Basic questions about evolution

- How complex can organisms become, in principle?
    - How can we define complexity?
      Solution: consider **channel capacity** instead
    - Are there different limits for prokaryotes and eukaryotes?
      Answer: yes, and the limit for eukaryotes is much larger than expected

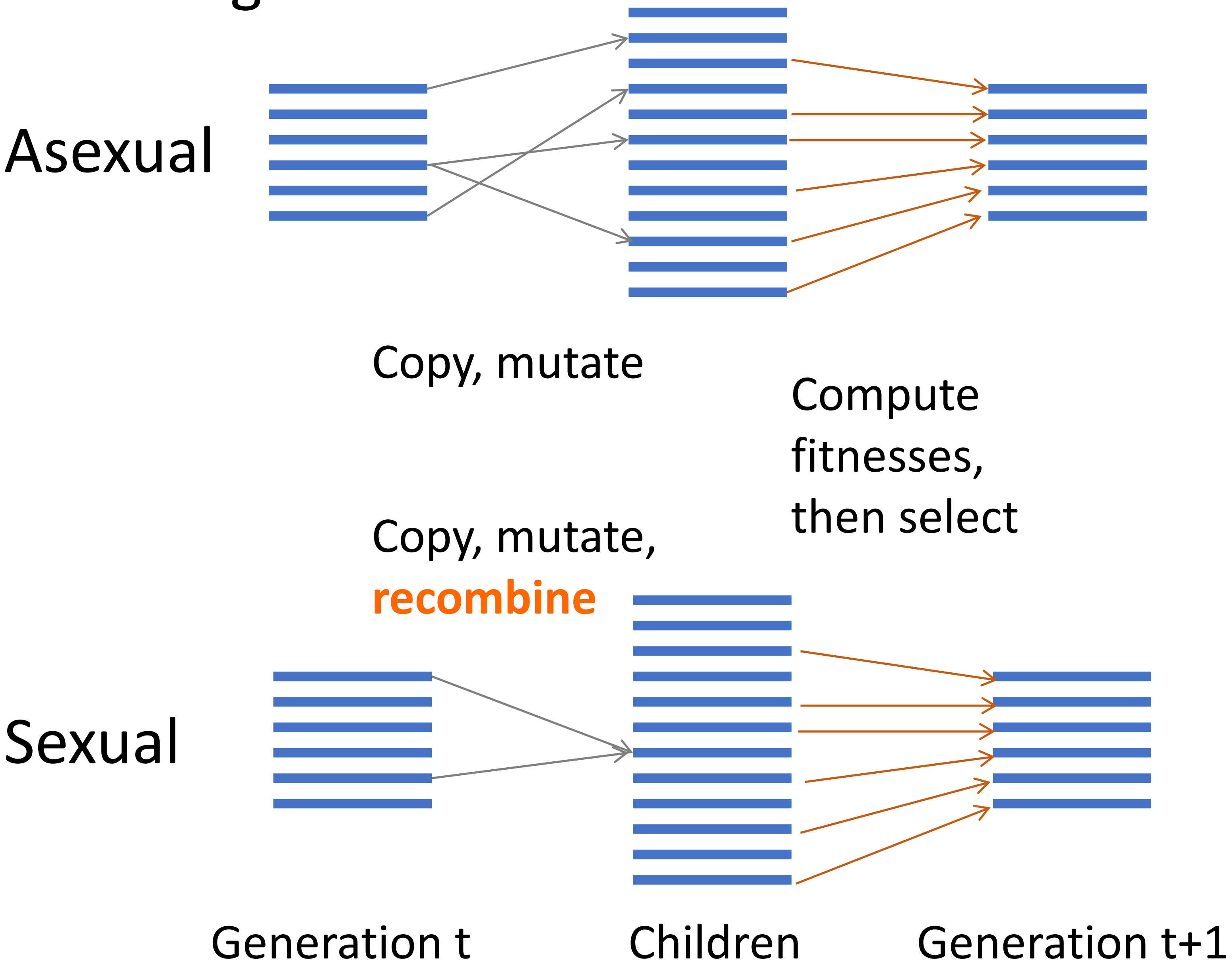
- Should we expect evolved systems to be different from designed systems?
    - An important question: will designed AI be different from our evolved intelligence?
    - Conventional answer is that evolved organisms have an evolutionary history, whereas machines can be designed de novo
    - Different answer: **Low-density code hypothesis**: greater organismal complexity achievable with genetic architectures analogous to error correcting codes

