Ancestral selection graph meets lookdown construction

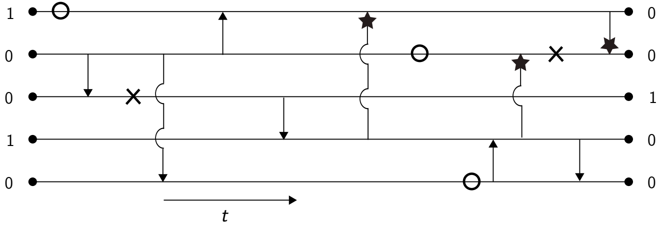
Ellen Baake

Bielefeld University

joint work with Ute Lenz, Sandra Kluth, and Anton Wakolbinger

PROBABILISTIC STRUCTURES
IN EVOLUTION
DFG SPP 1590

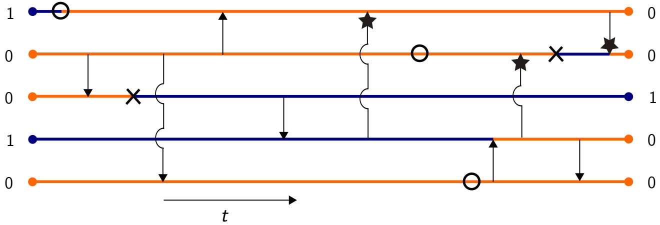
Moran model with mutation and selection



N individuals, types 0 ('good'), 1 ('bad')

- ⟶ neutral reproduction, rate 1 (for all individuals)
- ⟶★ selective reproduction, rate s (for 0 individuals)
- ○ mutation to 0, rate $u\nu_0$
- × mutation to 1, rate $u\nu_1$ $\quad(\nu_0 + \nu_1 = 1)$

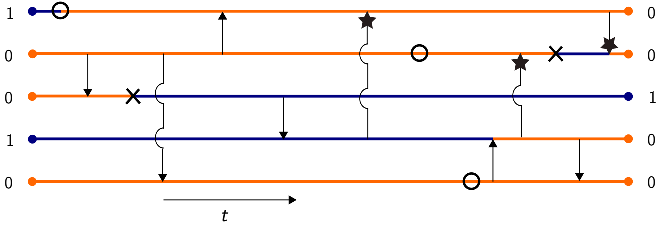
Moran model with mutation and selection



N individuals, types 0 ('good'), 1 ('bad')

- ⟶ neutral reproduction, rate 1 (for all individuals)
- ⟶★ selective reproduction, rate s (for 0 individuals)
- ○ mutation to 0, rate $u\nu_0$
- × mutation to 1, rate $u\nu_1$ $(\nu_0 + \nu_1 = 1)$

Moran model with mutation and selection



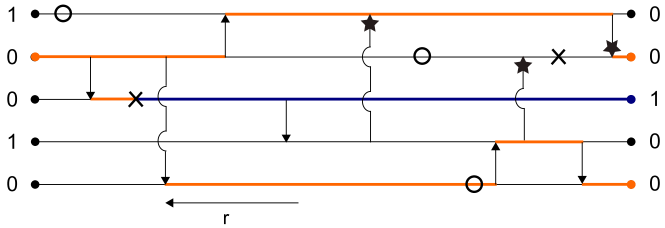
N individuals, types 0 ('good'), 1 ('bad')

- $\longrightarrow$ neutral reproduction, rate 1 (for all individuals)
- $\longrightarrow\bigstar$ selective reproduction, rate s (for 0 individuals)
- ○ mutation to 0, rate $u\nu_0$
- × mutation to 1, rate $u\nu_1$ $\quad (\nu_0 + \nu_1 = 1)$

X_t frequency of type-0 individuals at time t

diffusion limit: $t \to t/N$, $N \to \infty$ s.t. $Ns \to \sigma$, $Nu \to \theta$

Looking back



ancestors? genealogy? MRCA?

