

SFB 680  
Molecular Basis of  
Evolutionary Innovations

# Evolutionary dynamics in random fitness landscapes

Joachim Krug

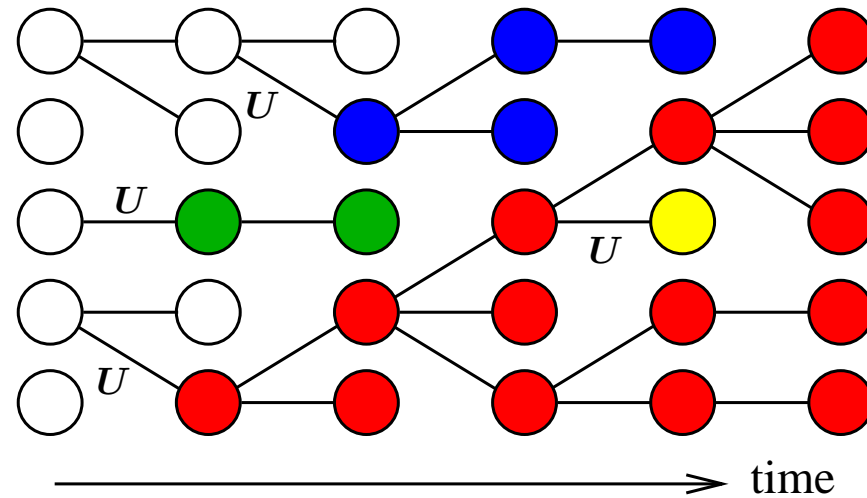
Institute of Theoretical Physics, University of Cologne

- **Random energy** and **house of cards** models of evolution
- Deterministic and stochastic evolutionary regimes  
Genetics **175**, 1275 (2007)
- The infinite sites model  
JSTAT (2008) P04014

Joint work with Kavita Jain and Su-Chan Park

# Evolution of asexual populations

**Basic model:** Wright-Fisher sampling of a finite population of size  $N$



- Each individual chooses an ancestor from the preceding generation
- Individual  $i$  is chosen with probability  $\sim w_i$  **Wrightian fitness**
- Mutations occur with probability  $U$  per individual and generation
- Two distinct sources of fluctuations ( $\sim 1/N, U$ )

# Sequence space and fitness model

- Each individual carries a genetic sequence of length  $L$

$$\sigma = (\sigma_1, \sigma_2, \dots, \sigma_L) \text{ with } \sigma_i = 0, 1 \quad \text{genotype}$$

- **Point mutations**  $\sigma_i \rightarrow 1 - \sigma_i$  occur with probability  $\mu$  per site
- The **Hamming distance**  $d(\sigma, \sigma')$  between two sequences  $\sigma, \sigma'$  is the number of letters in which they differ.
- Fitnesses  $w(\sigma)$  are uncorrelated random variables drawn from a common distribution  $g(w)$  default in this talk:  $g(w) = e^{-w}$

**House of cards model** of population genetics

J.F.C. Kingman, J. Appl. Prob. **15**, 1 (1978)