Differences between asexual and sexual organisms

Each difference is by multiple orders of magnitude

	Asexual	**Sexual**
Genome size:	Small	Large
Population size:	Large	Small
Generation time:	Short	Long
Gene regulation:	Simple	Complex
Development:	None	Complex

Genetic Algorithms



Asexual

Copy, mutate

Compute
fitnesses,
then select

Copy, mutate,
recombine

Sexual

Generation t

Children

Generation t+1

Selection by percentage match with Arbitrary Target Binary String

In the following experiments, we first choose a binary string as a ‘target’ genome

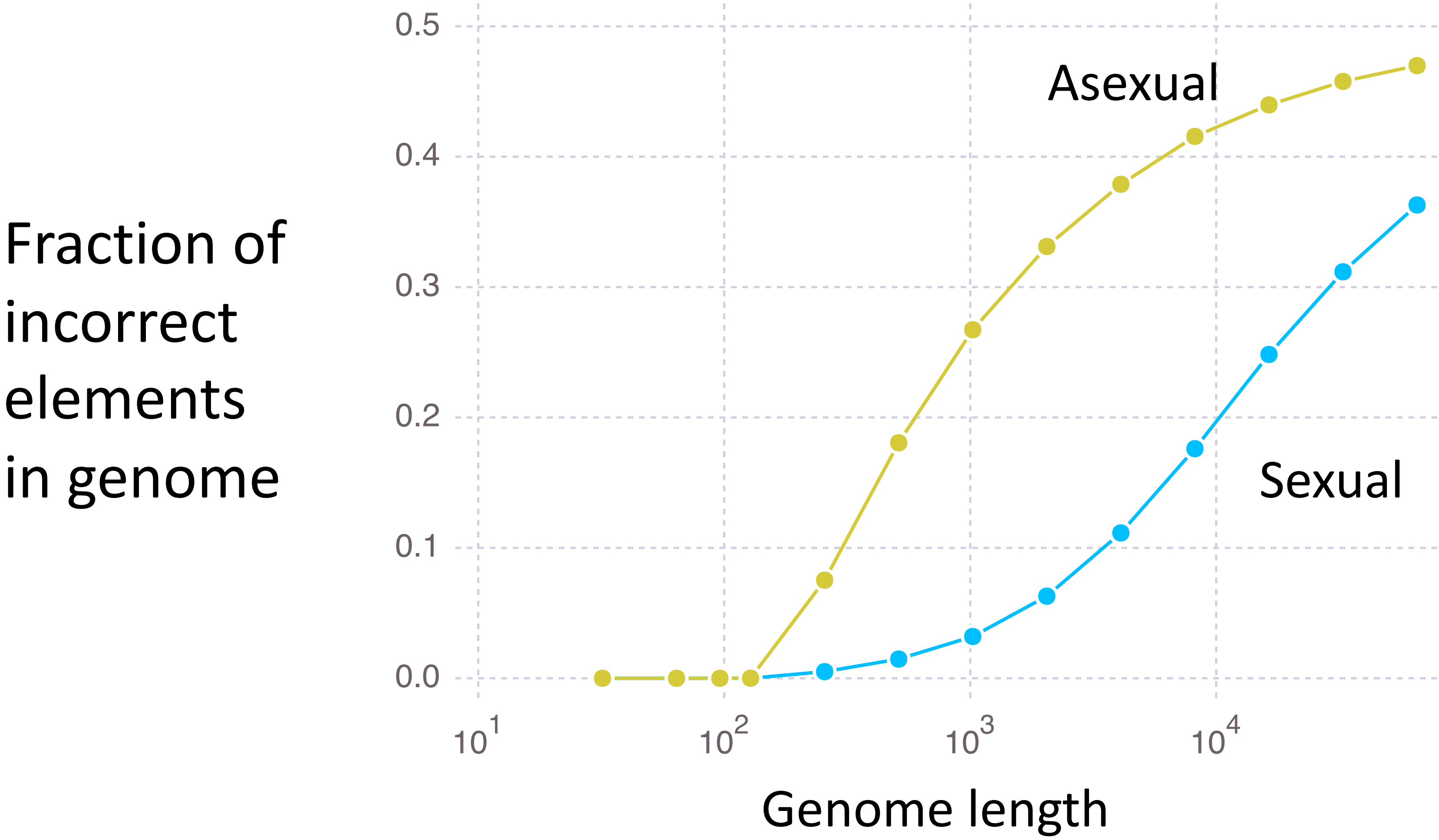
In each generation, breed 2N ‘children’, and then select the N that are closest to the target string.

