EVOLUTIONARY DYNAMICS IN FINITE POPULATIONS MIX RAPIDLY

Ioannis Panageas, Georgia Institute of Technology

Joint work (and slides) with:

Piyush Srivastava, Nisheeth K. Vishnoi

Finite Population Models for Evolution

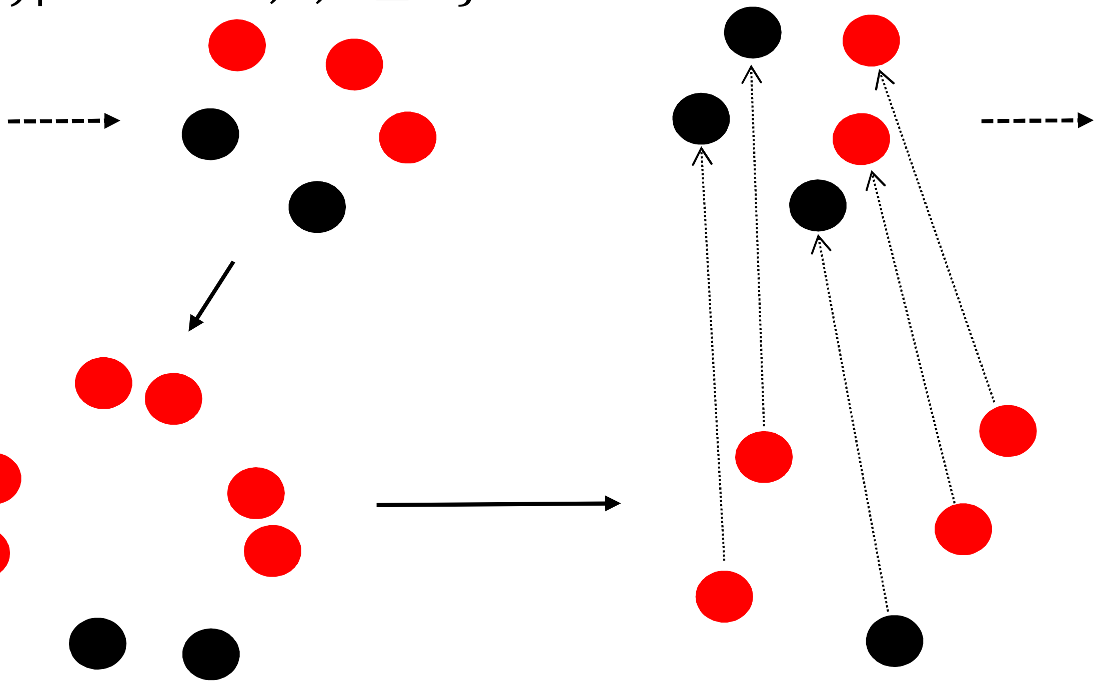
- **Large population with different (geno-)types**
  - Each type has a ***fitness*** – rate of reproduction
  - ***Mutations*** can change one to the other
  - Subject to evolutionary pressure of ***selection***

- Numerous models: e.g. Wright-Fisher ~1900s
  - Applications beyond evolution

- ***The problem:*** ***time to steady state/mixing time?***
  - ***Exponential state space***
  - Computationally relevant to *obtain statistics of steady state*

An Example

- **Types:** $\{\mathbf{0}, \mathbf{1}\}$; **Population size:** N
- **Mutation matrix:** $Q = \begin{bmatrix} \mathbf{1} - \boldsymbol{\mu} & \boldsymbol{\mu} \\ \boldsymbol{\mu} & \mathbf{1} - \boldsymbol{\mu} \end{bmatrix}$; $0 \to 1$ *w.p.* μ *and* $1 \to 0$ *w.p.* μ
- **Fitnesses:** $a_0 (= \boldsymbol{a})$, $a_1 (= \mathbf{1})$
- Markov chain on : $\{(s,t) | s + t = N, s,t \geq 0\}$



