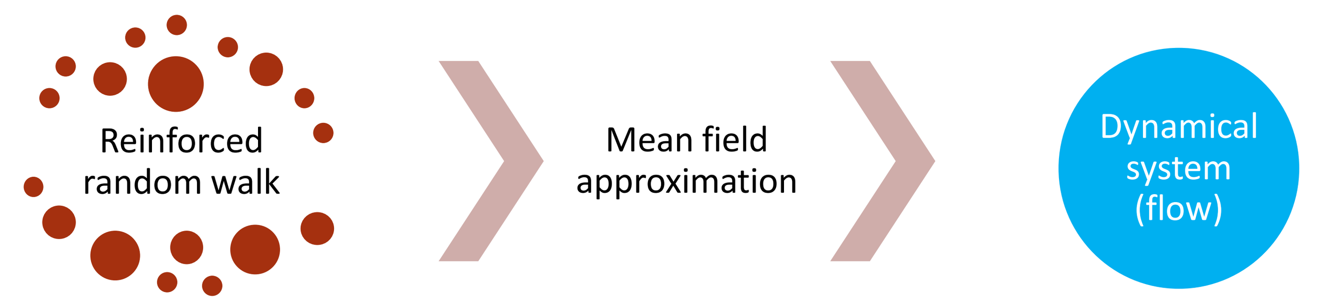
Escaping Saddle Points in Constant Dimensional Spaces: an Agent-based Modeling Perspective

Grant Schoenebeck, *University of Michigan*; Fang-Yi Yu, *Harvard University*

Pipeline to analyze multi-agent systems



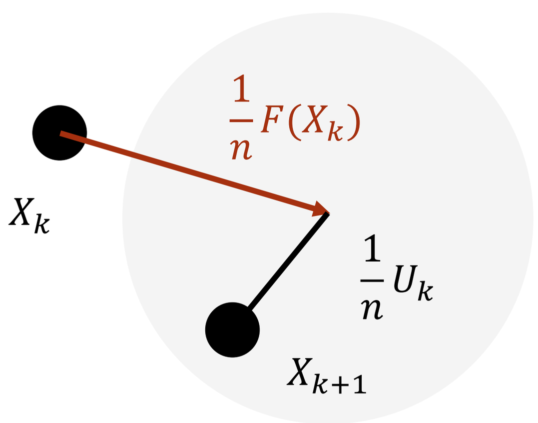
- Dynamics on social networks
- Evolutionary Game theory
- Stochastic Gradient Descent

- Escaping saddle points
- Phase portraits

Reinforced random walk with F

$$X_{k+1} - X_k = \frac{1}{n}\left(F(X_k) + U_k\right)$$

- Expected difference (drift), $F(X)$
- Unbiased noise (noise), U_k
- Step size, $1/n$



Examples

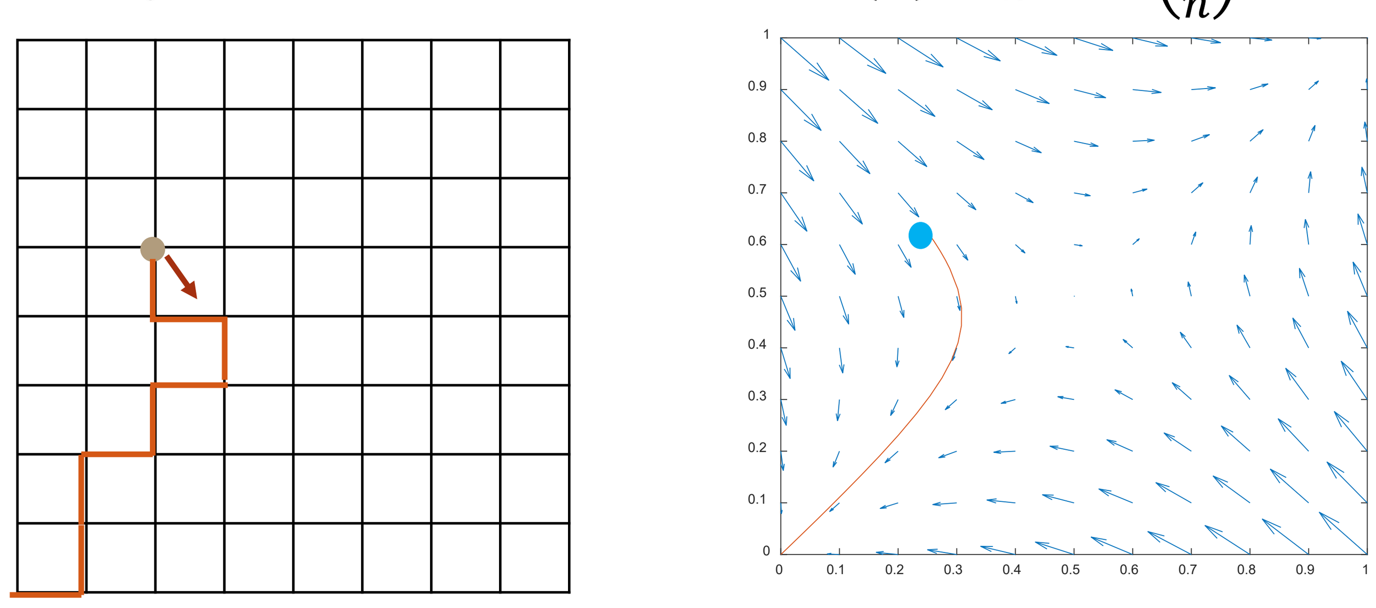
- Stochastic gradient descent
  - Objective function $H: \mathbb{R}^d \to \mathbb{R}$
  - Parameters $X_t \in \mathbb{R}^d$
  - $X_{k+1} = X_k - \eta\left(\nabla H(X_k) + U(X_k)\right)$
- Iterative majority on a complete graph with n nodes
  - Fraction of red opinion: $X_k \in [0,1]$
  - $X_{k+1} = X_k - \frac{1}{n}(1[X_k > 0.5] - X_k + noise)$

