Efficient Simulation Of A Simple Evolutionary System

Mahendra Duwal Shrestha

The University Of Tenessee

March 2, 2017

Outline

Background

Question 1: Distance between finite and infinite population

Question 2: Oscillation in finite population

Question 3: Oscillation in finite population under violation in mutation

Question 4: Oscillation in finite population under violation in crossover

Conclusion

Terms

Population P: a collection of length ℓ binary strings

Population vector $\mathbf{p}$: $\mathbf{p}_j$ is the proportion of string j in the population

If $P = 00, 01, 01, 10, 11, 11$, then $\mathbf{p}_3 = 2/6 = 1/3$

$\mathcal{R}$ denotes a set binary strings of length ℓ

- Addition and multiplication of elements in $\mathcal{R}$ are bitwise operations modulo 2
- $x = 1101$, $y = 1010$
- $x + y = 1101 + 1010 = 0111$
- $xy = 1101 \cdot 1010 = 1000$
- $\bar{x} = 0010$

Crossover & Mutation

Crossover : Choose parents u and v, exchange bits using crossover mask m:

$u' = um + v\bar{m}, v' = u\bar{m} + vm$

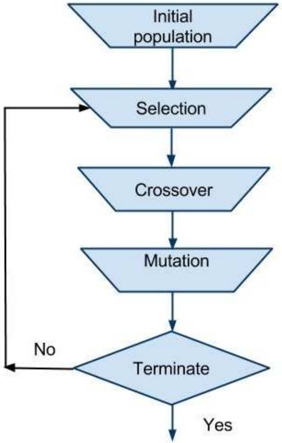
$u = \mathbf{11001011}, v = 11011111, m = 11110000$

$\{\mathbf{11001011}, 11011111\} \to \{\mathbf{1100}0000 + 00001111,\ 0000\mathbf{1011} + 11010000\} \to \{\mathbf{1100}1111, 1101\mathbf{1011}\}$

Mutation: Flip bits using mutation mask:

$x \to x + m$

Finite Population GA



Randomly select parents u and v

Crossover u and v to produce u' and v'

Keep one of u', v' , and mutate

Repeat above to form next generation

Repeat whole process until system stops to improve or threshold is reached

Infinite Population Model

Population is modeled as a vector $\mathbf{p}$

$\mathcal{G}$ maps $\mathbf{p}$ to the next generation

$\mathcal{G}(\mathbf{p})_j$ = probability that string j occurs in the next generation

The infinite population model is the sequence

$\mathbf{p} \to \mathcal{G}(\mathbf{p}) \to \mathcal{G}(\mathcal{G}(\mathbf{p})) \to \cdots$

Random Heuristic Search

τ is a stochastic transition rule that maps $\mathbf{p}$ to $\mathbf{p}$

For a finite population, sequence $\mathbf{p}, \tau(\mathbf{p}), \tau^2(\mathbf{p}), \cdots$ forms Markov chain

$\tau(\mathbf{p})$ cannot be predicted with certainty

$\mathcal{G}(\mathbf{p})$ is the expected next generation $\mathcal{E}(\tau(\mathbf{p}))$

The variance in the next generation is

$$\mathcal{E}(\|\tau(\mathbf{p}) - \mathcal{G}(\mathbf{p})\|^2) = \frac{1 - \|\mathcal{G}(\mathbf{p})\|^2}{r}$$

Question 1: Distance Between Finite and Infinite Population

Chebyshev's inequality $\rightarrow$

$$\|\tau(\mathbf{p}) - \mathcal{G}(\mathbf{p})\| \leq \frac{k}{\sqrt{r}}$$

Jensen's inequality $\rightarrow$

$$\mathcal{E}(\|\tau(\mathbf{p}) - \mathcal{G}(\mathbf{p})\| \leq \frac{\sqrt{1 - \|\mathcal{G}(\mathbf{p})\|^2}}{\sqrt{r}}$$

Geometric point of view $\rightarrow$

$$\|\tau(\mathbf{p}) - \mathcal{G}(\mathbf{p})\| = O(r)$$

Does the distance decrease in practice like $1/\sqrt{r}$?