Not intended to be a realistic model: it is a limitative model.

Every assumption is optimistic: in this way, we can get *upper bounds* on achievable complexity.

Bit-error-rate vs genome length

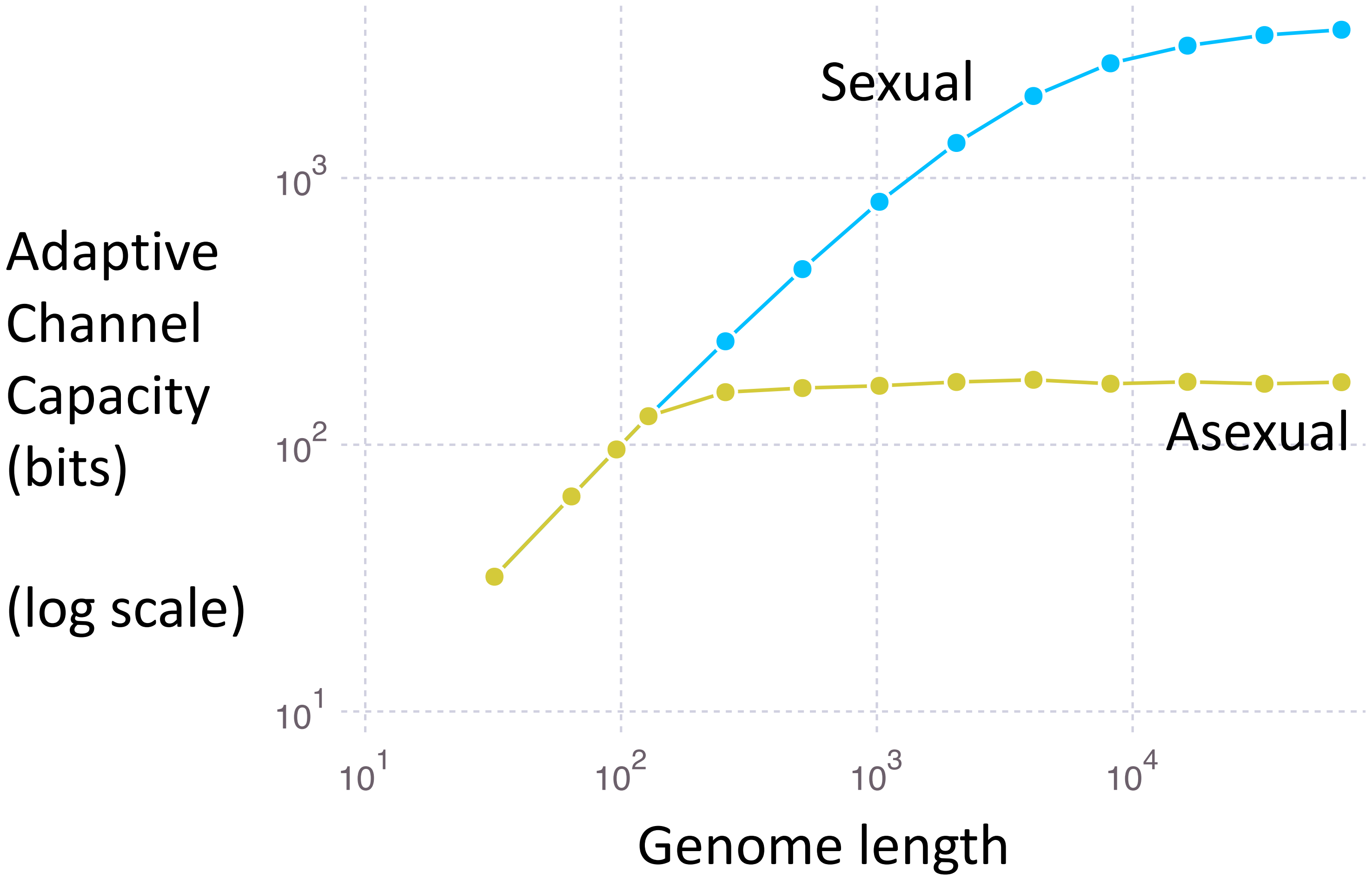
Population size: N = 100 Mutation rate: u = 0.01