Explosion of State Space

- **Parameters**
  - Types $= \{1, \ldots, m\}$
  - Q: $m \times m$ stochastic **mutation** matrix
  - $\{a_\sigma\}$ for $\sigma \in \{1, \ldots, m\}$ – **fitness** landscape – diagonal matrix A
  - $N =$ Population Size
  - $X_\sigma(t)$ - number of σ at time t; random variable
  - $\sum_\sigma X_\sigma(t)$ - total population at time t is $\boldsymbol{N}$ $\forall \boldsymbol{t}$

- **The Markov Chain**
  - Reproduction - Each σ produces a_σ copies of itself ; $a_\sigma X_\sigma(t)$
  - Selection - N genomes chosen at random with replacement
  - Mutation - Each selected genome mutates according to Q
  - State space of the chain $\sim N^m$. $m = 30, N = 10{,}000; \gg 2^{350}$!

Convergence and the Mixing Time

- When $\mu > 0$, chain is **ergodic**

- **Mixing time:** $t_{mix}(\varepsilon)$: min. t s.t. the distance between the distribution at time t and the steady state is at most ε – no matter where you start
  - Simulating and waiting for statistics to stabilize can be misleading

- ***Upper bound on the mixing time***
  - translates to ***computational efficiency*** of sampling from steady state
  - Addresses the question: ***the speed of evolution***

Prior Work and Our Result

- **[DSV '12]** mixing time $\sim \log N$ **-** restrictive parameter setting

- **[Vishnoi '15]** **m=2,** for all constant a, μ; $\sim \log N$

- **[This paper] mixing time** $\sim \log N$ **for all** $\mathbf{m}, \boldsymbol{Q}, \boldsymbol{A}$
  - ***evolution guided by any nice dynamical system***
  - sampling from steady state efficient!

- Mixing time bound derived using geometry of the guiding dynamical system

Markov Chains and Dynamical Systems

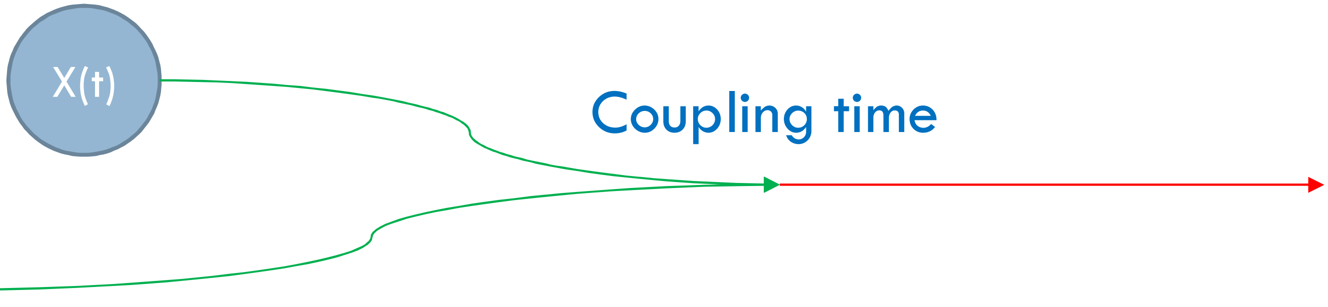
- **Recall:** $X_\sigma(t)$ - number of σ at time t

- **Observation 1:**
  - $\frac{\mathrm{E}[X(t+1)\,|X(t)]}{N} = \frac{QA\,X(t)}{\|QA\,X(t)\|_1} = \boldsymbol{f}(X(t)/N)$
    - $f: \Delta_m \mapsto \Delta_m$ - dynamical system over simplex
    - f captures the *expected motion* of the MC

If $QA > 0$, f has a ***unique stable fixed point*** $\boldsymbol{v}$ inside the simplex

Markov Chains and Dynamical Systems

- **Coupling method:** Show that ***coupled*** copies $\boldsymbol{X(t)}, \boldsymbol{Y(t)}$ collide fast
  - Coupling time $\leq T$ $(prob \geq \frac{3}{4}) \Rightarrow$ Mixing time $\leq T$