Mean field approximation

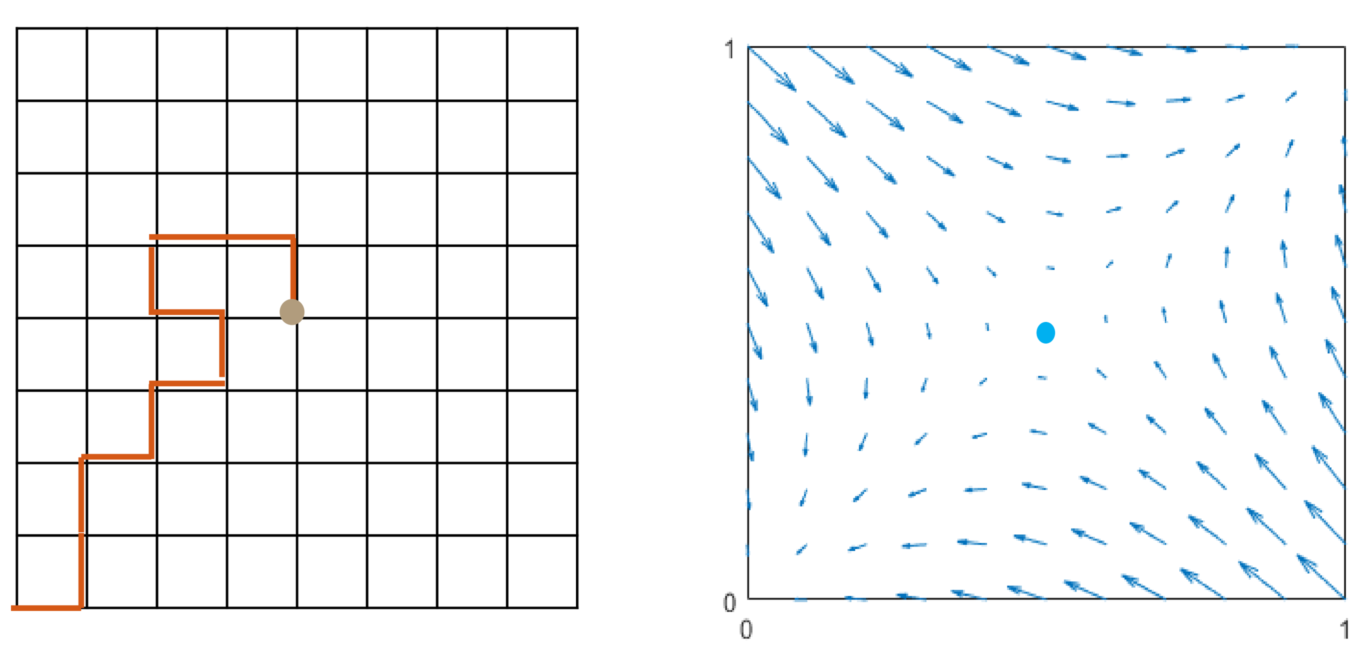
$$x'(t) = F(x(t))$$

Limit of mean field approximation

If X_0 is a regular point, for $k = O(n)$, $X_k \approx x\left(\frac{k}{n}\right)$