History

1981 neutral case ($\sigma = 0$): Kingman's coalescent;
genealogy independent of types

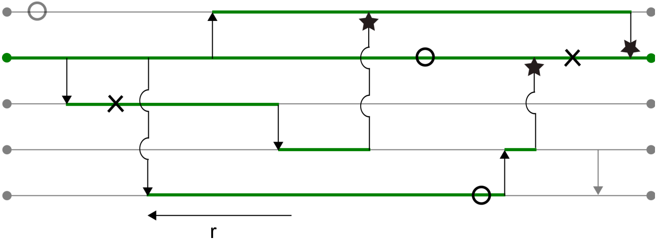
–1996 coalescent with selection ($\sigma > 0$):
considered impossible due to principal reasons

1997 Krone and Neuhauser, ancestral selection graph (ASG)

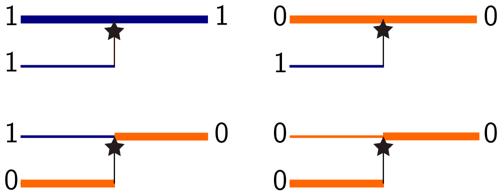
1999 Donnelly and Kurtz, lookdown construction (LD)
with selection

Ancestral selection graph: Basic idea

follow back all potential ancestors.....

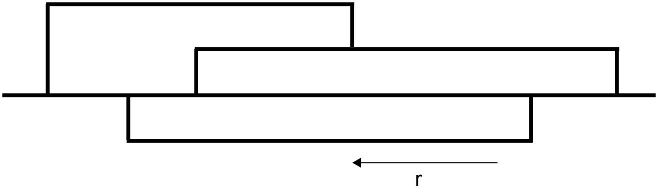


.... and keep in mind pecking order



Ancestral selection graph: Basic idea

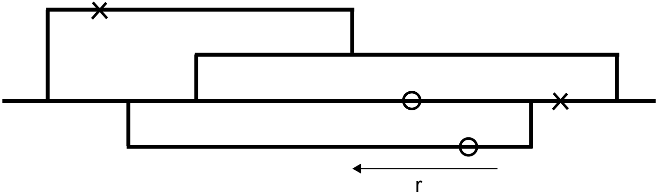
step 1: backward, w/o types



(branching: rate σ per line; coalescence: rate 1 per ordered pair)

Ancestral selection graph: Basic idea

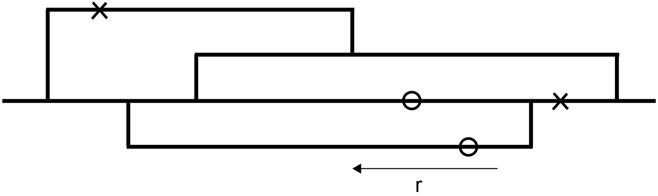
step 1: backward, w/o types



(mutation: rates $\theta\nu_1, \theta\nu_0$ per line)

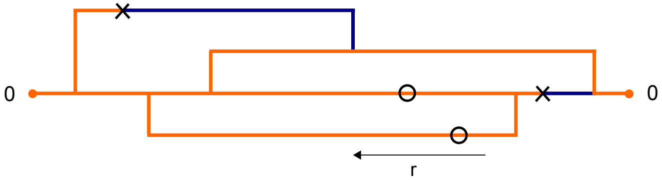
Ancestral selection graph: Basic idea

step 1: backward, w/o types



(mutation: rates $\theta\nu_1, \theta\nu_0$ per line)

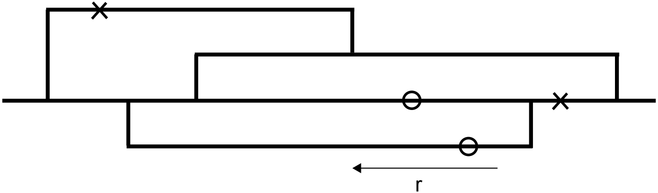
step 2: forward, with types (assigned at $t = 0$ according to X_0)



(resolve branching events)

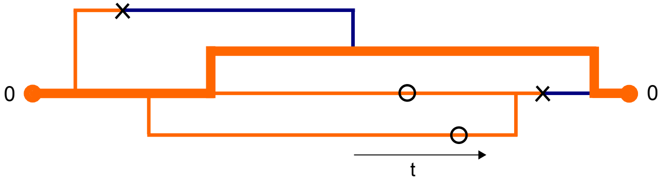
Ancestral selection graph: Basic idea

step 1: backward, w/o types



(mutation: rates $\theta\nu_1, \theta\nu_0$ per line)