Diploid Population Model

Diploid genome: $\alpha = \langle \alpha_0, \alpha_1 \rangle$

$\mathbf{q}^n \rightarrow$ population at generation n

$\mathbf{q}^n_\alpha \rightarrow$ prevalence of diploid α

$t_\alpha(g) \rightarrow$ probability that gamete g is produced from parent α

$$\mathbf{q}^{n+1}_\gamma = \sum_\alpha \mathbf{q}^n_\alpha \, t_\alpha(\gamma_0) \sum_\beta \mathbf{q}^n_\beta \, t_\beta(\gamma_1)$$

Diploid Model Reduction to Haploid Model

Diploid distribution in terms of haploids

$$q^n_{\langle\gamma_0,\gamma_1\rangle} \;=\; p^n_{\gamma_0}\, p^n_{\gamma_1}$$

Haploid distribution in terms of diploids

$$\mathbf{p}^n_g \;=\; \frac{1}{2}\sum_{\alpha_0,\,\alpha_1} \mathbf{q}^n_{\langle\alpha_0,\alpha_1\rangle}([g=\alpha_0]+[g=\alpha_1])$$

Evolution equation in terms of haploid distributions **p**,

$$\mathbf{p}^{n+1}_{\gamma_0} \;=\; \sum_{\alpha_0,\,\alpha_1} \mathbf{p}^n_{\alpha_0}\,\mathbf{p}^n_{\alpha_1}\, t_{\langle\alpha_0,\,\alpha_1\rangle}(\gamma_0)$$

Matrix form:

$$\mathbf{p}'_g \;=\; \mathbf{p}^T M_g\, \mathbf{p}$$

where

$$\Big(M_g\Big)_{u,v} \;=\; t_{\langle u,v\rangle}(g)$$

Specialization to Vose's Haploid Model

Mutation:

$$\boldsymbol{\mu}_i = (\mu)^{\mathbf{1}^T i}(1-\mu)^{\ell-\mathbf{1}^T i}$$

Crossover:

$$\boldsymbol{\chi}_i = \begin{cases} \chi c_i & \text{if } i > 0 \\ 1-\chi+\chi c_0 & \text{if } i = 0 \end{cases}$$

For uniform crossover, $c_i = 2^{-\ell}$

$$t_{\langle u,v \rangle}(g) = \sum_{i \in \mathcal{R}} \sum_{j \in \mathcal{R}} \sum_{k \in \mathcal{R}} \boldsymbol{\mu}_i \boldsymbol{\mu}_j \frac{\boldsymbol{\chi}_k + \boldsymbol{\chi}_{\overline{k}}}{2} [k(u+i) + \overline{k}(v+j) = g]$$

where $u, v \in \mathcal{R}$

Walsh Basis

$W_{n,t} = N^{-1/2}(-1)^{n^T t}$ where $N = 2^\ell$

$\widehat{A} = WAW$

$\widehat{w} = Ww$

Mixing matrix in Walsh basis

$$\widehat{M}_{u,v} \;=\; 2^{\,\ell-1}\,[u\,v = \mathbf{0}]\,\widehat{\boldsymbol{\mu}}_u\,\widehat{\boldsymbol{\mu}}_v \sum_{k\in\overline{u+v}\mathcal{R}} \boldsymbol{\chi}_{k+u} + \boldsymbol{\chi}_{k+v}$$

Evolution eqn in Walsh basis

$$\widehat{\mathbf{p}}_g^{\,\prime} \;=\; 2^{\,\ell/2} \sum_{i\in g\mathcal{R}} \widehat{\mathbf{p}}_i\,\widehat{\mathbf{p}}_{i+g}\,\widehat{M}_{i,i+g}$$

where $g\mathcal{R} = \{g\,i \mid i \in \mathcal{R}\}$

Computational Advantages

Specialization simplifies computation, which otherwise for diploid case would have been impractical

Only one mixing matrix as opposed to 2^ℓ is needed to compute next generation

For $\ell = 14$, using 2^{14} mixing matrices with each having $2^{14} \cdot 2^{14}$ entries would require 32 TB of memory, whereas one mixing matrix requires only 2 GB

Distance Computation

Naive computation

$$\|\mathbf{f}-\mathbf{q}\|^2 = \sum_{\alpha}(\mathbf{f}_\alpha - \mathbf{q}_\alpha)^2 \longrightarrow 2^\ell \cdot 2^\ell \textit{ terms}$$

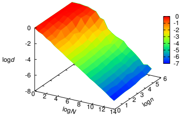
Our implementation

$$\begin{aligned}\|\mathbf{f}-\mathbf{q}\|^2 &= \sum_{\alpha\notin S_\mathbf{f}}(\mathbf{f}_\alpha - \mathbf{q}_\alpha)^2 + \sum_{\alpha\in S_\mathbf{f}}(\mathbf{f}_\alpha - \mathbf{q}_\alpha)^2 \\ &= \sum_{g}^{2}(\mathbf{p}_g)^2 + \sum_{\alpha\in S_\mathbf{f}} \mathbf{f}_\alpha(\mathbf{f}_\alpha - 2\mathbf{q}_\alpha) \to 2^\ell + |S_\mathbf{f}| \textit{ terms}\end{aligned}$$

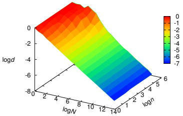
$$S_\mathbf{f} = \{\alpha \,|\, \mathbf{f}_\alpha > 0\}$$