Same data replotted as channel capacity vs genome length

$$\text{Channel Capacity} \geq L(1 - h_2(p))$$

where p is error rate, and $h_2(p)$ is entropy in bits



The Drosophila Game: a thought-experiment

Suppose two geneticists – A and B – wish to communicate.
First, they agree on a code-book of varieties of Drosophila

- could use binary traits: red/black eyes, long/stubby wings, smooth/hairy, etc
- could use a binary code on polymorphic SNPs

A decides on the message to send.
She then breeds a population of N Drosophila in her office; after each generation she has 2N, and she squashes half of them, and continues to breed the rest.

By keeping the flies most similar to the variety she wants, A can breed her population to have the binary trait values she desired.

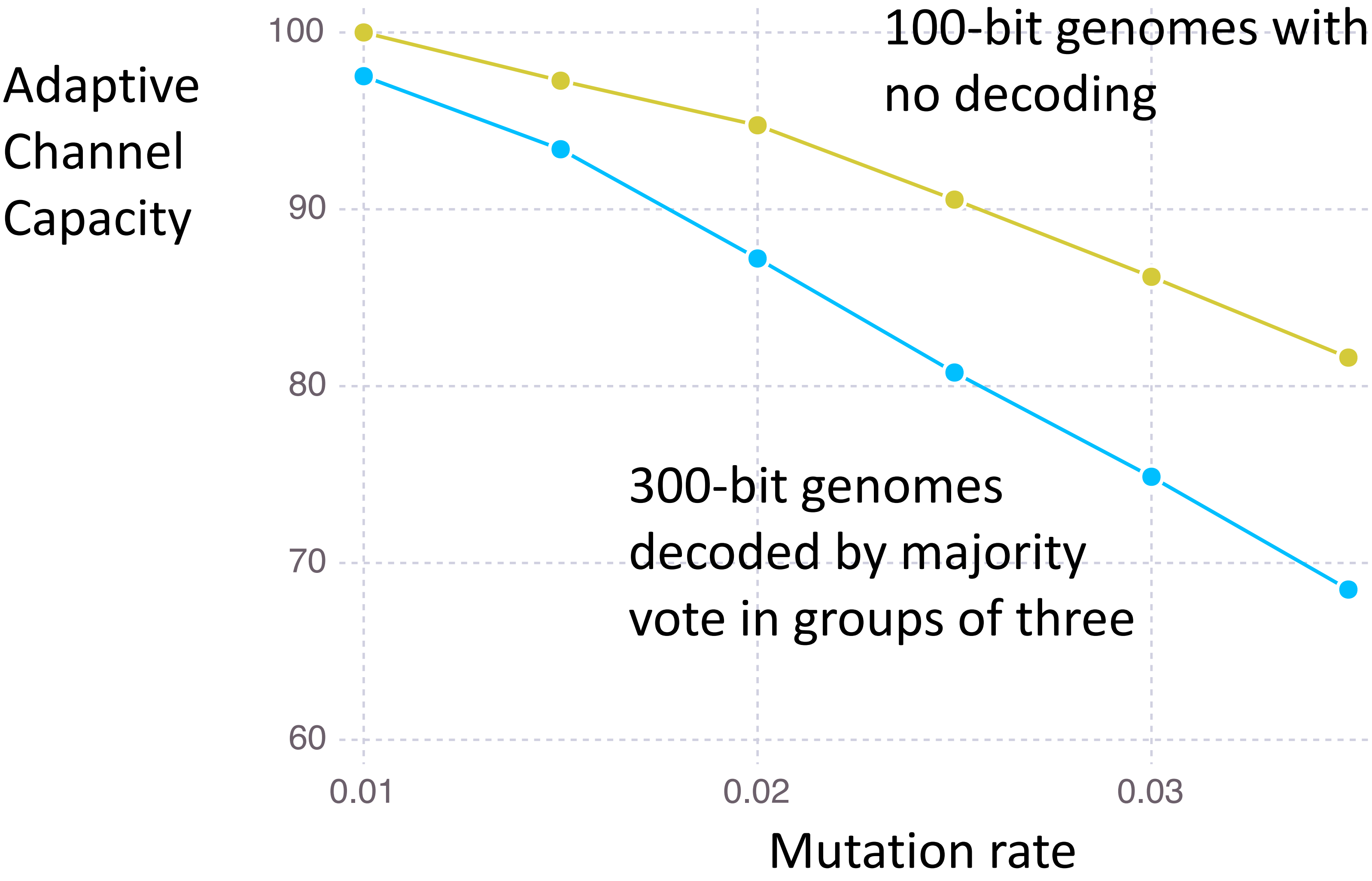
Finally A sends her message: she releases her final selected population of N flies, and B catches one.

B examines (or sequences) the fly, and determines which variety in the code-book it belongs to, and so decodes A’s message.

