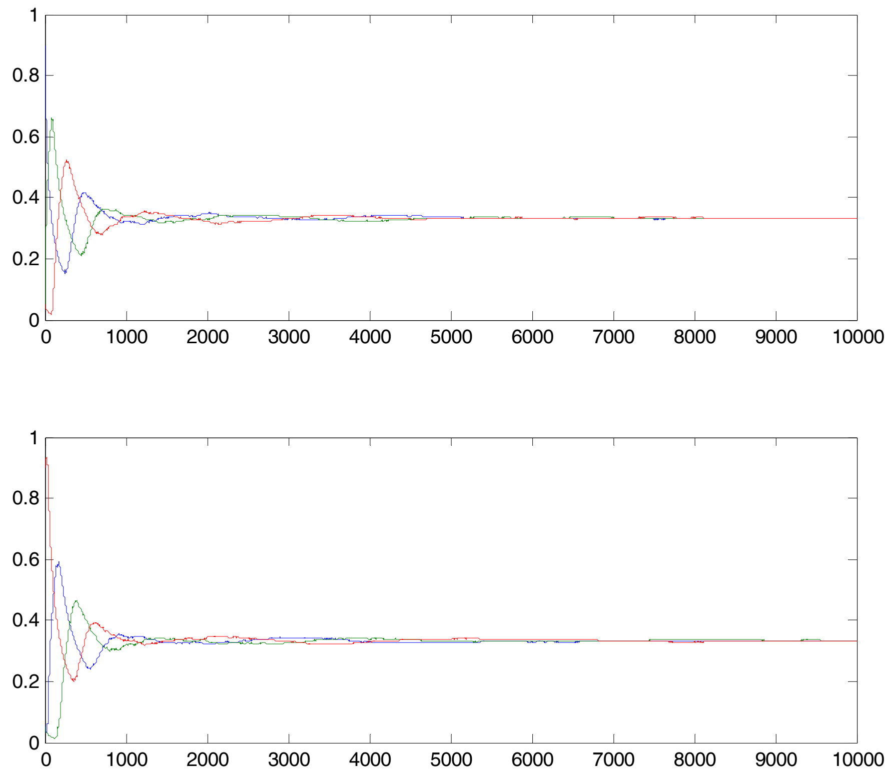
TF vs FP. Avg rewards = (1/3,1/3)

Simulations: TF vs TF



- Thin lines: Tracking forecasts
- Thick lines: Empirical frequencies “flatten out”

Simulations: TF/FP vs TF/FP



- Forecast = Convex combination of TF & Empirical Frequencies
- Analysis: Characterize when convergence to Nash Equilibrium possible
- Similar convergence for Jordan Game

Final Remarks

- Recap:
  - Tracking forecast & trade-offs
  - Parallel efforts: Game theoretic methods for networked systems

- Key issue: *Equally capable/rational agents*
  - e.g., Foster & Young, 2001: “On the impossibility of predicting the behavior of rational agents”