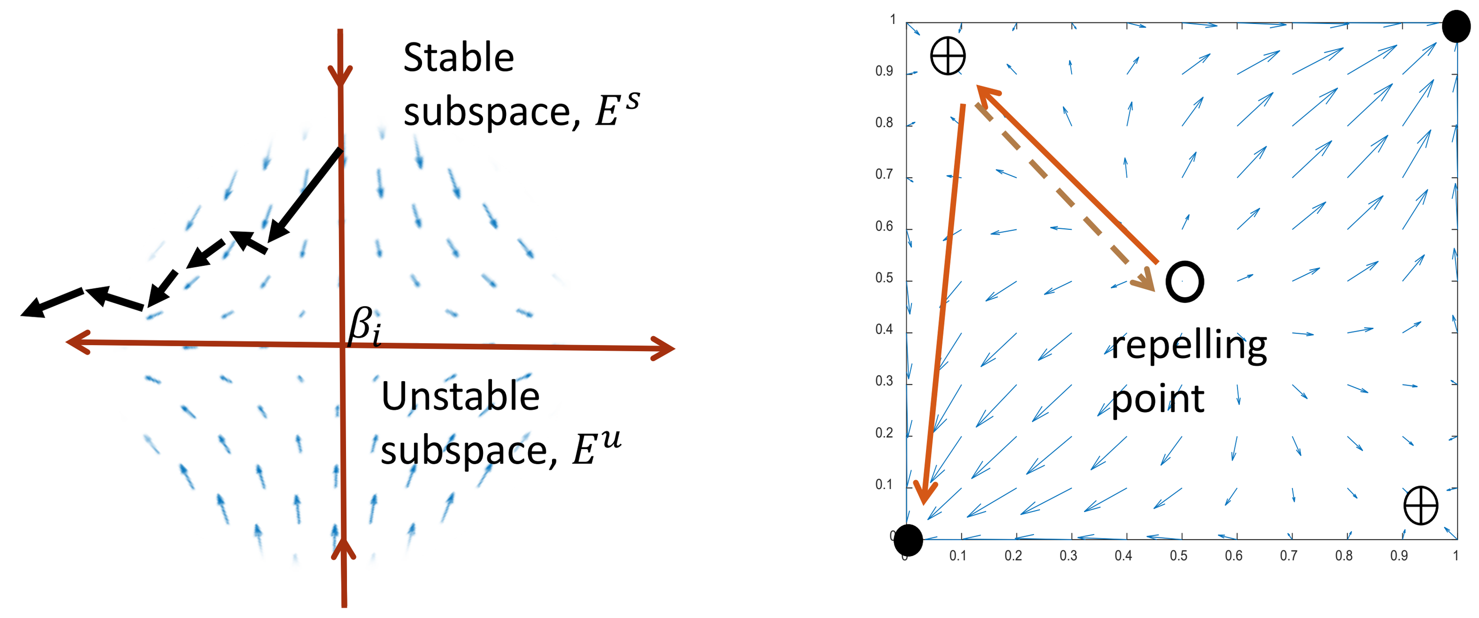


What would happen if X_0 is at a fixed point?



Theoretical results

1. Local: X_k escapes a saddle point β_i in $\Theta(n \log n)$ steps if F is smooth and U is noisy enough.



Local Global

2. Global: X_k reaches an attracting fixed point in $\Theta(n \log n)$ steps if F is gradient-like.

(Dis)agreement between communities

Echo chamber

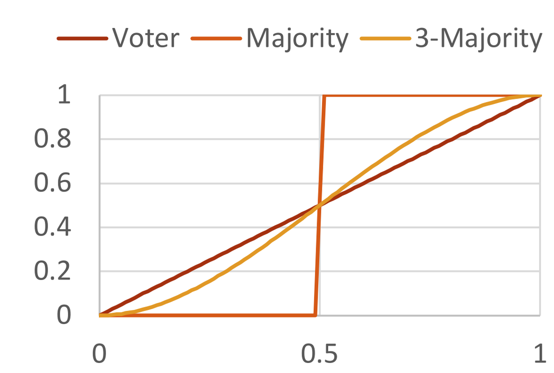
- Beliefs are amplified through interactions in segregated systems
- What is the consensus time given a rich-get-richer opinion formation and the level of intercommunity connectivity?