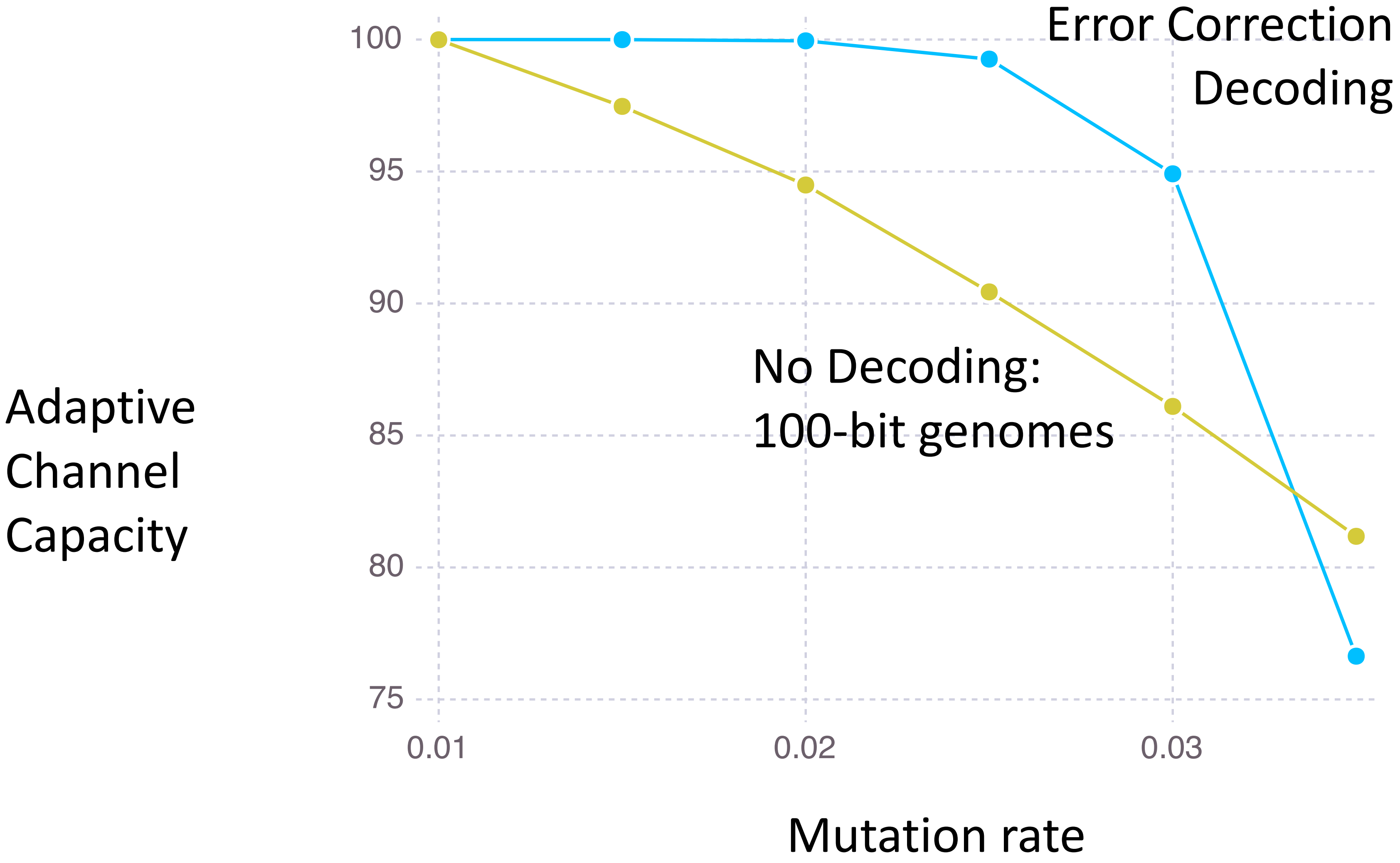
If A can reliably set, say, 10 different binary traits, she can produce 1024 different varieties of fly, and so send 10 bits.

The **channel capacity** measures how precisely the flies can adapt to A’s selective breeding