Convergence

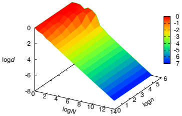
$$\chi = 0.1, \ \mu = 0.001$$



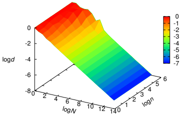
(a) $\ell = 4$



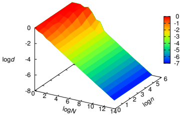
(b) $\ell = 6$



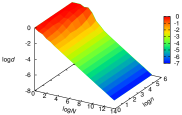
(c) $\ell = 8$



(d) $\ell = 10$



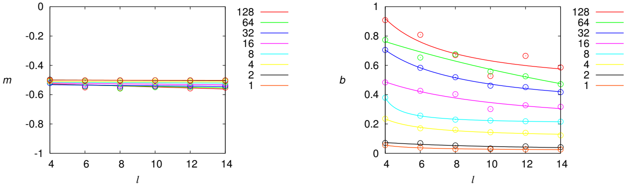
(e) $\ell = 12$



(f) $\ell = 14$

Regression

$$\log d = m \log N + b$$



(a) Slope m **(b)** Intercept b

Figure : Regression parameter for generation $n \in \{1, 2, 4, 8, 16, 32, 64, 128\}$

$d \approx N^m e^b$

From figure (a) above, $m \approx -0.5$

$d \approx k/N$

Distance: Conclusion

Vose's infinite population model makes computation in diploid case efficient by reducing to the haploid case

Distance between finite diploid population and infinite diploid population can decrease like $1/\sqrt{N}$

Question 2

Oscillations in Finite Population Evolution

Limits

The sequence $\mathbf{p}$, $\mathcal{G}(\mathbf{p})$, $\mathcal{G}^2(\mathbf{p})$, $\cdots$ may converge to a fixed point

$\mathcal{G}(\omega) = \lim_{n\to\infty} \mathcal{G}^n(\mathbf{p}) = \omega$

But under some circumstances, the sequence converges to a periodic orbit that oscillates between two fixed points, $\mathbf{p}^*$ and $\mathbf{q}^*$

Periodic Orbit: Necessary and Sufficient Conditions

For some $g \in \mathcal{R}, g \neq 0$

$$
\begin{aligned}
-1 &= \sum_j (-1)^{g^T j} \boldsymbol{\mu}_j \\
1 &= \sum_{k \in \bar{g}\mathcal{R}} \chi_{k+g} + \chi_k
\end{aligned}
$$

Infinite populations converge to a periodic orbit

Do finite populations also exhibit oscillation from random initial populations?

Previous Works on Oscillation

Akin (1982) proved existence of cycling for continuous-time 2-bit diploid model

Hasting (1981) studied cycling in populations with infinite 2-bit diploid population model

Wright and Bidwell (1997) provided examples of cycling in an infinite haploid model with crossover and mutation for 3 bit and 4 bit populations

Wright and Agapie (2001) described cycling in infinite population for up to 4 bits, and also presented data for cycling in finite population

- Akin considered continuous time model and we consider discrete time model
- Hastings' study was limited to two bits length, and includes only crossover but not mutation
- Examples provided by Wright and Bidwell were for specific set parameter values for crossover, mutation and fitness
- Wright and Agapie used dynamic mutation that depends upon where population is in the population space
- Another difference between Wright and Agapie's work and ours is fitness
- We study cyclic behavior for:
  - fixed fitness function and random: initial population, mutation and crossover distribution
  - higher bit length (up to 14)
  - both haploid and diploid populations, and for both finite and infinite populations
- We also visualize oscillation

Simulation

Simulations were run for both haploid and diploid populations

Random initial population

$\ell \in 8, 10, 12, 14$

$N = 4096, 40960, 81920$

To visualize oscillation, distance between fixed points and population are plotted