**Random energy model** of spin glass physics

B. Derrida, Phys. Rev. B **24**, 2613 (1981)

- Both consider (explicitly or implicitly)  $N \rightarrow \infty$

# Classification of evolutionary regimes

K. Jain, JK, Genetics **175**, 1275 (2007)

**Parameters:** Population size  $N$ , mutation probability  $\mu$ , sequence length  $L$

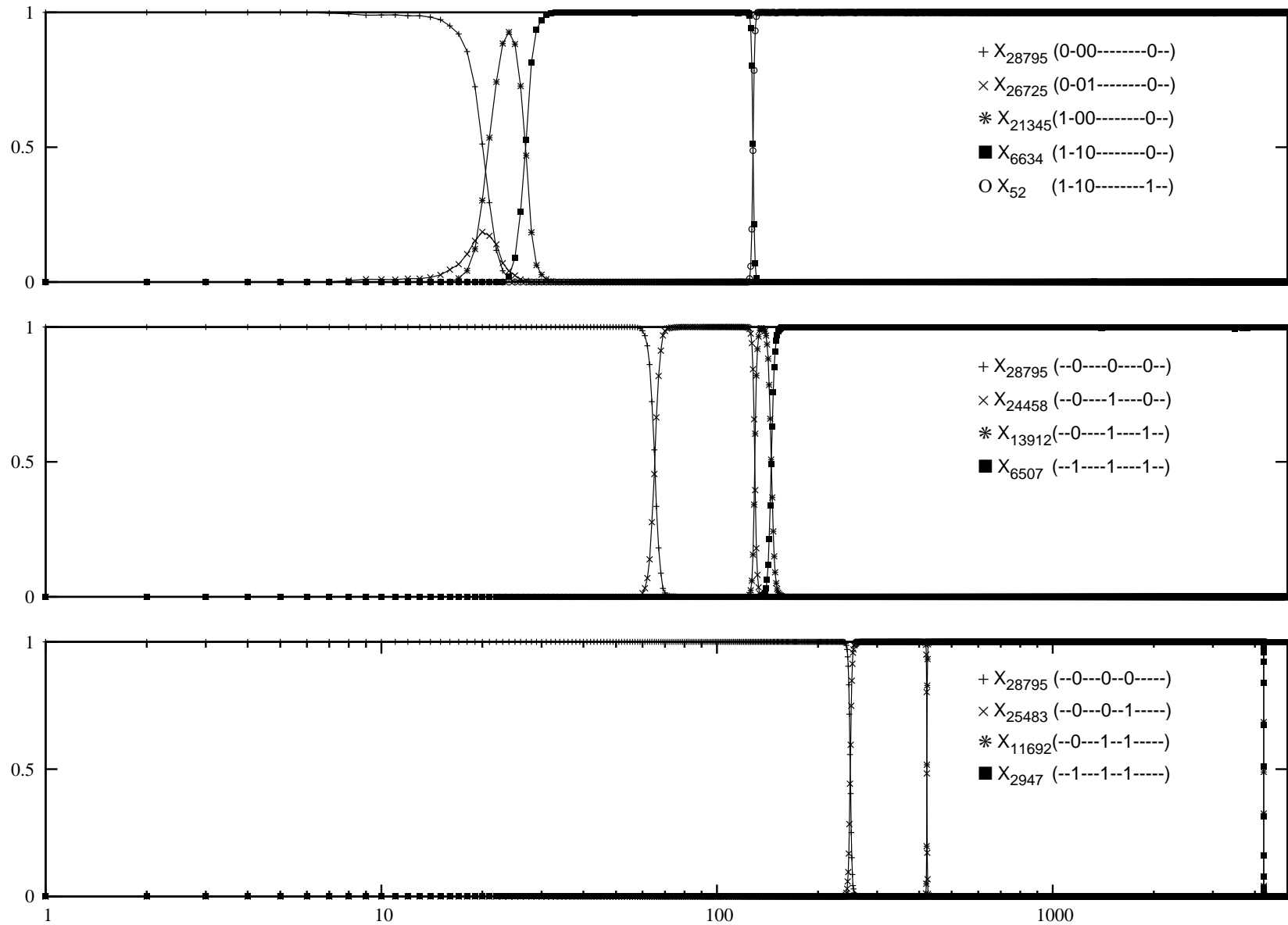
- $LN\mu$ : Number of mutants produced per generation
- $LN\mu \ll 1$ : Mutations are rare  $\Rightarrow$  population occupies a single site in sequence space and performs an uphill **adaptive walk**
- $1 \ll LN\mu \ll L$ : Stochastic regime with **interacting clones**
- $LN\mu > L$ : **Locally deterministic** evolution within a shell of size