Node Dynamic $\mathrm{ND}(G, f_{ND}, X_0)$

- Parameters
  - Fixed a (weighted) graph $G = (V, E)$
  - update function f_{ND}
  - initial configuration $X_0: V \mapsto \{0,1\}$
- At round t,
  1. A node v is picked uniformly at random
  2. $X_t(v) = 1$ *w.p.* $f_{ND}\left(r_{X_{t-1}(v)}\right)$; $= 0$ otherwise

Fraction of red neighbors

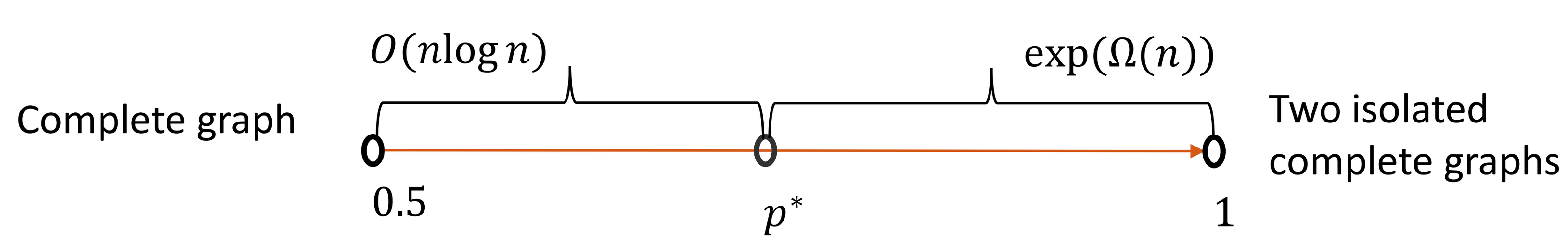
- Examples of ND
  - Voter model
  - Iterative majority
  - 3-majority

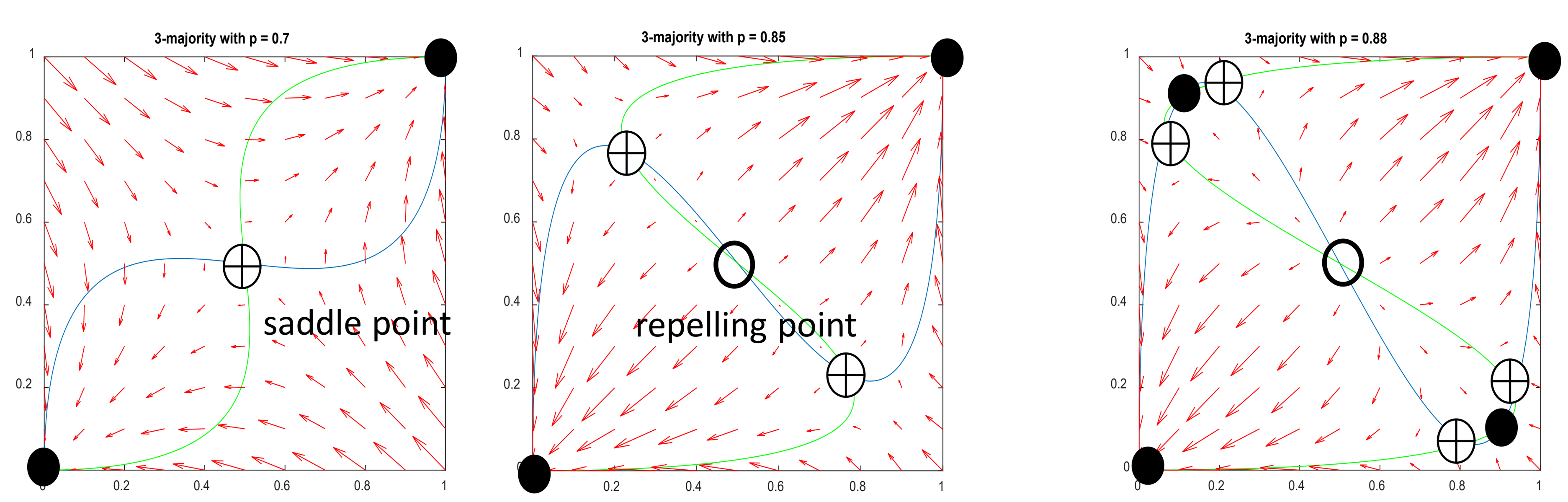


Planted Community $K(n, p)$

- Two communities with equal size
- An edge has weight p if in the same community and $1 - p$ o.w.

Theorem: Given a smooth rich-get-richer function $f_{ND} \in \mathcal{C}^2$, and a planted community graph $G = K(n, p)$, The maximum expected consensus time of $\mathrm{ND}(G, f_{ND}, X_0)$ has two cases:





Attracting point