Oscillation

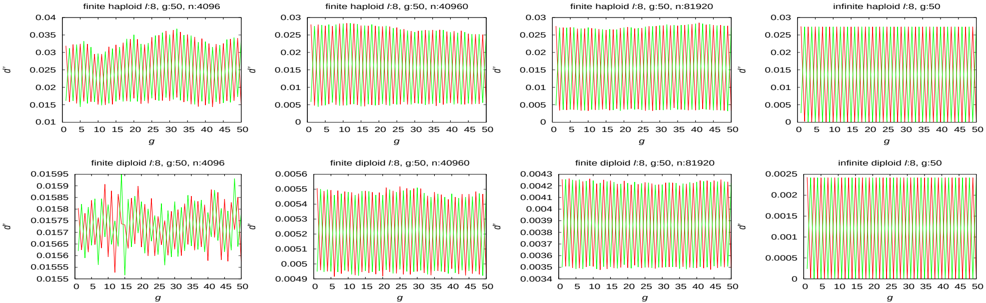


Figure : Infinite and finite population behavior for genome length $\ell = 8$

Oscillation

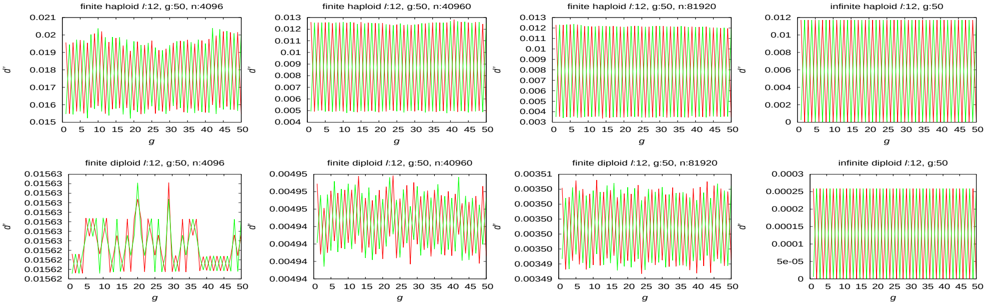
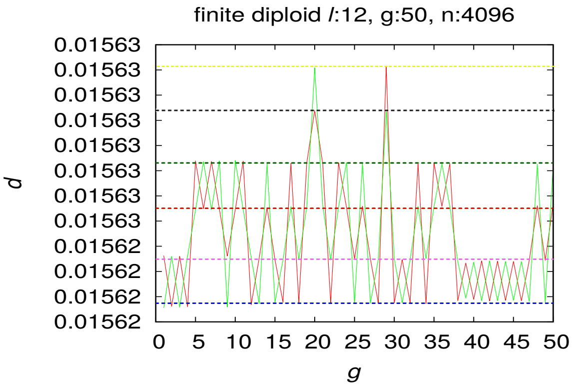


Figure : Infinite and finite population behavior for genome length $\ell = 12$

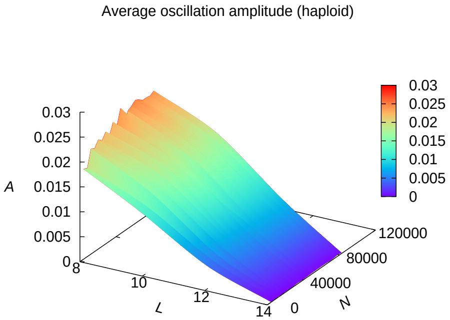
Oscillation: Unusual Behavior



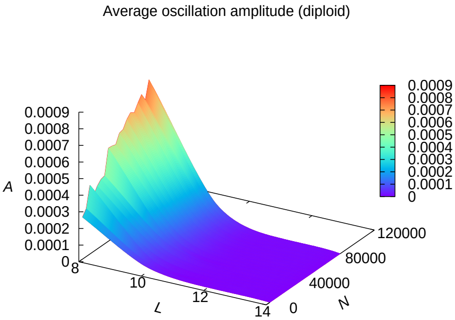
(a)

Figure : Oscillation between different points

Oscillation Amplitude



(a)



(b)

Figure : Average oscillation amplitude

Oscillation: Conclusion

Finite population evolution exhibits approximate oscillations

As ℓ increases, oscillation amplitude decreases

As population size increases, oscillation amplitude increases and randomness decreases

Finite population can also oscillate between different pairs of points for diploid population of smaller size and larger ℓ

Question 3

Oscillations Under Violation in Mutation

Robustness of Finite Population

A Markov chain is said to be *irreducible* if it is possible to get to any state from any state

A Markov chain is *aperiodic* if it can return to state i at irregular times

Markov chain is *regular* if it is both irreducible and aperiodic

Positive steady state distribution exists if Markov chain is regular

No periodic orbit exists for finite population

Can finite population exhibit approximate oscillations?