$$d_{\text{eff}} \sim \ln N / |\ln \mu|$$

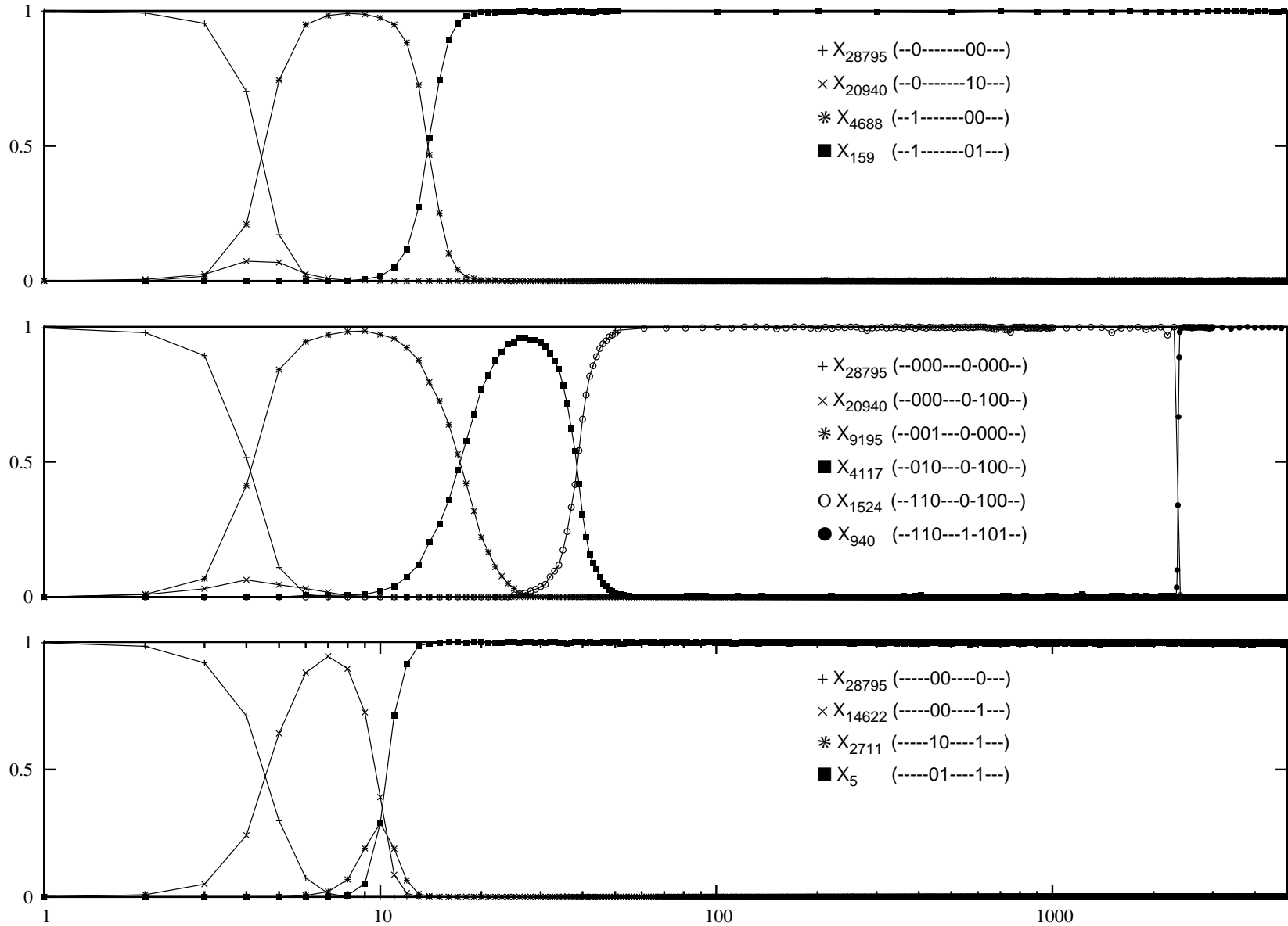
estimated from quasispecies theory

- $N \gg |\mu|^{-L}$ : Deterministic **quasispecies** dynamics K. Jain, JK, JSTAT (2005)

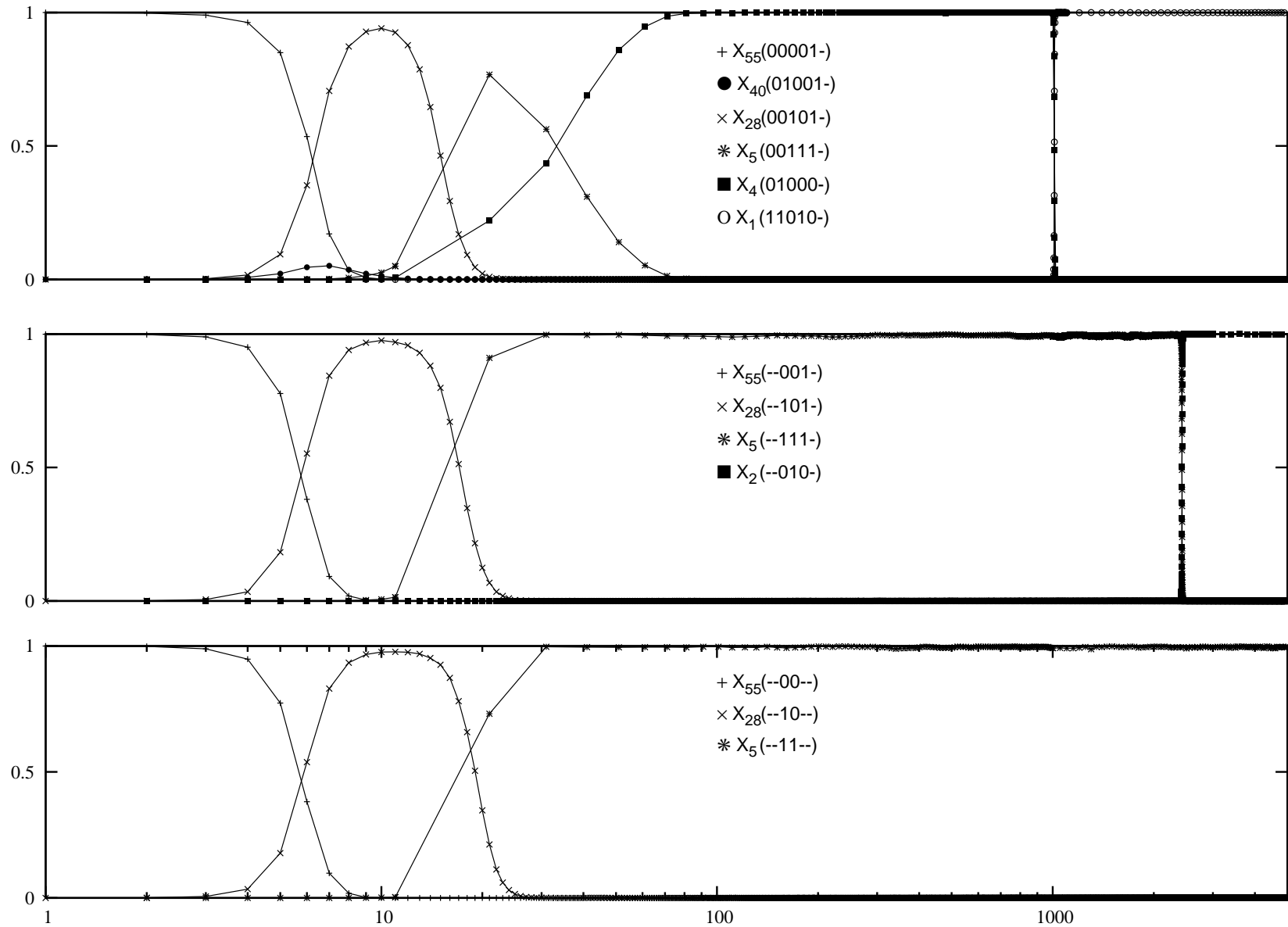
# Adaptive walk: $L = 15, N = 1024, \mu = 10^{-5}, 10^{-6}, 10^{-7}$