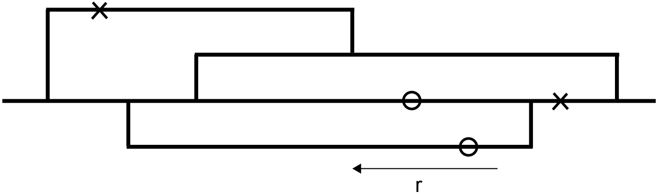
step 2: forward, with types (assigned at $t = 0$ according to X_0)



(identify ancestral line)

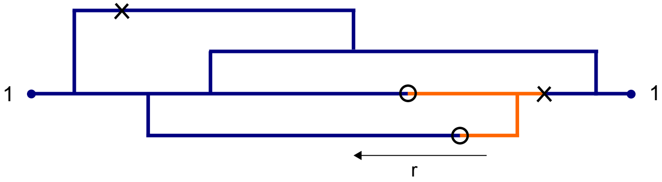
Ancestral selection graph: Basic idea

step 1: backward, w/o types



(mutation: rates $\theta\nu_1, \theta\nu_0$ per line)

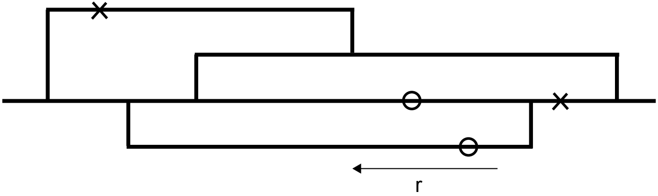
step 2: forward, with types (assigned at $t = 0$ according to X_0)



(resolve branching events)

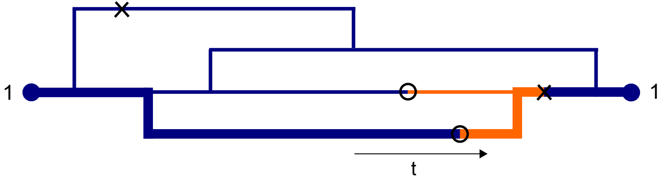
Ancestral selection graph: Basic idea

step 1: backward, w/o types



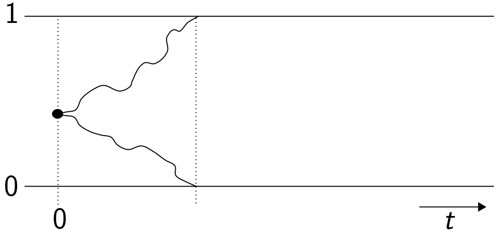
(mutation: rates $\theta\nu_1, \theta\nu_0$ per line)

step 2: forward, with types (assigned at $t = 0$ according to X_0)

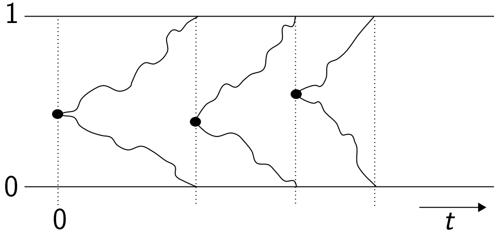


(identify ancestral line)

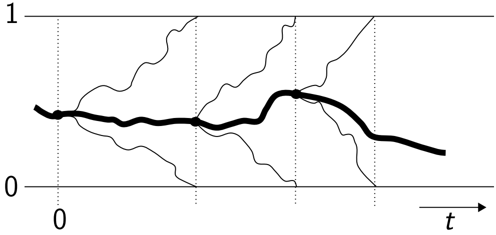
Immortal line



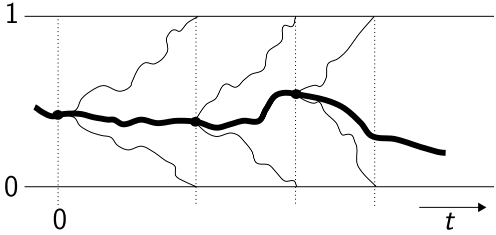
Immortal line



Immortal line



Immortal line



$$h(x) := \mathbb{P}(\text{'winner' at } t=0 \text{ is of type } 0 \mid X_0 = x)$$

$$= \sum_{n \geq 0} a_n x (1-x)^n$$ (Fearnhead 2002, Taylor 2007)