Violation in Mutation

Violation ϵ is introduced in $\boldsymbol{\mu}$

$\boldsymbol{\mu}_0 = \epsilon$

$\boldsymbol{\mu}_i := (1 - \epsilon)\boldsymbol{\mu}_i$

This modification makes the Markov chain regular

No periodic orbits for finite population

No periodic orbits for infinite population

Simulation

$\epsilon = 0.01,\ 0.1,\ 0.5$

$\ell\ =\ \{8, 10, 12, 14\}$

$N\ =\ \{4096, 40960, 81920\}$

Distances of population to limits p^* and q^* without violation are plotted

Results: Violation in Mutation

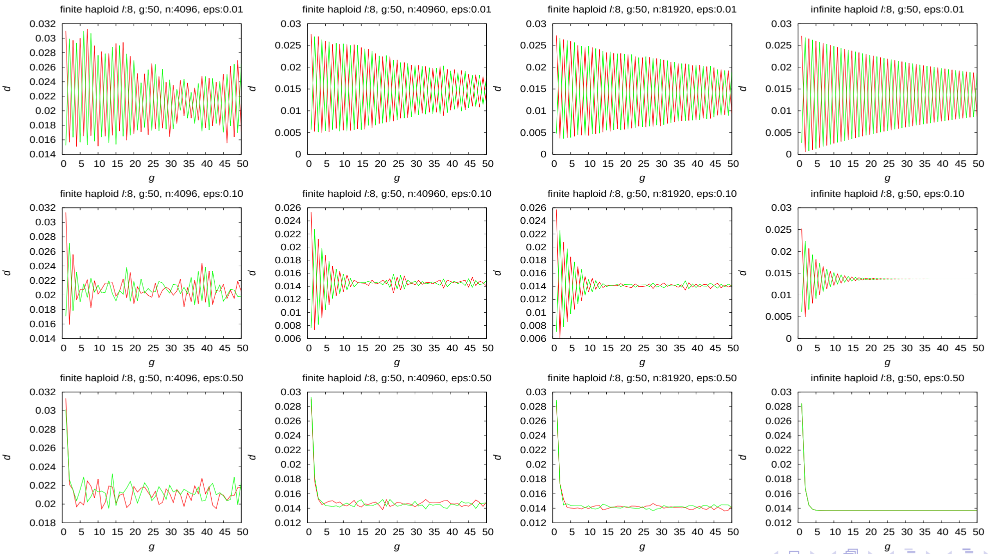


Figure : Infinite and finite haploid population behavior for μ violation and $\ell = 8$

Results: Violation in Mutation

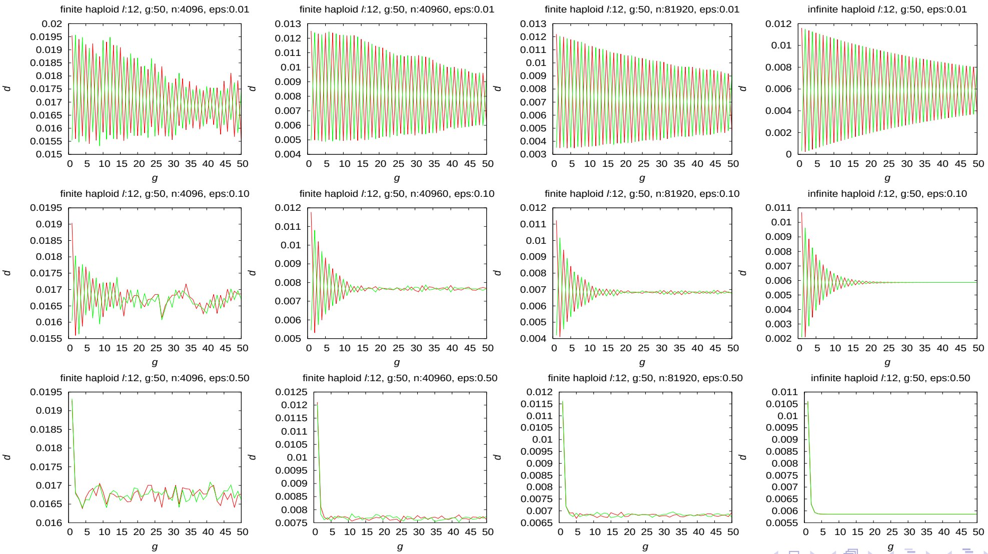


Figure : Infinite and finite haploid population behavior for μ violation and $\ell = 12$

Results: Violation in Mutation

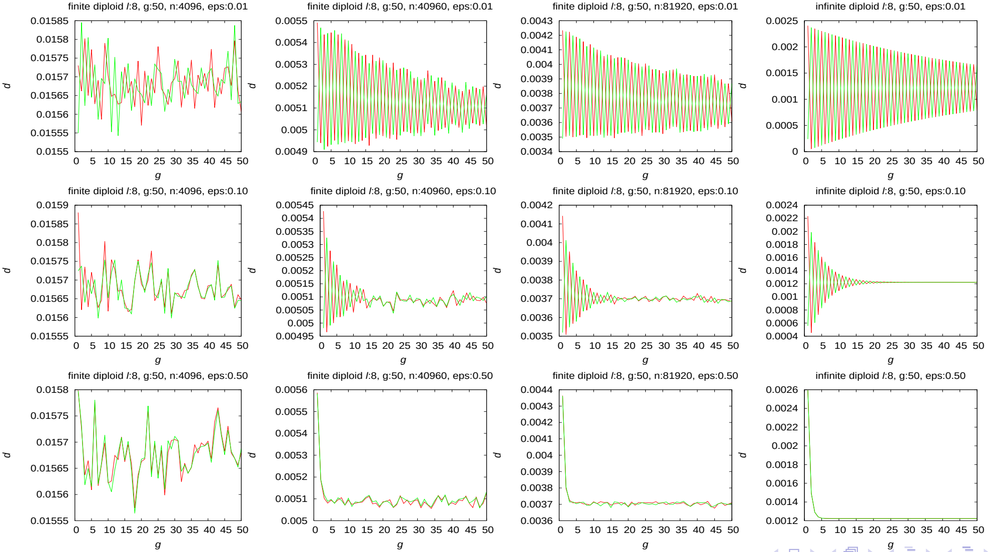


Figure : Infinite and finite diploid population behavior for μ violation and $\ell = 12$