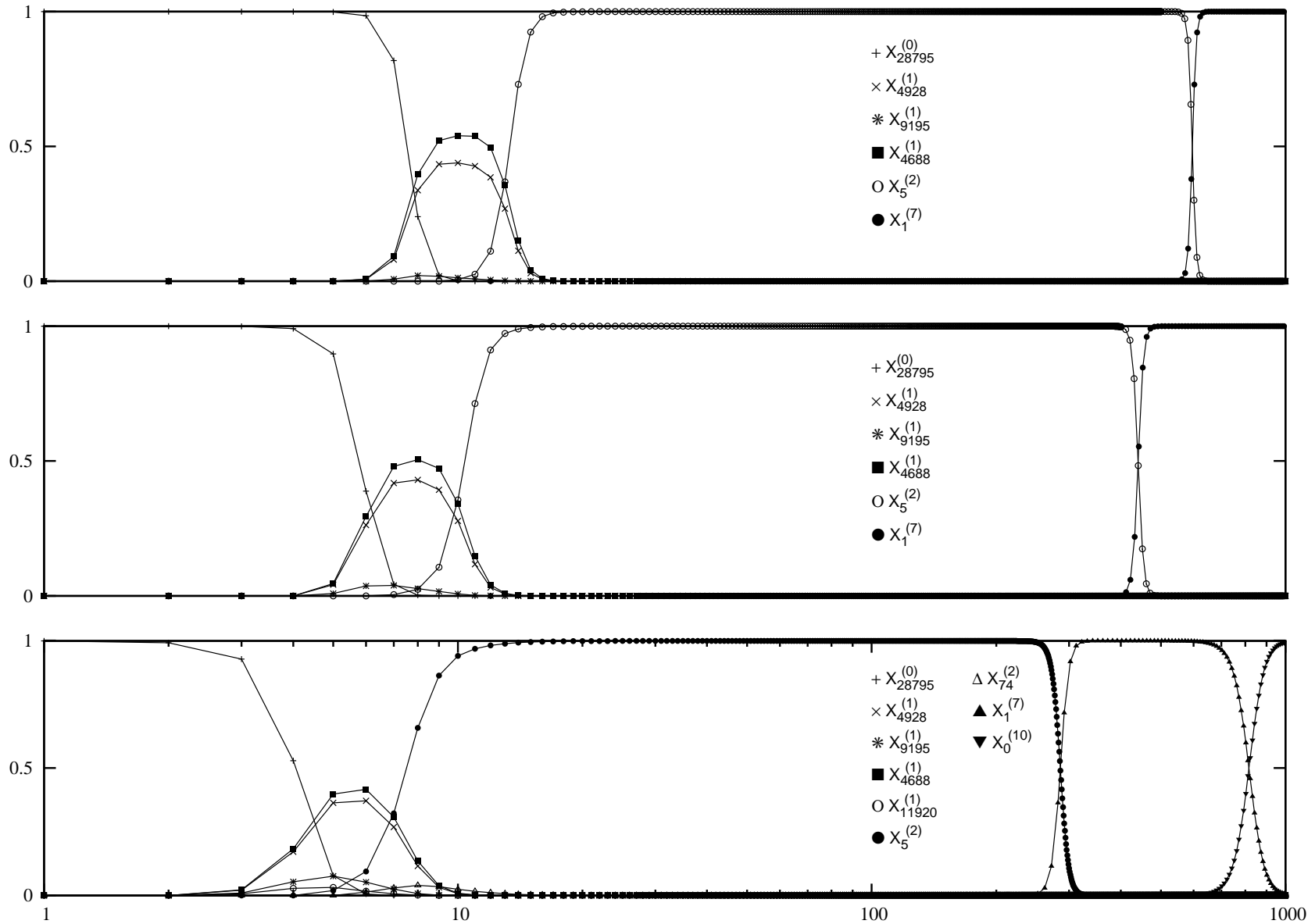
# “Clonal interference”: $L = 15, N = 1024, \mu = 10^{-4}$



# Locally deterministic evolution: $L = 6, N = 16384, \mu = 10^{-4}$



# Deterministic quasispecies dynamics: $L = 15, \mu = 10^{-8}, 10^{-6}, 10^{-4}$





# The infinite sites limit

S.C. Park, JK, JSTAT (2008) P04014

- Take  $L \rightarrow \infty$ ,  $\mu \rightarrow 0$  with fixed

$$U = 1 - (1 - \mu)^L \approx 1 - e^{-\mu L}$$

$\Rightarrow$  each mutation creates a new genotype with a fitness drawn randomly from the **mutation distribution**  $g(w)$

- In the limit  $N \rightarrow \infty$  the population fitness distribution evolves according to

$$f_{t+1}(w) = (1 - U) \frac{w f_t(w)}{\bar{w}_t} + U g(w) \quad \bar{w}_t : \text{mean fitness}$$

- Mutation-selection balance for  $g(w)$  with bounded support Kingman (1978)
- For unbounded  $g(w) \sim \exp[-(w/w_0)^\beta]$  mean fitness grows as

$$\bar{w}_t \approx C_\beta w_0 (1 - U) t^{1/\beta} \quad 1 - U : \text{mutational load}$$