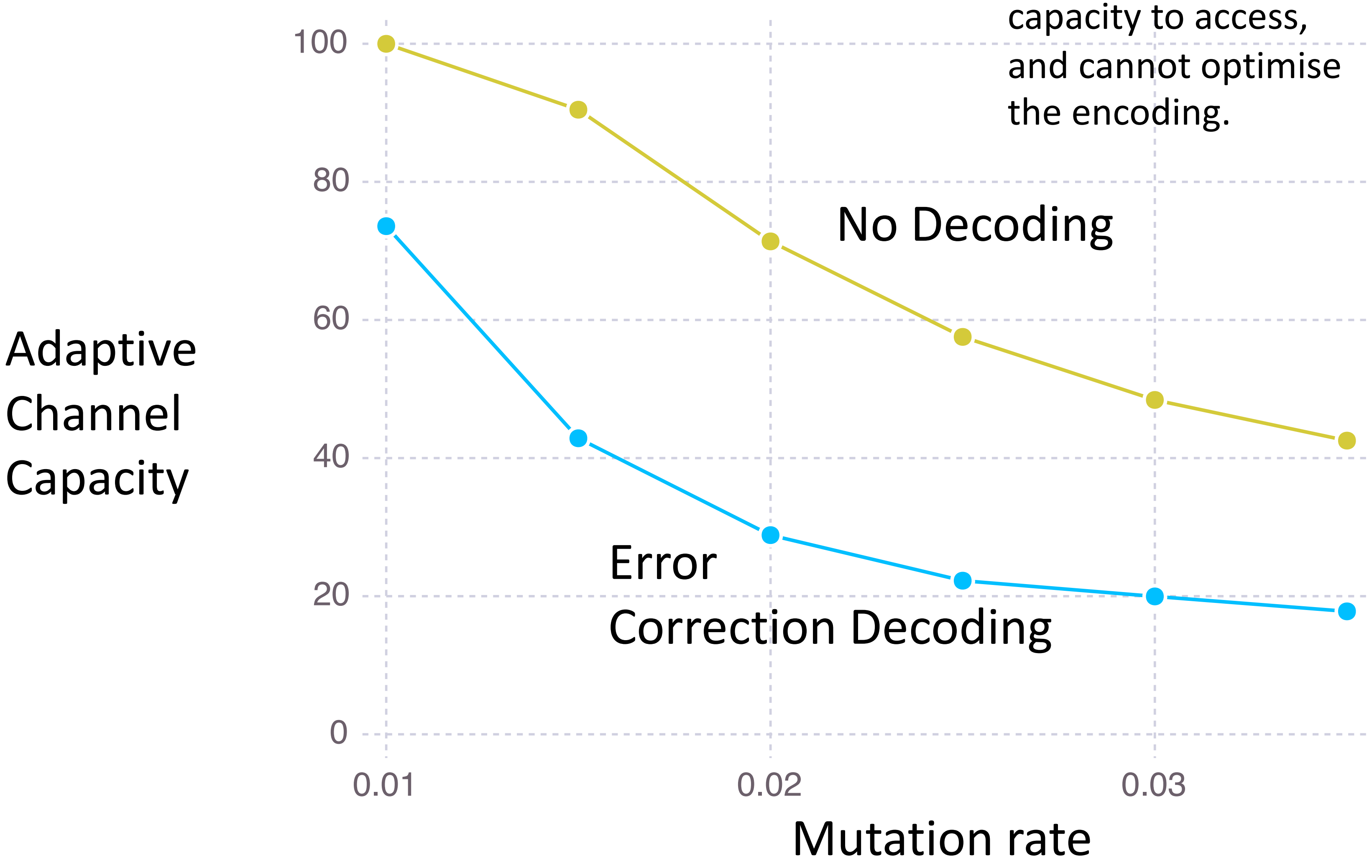
Simple redundancy: 300 bit genome, decoded to 100 bits by majority votes on 100 separate groups of 3 bits.



Non-linear decoding: 300 bit genomes, decoded to 100 bits using much-simplified LDPC decoding



LDPC decoding compared with no decoding for **Asexual GA.**



Asexual GA has no adaptive channel capacity to access, and cannot optimise the encoding.

Evolutionary roles of ‘entry-level’ learning

Situation-adaptive	Organisms born into **different environments**, and must learn from environment to succeed
Decompression	Learning decodes a compact description of behaviour, e.g. as a reward system.
Error-correction	Learning resolves inconsistencies in genetic specification of behaviour, enabling genetic encoding of behaviour of greater complexity

Decompression and error correction valuable even if all organisms in **same environment**;

EBT: Exchangeable Breeding with Tournaments

A tool for studying mutation-selection equilibrium

‘Genetic algorithm’ that satisfies detailed balance.