Results: Violation in Mutation

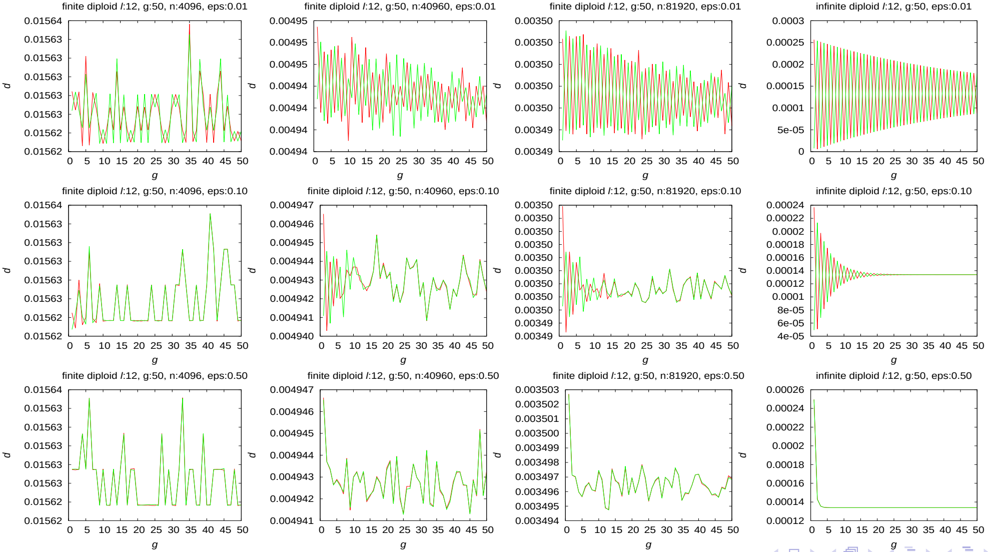


Figure : Infinite and finite diploid population behavior for μ violation and $\ell = 12$

Violation in Mutation: Conclusion

Necessary condition for mutation distribution is violated so that no periodic orbit exists for infinite population

Violation makes Markov chain regular so that no periodic orbit exists for finite population

Finite populations exhibit approximate oscillation even if Markov chain is regular when violation is small

If violation is large, then finite population oscillation decreases

As string length increases, oscillation degrades

Question 4

Oscillations under Violation in Crossover

Robustness of Finite Population

Violation in crossover condition means no periodic orbit exists for infinite population

But we don't know if Markov chain is regular

Can finite population exhibit approximate oscillation?

Violation in Crossover

Violation ϵ is introduced in χ

$\chi_i := (1 - \epsilon)\chi_i$

$\chi_j = \epsilon$ $\qquad$ j is chosen where $\chi_j = 0$

Simulation

$$\epsilon = \{0.01, 0.1, 0.5\}$$

$$\ell = \{8, 10, 12, 14\}$$

$$N = \{4096, 40960, 81920\}$$

Distances of population to limits $\mathbf{p}^*$ and $\mathbf{q}^*$ without violation are plotted

Results: Violation in Crossover

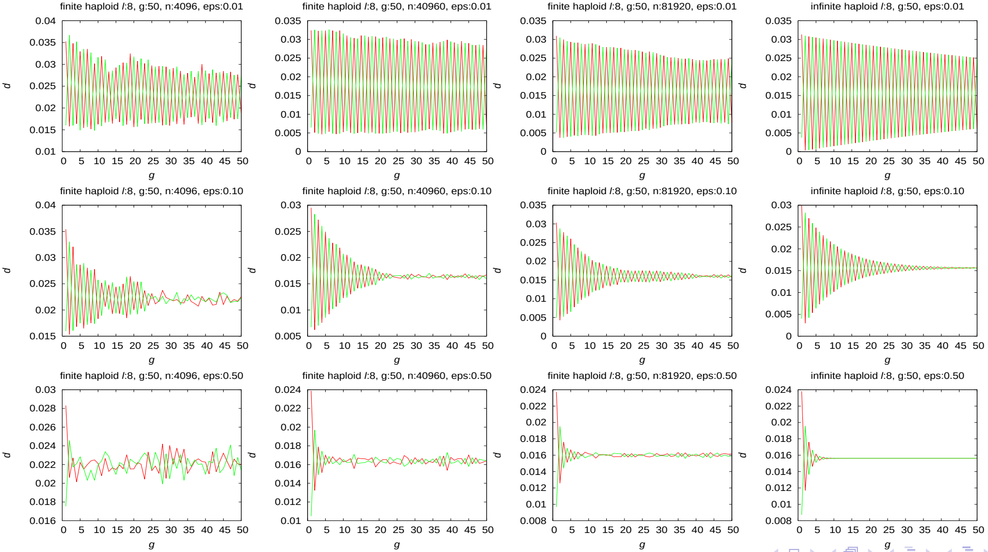


Figure : Infinite and finite haploid population behavior for χ violation and $\ell = 8$