## Finite populations and records

- At long times beneficial mutations are rare events:

$$U_b(t) = U \text{Prob}[w > \overline{w}_t] = U \int_{\overline{w}_t}^{\infty} dw g(w) \rightarrow 0 \text{ for } t \rightarrow \infty$$

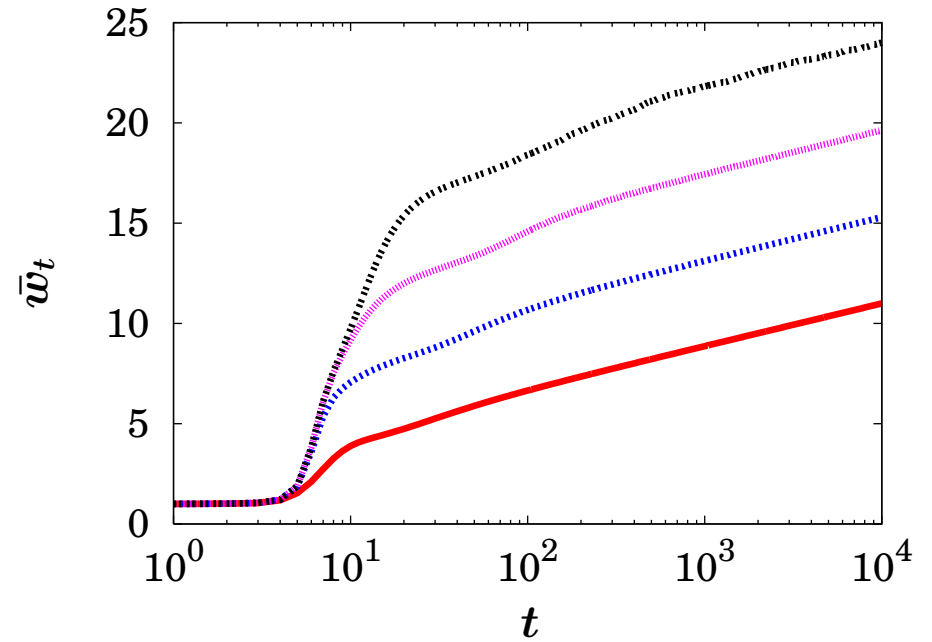
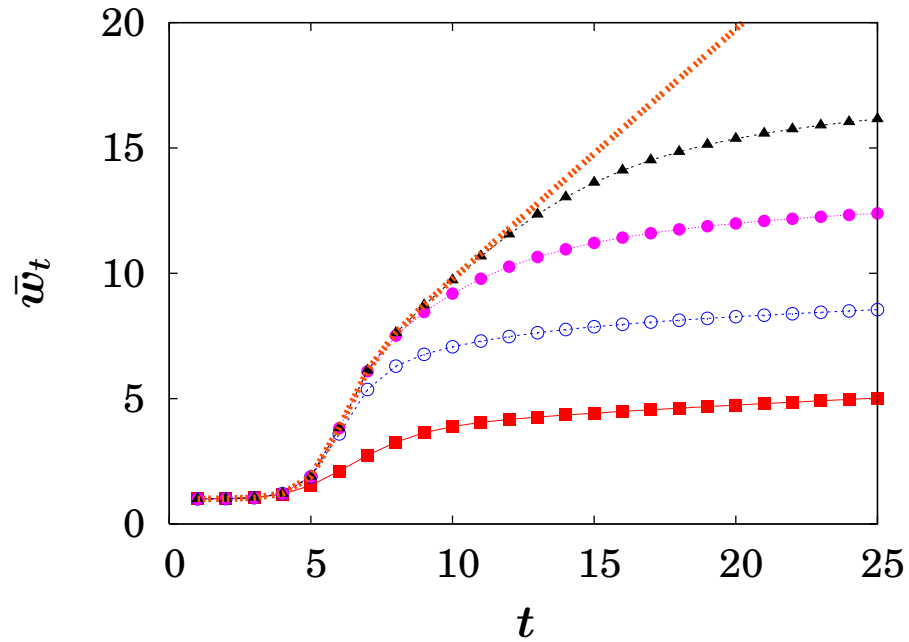
- For  $U \ll 1$  the effect of deleterious mutations can be neglected as well  
 $\Rightarrow$  approximation by a **diluted record process**  $w_t^{\text{DRP}}$ , in which mutants of fitness  $w' > w$  replace current genotype  $w$  with the **fixation probability**

$$\pi(s) = 1 - e^{-2s}, \quad s = w'/w - 1.$$

- To leading order  $\overline{w}_t^{\text{DRP}}$  is equal to the largest fitness value encountered up to time  $t$  [=standard record process], with corrections that can be systematically computed
- Deleterious mutations rescale the fitness according to