Generation t

Exchangeably sample one genome

Reject one genome

Generation t+1

$$g_{N+1} \sim p_B(\cdot \mid g_1, \ldots, g_N)$$

Reject genome k with probability $$\frac{\frac{1}{f_k}}{\frac{1}{f_1} + \cdots + \frac{1}{f_{N+1}}}$$

Exchangeable Sampling

“Breeding” is by sampling a genome, conditional upon existing genomes. Joint “breeding distribution” should be exchangeable:

$$p_B(g_1, \ldots, g_N) = p_B(g_{\sigma_1}, \ldots, g_{\sigma_N})$$

For a ‘Genetic Algorithm’, use a breeding distribution that is a product of Beta-Bernoulli distributions (or Dirichlet Processes).

That is, each element of a new genome is sampled from an independent Dirichlet process: the ‘genome’ consists of a vector containing a sample from each DP.

Tournament Selection by ‘loser ticket’

If genomes i and j have a ‘tournament’

$$Pr(\text{j wins tournament against i}) = \frac{f_j}{f_i + f_j}$$

Suppose there is one ‘loser ticket’. After many tournaments in which the genome currently holding the loser ticket ‘fights’ another genome, and the loser gets the ticket, limit distribution is:

$$Pr(\text{j holds loser ticket}) = \frac{\frac{1}{f_j}}{\frac{1}{f_1} + \cdots + \frac{1}{f_{N+1}}}$$

Key point: Stationary Distribution of EBT can be written in closed form

EBT is a Markov Chain of populations.

Stationary distribution is:

$$\pi(g_1, \ldots, g_N) \propto p_B(g_1, \ldots, g_N) f(g_1) \cdots f(g_N)$$

Analogous to a Bayesian posterior, with
‘breeding distribution’ as prior, and
fitness as likelihood.

Proof of detailed balance of EBT:

Let $G = \{g_1, g_2, \ldots, g_{N+1}\}$

To prove: $\pi(G_{\setminus N+1}) T(G_{\setminus N+1} \to G_{\setminus i}) = \pi(G_{\setminus i}) T(G_{\setminus i} \to G_{\setminus N+1})$

Proof:

$$\pi(G_{\setminus N+1}) T(G_{\setminus N+1} \to G_{\setminus i}) =$$

$$p_B(G_{\setminus N+1})\ f_1 \cdots f_N\ \ p_B(g_{N+1} \mid G_{\setminus N+1}) \frac{\frac{1}{f_i}}{\frac{1}{f_1} + \cdots + \frac{1}{f_{N+1}}}$$

$$= p_B(G)\ \ \frac{f_1 \cdots f_{N+1}}{f_i f_{N+1}}\ \ \frac{1}{\frac{1}{f_1} + \cdots + \frac{1}{f_{N+1}}}$$

which is symmetric between g_i and g_{N+1}.

Genetic algorithm with population size 1

Breed **two** new genomes,
discard existing genome

Select between two
new genomes.
Bad algorithm!

EBT with population size 1

Sample **one** new genome,
keep existing genome

Metropolis-Hastings
(original version)

Conclusions and Further Questions

A suggestive difference between the asexual and sexual genetic algorithms is that recombination gives orders of magnitude more adaptive channel capacity.

A small demonstration that 'adaptive channel capacity' can be accessed with suitable decoder.

Question: Is one function of the elaborate control of gene expression in eukaryotes, and – at a higher level – of individual cognitive development, to 'decode' highly distributed genetic information?

Secondly, Exchangeable Breeding with Tournaments (EBT) is an alternative formulation of GAs, that satisfies detailed balance.

The mechanisms of non-parametric Bayesian MCMC using priors based on Dirichlet distributions can be plausibly interpreted as 'genetic algorithms'.

Links GAs and MCMC, and enables integrated probability models of evolution and individual learning.

50 million years ago, dog-sized quadrupeds living near shallow streams, which sometimes dried up. Not good swimmers.

Only 10 million generations

Population size ~ 1 million

So 10^{13} individuals in total

Within this small number of generations, they developed spectacular adaptations…