Results: Violation in Crossover

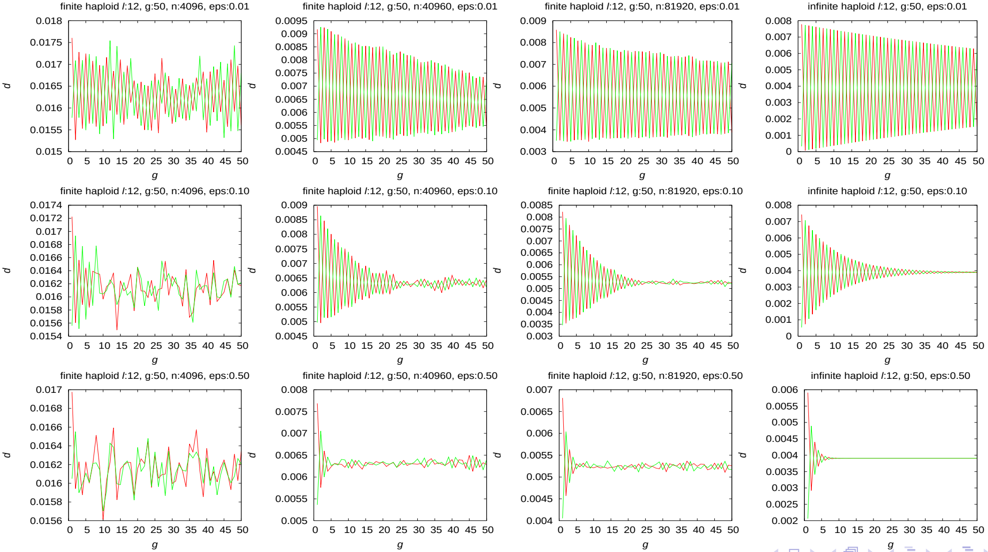


Figure : Infinite and finite haploid population behavior for χ violation and $\ell = 12$

Results: Violation in Crossover

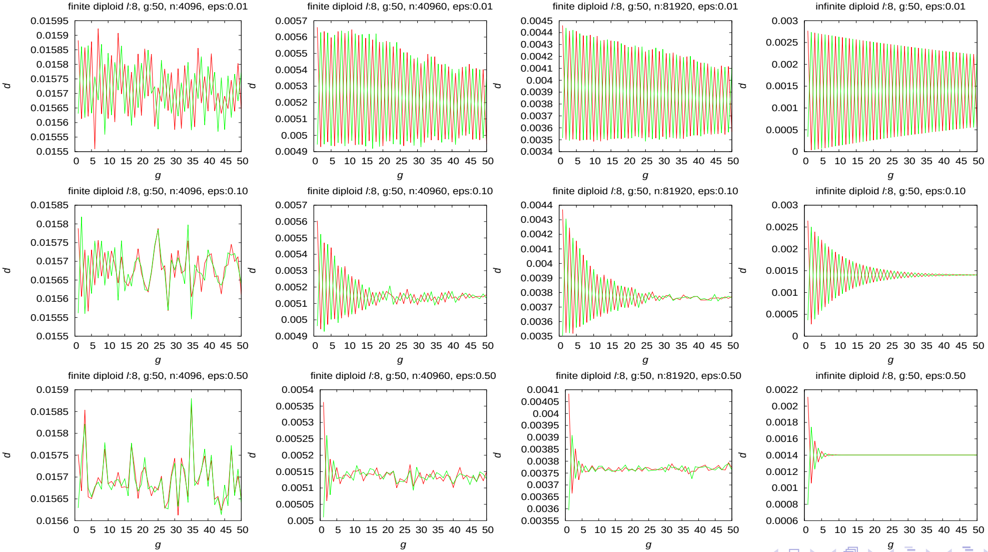


Figure : Infinite and finite diploid population behavior for χ violation and $\ell = 12$