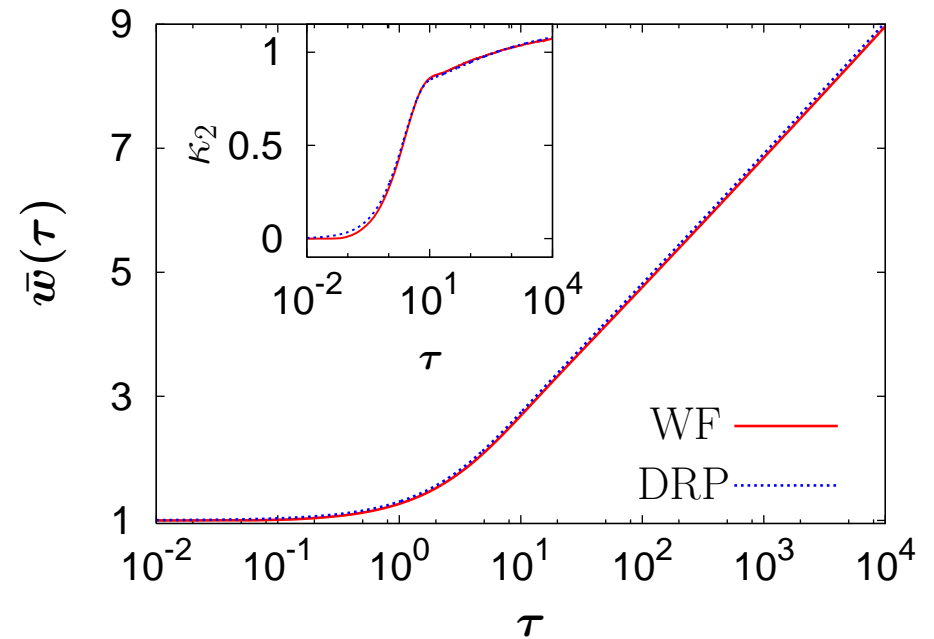
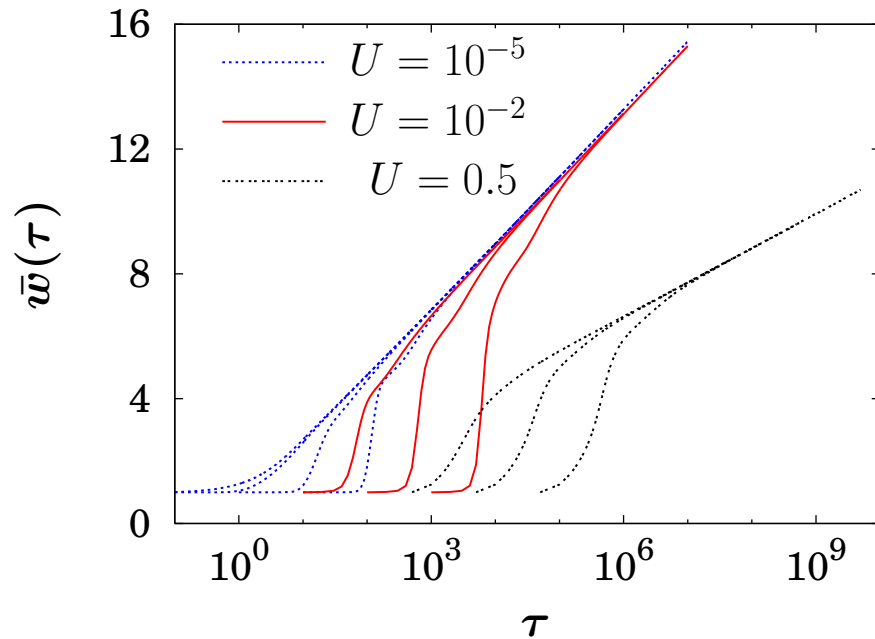
$$\overline{w}_t \approx (1 - U) \overline{w}_t^{\text{DRP}} \approx (1 - U) \ln(NUt) \text{ for } g(w) = e^{-w}$$