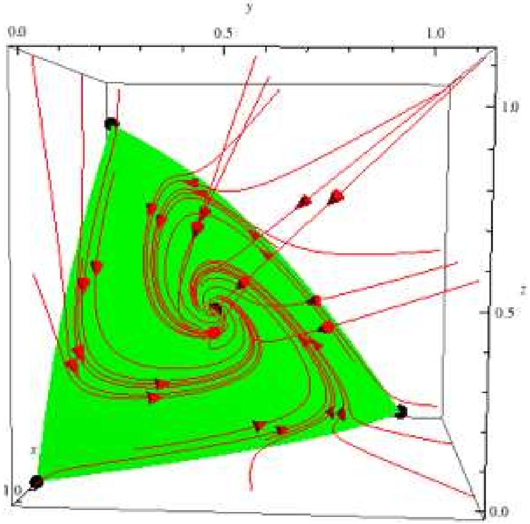
- **Key:** Use driving system f to construct coupling
  - Try to use stability of f + concentration to make the copies collide

Coupling and the driving system

- **Condition on f:** f has a stable fixed point v
- ***Stability*** is determined by the ***Jacobian J*** of f near $\boldsymbol{v}$
  - $|\lambda_{max}(J)| < \rho < 1$ for all $x \sim v$

Observation:

- f quickly drives $X(t), Y(t)$ within dist. ${}^{1}/_{\sqrt{N}}$ to v
- Due to variance, it cannot drive them closer!



But: Coupling method requires a collision

$$\|\boldsymbol{X}(\boldsymbol{t}) - \boldsymbol{Y}(\boldsymbol{t})\|_1 < 1/N$$

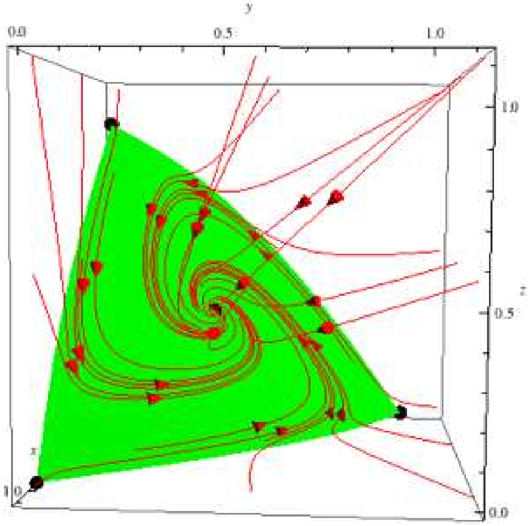
Coupling and the Dynamical System

- **Couple chains using the driving system f**

$$E[\|X(t+1) - Y(t+1)\|_1 | X(t), Y(t)] = \|f\left(X(t)\right) - f\left(Y(t)\right)\|_1$$
$$\leq \|J\|_{1-1}\|X(t) - Y(t)\|_1$$

- But f can actually drive the chains farther in one step!
  - This problem does not arise when $m = 2$

- Must use k steps together - the above equation does not chain!
  - $|\lambda_{max}(J(x))| < 1$ for all $x \sim v$
  - $\Rightarrow$ For some k, $\|J^k(x)\|_{1-1} < \rho^k < 1$, $x \sim v$
  - Need to handle stochastic noise
    - $X(t+i+1) = f\left(X(t+i)\right) + u_i$
    - $Y(t+i+1) = f\left(Y(t+i)\right) + v_i$



Perturbed coupled evolutions

- $\|\boldsymbol{X}(\boldsymbol{t}+\boldsymbol{k})-\boldsymbol{Y}(\boldsymbol{t}+\boldsymbol{k})\|_{\mathbf{1}} \leq \rho^{k}\|\boldsymbol{X}(\boldsymbol{t})-\boldsymbol{Y}(\boldsymbol{t})\|_{\mathbf{1}}$

  provided stochastic noise remains small

  relative to current distance between copies

- Effect of noise worsens as chains come closer!

- Multi-stage argument using stability
  - $\frac{\log N}{N}$ distance w.h.p in $O(\log N)$ steps
  - $\frac{1}{N}$ distance with const. probability in $O(\log\log N)$ steps
  - Collision at next step with constant probability

Discussion and Open Questions

- Improve dependency on parameters?

- Characterize the mixing time of the described generic model for the following cases:
  - Function f has at least two ***attractive*** fixed points?
  - Limit cycles?

- Use tools from Dynamical systems to analyze MC's?