Results: Violation in Crossover

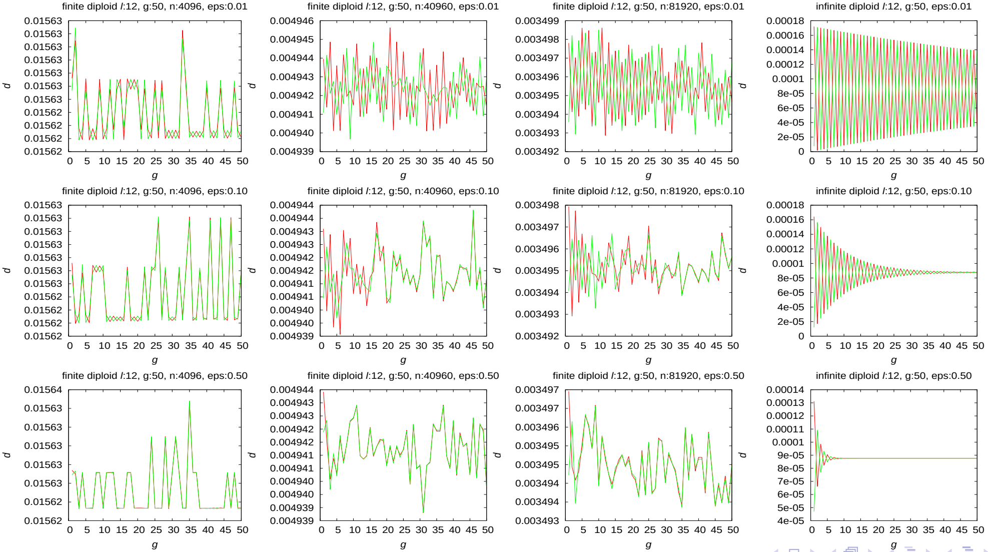


Figure : Infinite and finite diploid population behavior for χ violation and $\ell = 12$

Violation in Crossover: Conclusion

Finite population exhibits approximate oscillation if the violation is small

If violation is large, then finite population oscillation decreases

Rate of damping of oscillation is slower than in violation in mutation case

More randomness are observed with violation in crossover than in mutation

Randomness increases as string length increases

Conclusion

Vose's haploid model makes computation efficient in diploid case by reducing to haploid case

Distance between finite population and infinite population can decrease like $1/N$

When infinite populations oscillate, finite populations exhibit approximate oscillation

When Markov chain is regular, finite population exhibits approximate oscillation for small mutation violation

Finite populations exhibit approximate oscillation for small crossover violation

Thank You!!

Questions?

Chebyshev's Inequality

Let $\epsilon = f(r)/\sqrt{r}$, where $f(r)$ grows arbitrarily slowly such that

$$\lim_{r\to\infty} f(r) = \infty$$

and

$$\lim_{r\to\infty} f(r)/\sqrt{r} = 0.$$

From Chebyshev's inequality,

$$\lim_{r\to\infty} P(\|\tau(\mathbf{p}) - \mathcal{G}(\mathbf{p})\| \geq \epsilon) \;\leq\; \lim_{r\to\infty} \frac{1 - \|\mathcal{G}(\mathbf{p}\|^2}{f(r)^2} \;=\; 0$$

This suggests the distance between $\tau(\mathbf{p})$ and $\mathcal{G}(\mathbf{p})$ might decrease as $1/\sqrt{r}$

Jensen's Inequality

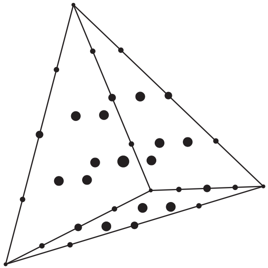
Let η be the random variable $\|\tau(\mathbf{p}) - \mathcal{G}(\mathbf{p})\|$, and convex function be $\phi(x) = x^2$

Then from Jensen's Inequality,

$$\mathcal{E}(\|\tau(\mathbf{p}) - \mathcal{G}(\mathbf{p})\|) \;=\; \mathcal{E}(\eta) \leq \sqrt{\mathcal{E}(\eta^2)} \;=\; \frac{\sqrt{1 - \|\mathcal{G}(\mathbf{p})\|^2}}{\sqrt{\mathbf{r}}}$$

This also suggests the distance might decrease as $1/\sqrt{r}$

Population Points



Finite populations are represented by dots

Infinite population can be any where in the space

Distance between a finite population and an infinite population is $O(1/\sqrt{r})$

This suggests the distance between $\tau(\mathbf{p})$ and $\mathcal{G}(\mathbf{p})$ might decrease as $1/\sqrt{r}$

History

- Haldane, in 1932, summarized basic population genetics models : Wright, Fisher and Haldane
- Several people working with evolution-inspired algorithms in the 1950s and the 1960s Box (1957), Friedman(1959), Bledsoe (1961), Bremermann (1962), and Reed, Toombs and Baricelli (1967)
- In 1960s and 1970s, Holland and colleagues formalized and promoted population based algorithms with crossover and mutation
- Vose (1999) presented efficient methods for computing with a haploid model using mask-based operators introduced by Geiringer (1944)