# Simulations: Finite vs. infinite populations



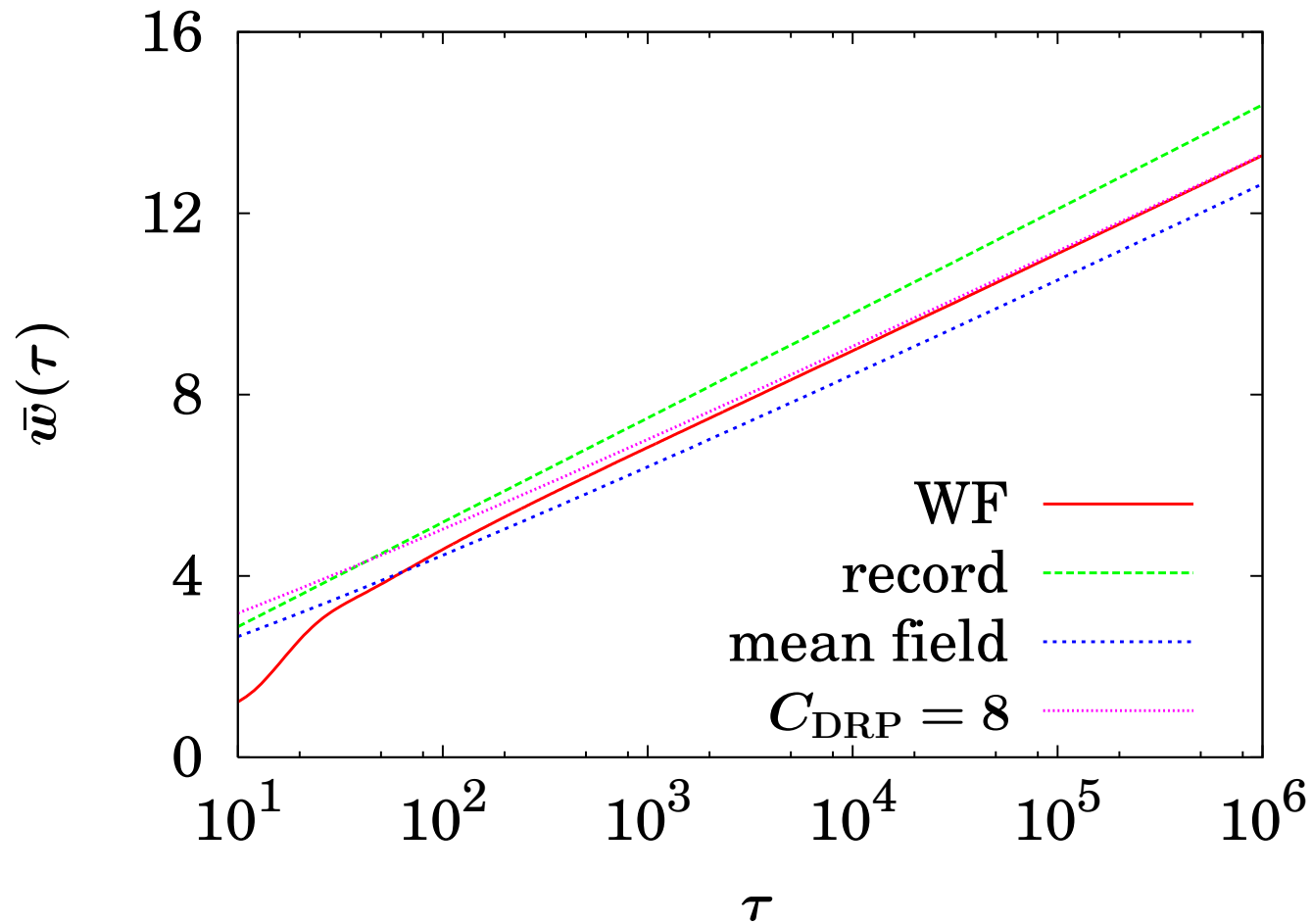
$$U = 0.01, N = 10^3, 10^5, 10^7, 10^9, \infty$$

# Comparison to the diluted record process



- scaled time  $\tau = NUt$
- fitness variance  $\kappa_2 \rightarrow \text{const.}$

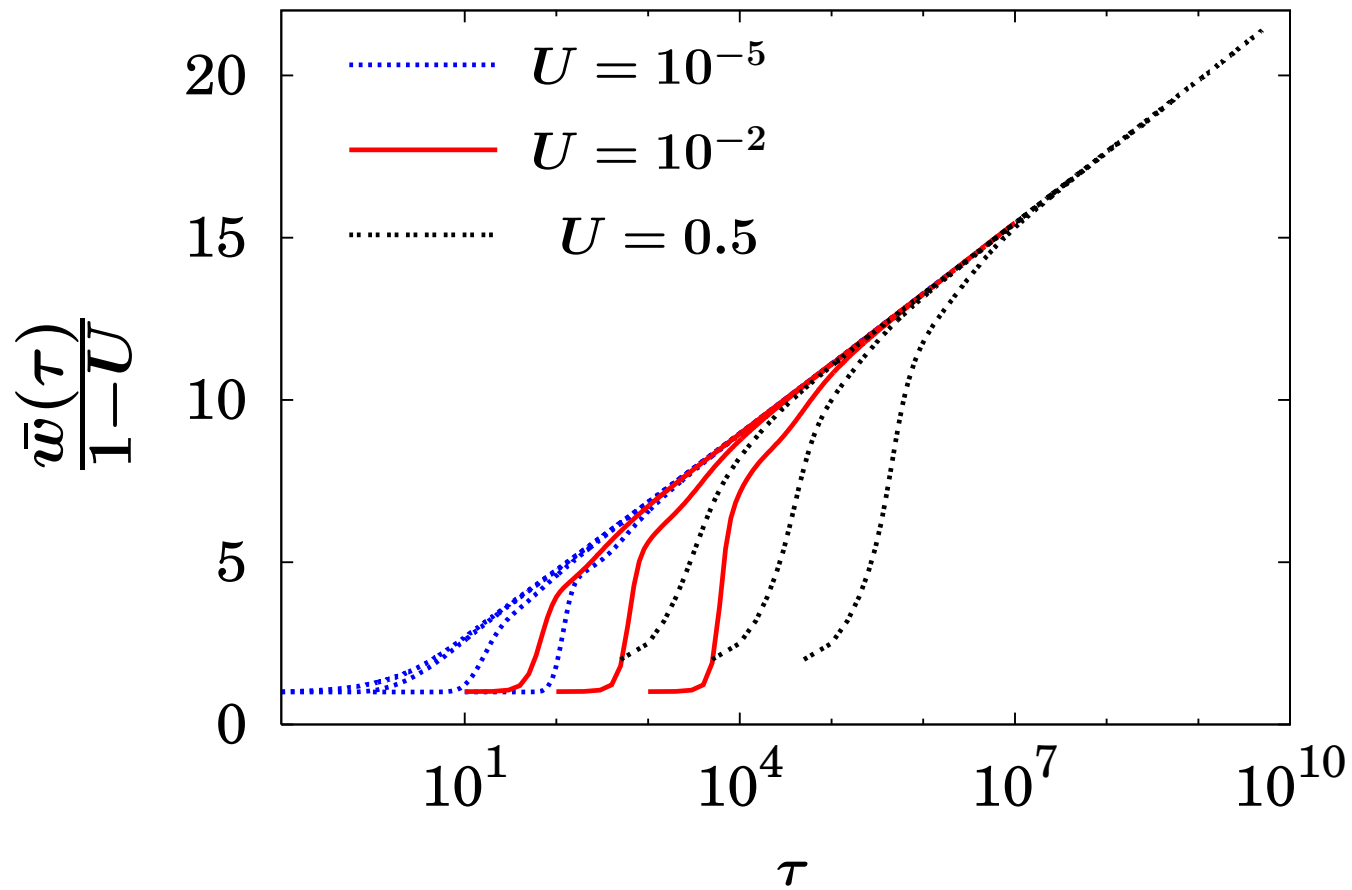
# Diluted record process: Bounds and approximations



record process: upper bound

mean field approximation: lower bound

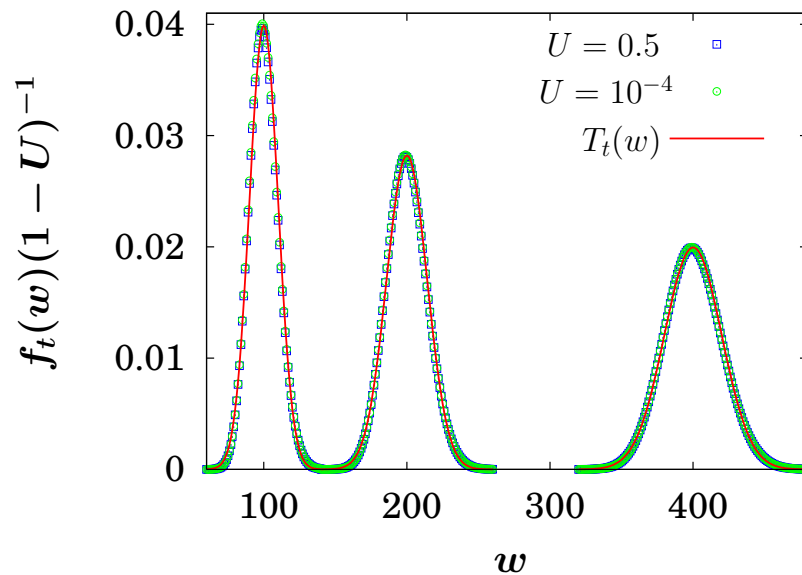
# Finite populations at arbitrary $U$



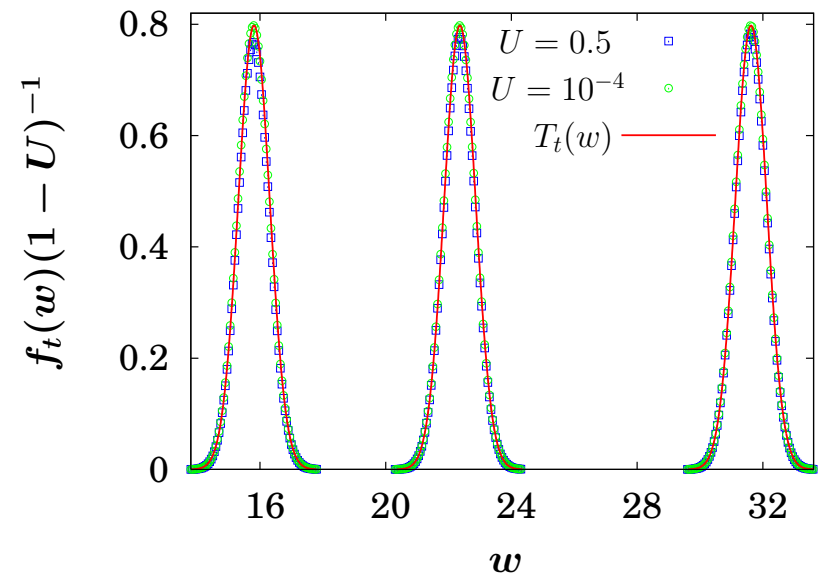
$$N = 10^3, 10^4, 10^5$$

# Bimodality of fitness distribution

exponential  $g(w)$



Gaussian  $g(w)$



- Asymptotic decomposition

$$f_t(w) \approx Ug(w) + (1-U)T_t(w)$$

with a “traveling wave” contribution  $T_t(w)$  holds for finite and infinite populations

## Summary

- Stochastic and deterministic regimes of evolution in fitness landscape with maximal ruggedness (**epistasis**)
- Two solvable limits:
  - $N \rightarrow \infty \Rightarrow$  deterministic quasispecies theory
  - $L \rightarrow \infty \Rightarrow$  infinite sites model, stochastic dynamics related to records
- Nonmonotonic dependence on mutation rate

## Future directions

- Infinite sites limit for fitness landscapes of intermediate epistasis (e.g. **NK-model**)
- Effects of sexual reproduction/recombination
- Empirical fitness landscapes from evolution experiments