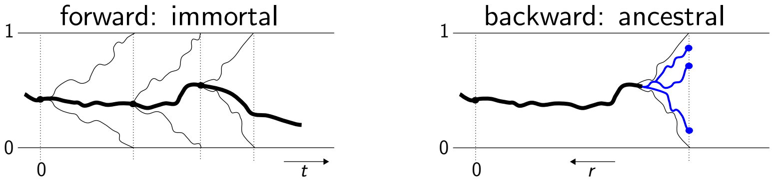
$$(n+1+\theta+\sigma) a_n = (n+1+\theta\nu_1) a_{n+1} + \sigma a_{n-1}$$

$$1 = a_0 \geq a_1 \geq \cdots, \qquad \lim_{n\to\infty} a_{n+1}/a_n = 0$$

bias towards type 0 !

probabilistic meaning, graphical approach ???

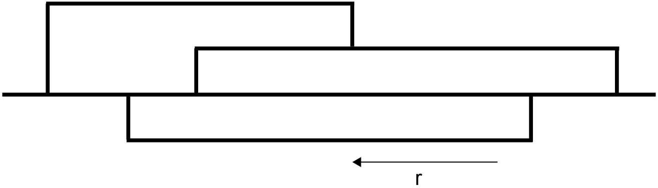
Immortal and ancestral line



immortal line:

- each ancestral line coalesces into it
- ancestral line of an individual sampled at a late time
- back to ASG

Number of lines in ASG



K_r number of lines in ASG at time $r = -t$
$(K_r)_{r \geq 0}$ birth-death process with rates

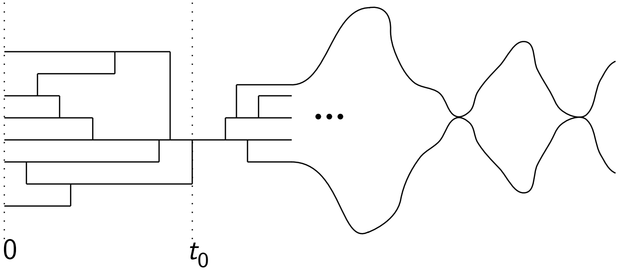
$$q_K(n, n+1) = n\sigma, \quad q_K(n, n-1) = n(n-1), \qquad n = 1, 2 \dots$$

and (reversible) equilibrium distribution $(r \to \infty)$

$$\mathbb{P}(K = n) = \frac{\sigma^n}{n!(\exp(\sigma) - 1)}, \quad n \in \mathbb{N}$$

(i.e., Poisson(σ) conditioned to $\{1, 2, \dots\}$)

Ancestral line

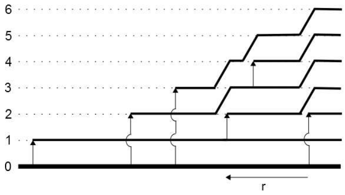


- $(K_r)_{r \geq 0}$ has bottlenecks
- identify true ancestor of first bottleneck individual
- assign types to K_0 lines (stationary!) at $t = 0$ (by drawing iid according to X_0)
- propagate types, apply pecking order (confusing!)

$\rightsquigarrow$ bring some order into the picture!

Lookdown construction ($\sigma = 0$)

Donnelly & Kurtz 1999:



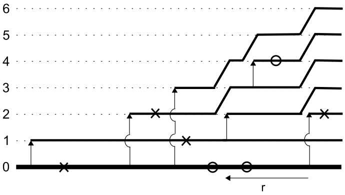
lines on levels $\in \mathbb{N}_{\geq 0}$:

- immortal line on level 0 at all times
- exchangeability preserved
  $\rightsquigarrow$ assign types iid according to X_0

$$h(x) = \mathbb{P}\left(\text{level 0 is 'good'} \mid X_0 = x\right) = x$$

Lookdown construction ($\sigma = 0$)

Donnelly & Kurtz 1999:



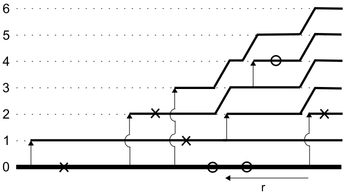
lines on levels $\in \mathbb{N}_{\geq 0}$:

- immortal line on level 0 at all times
- exchangeability preserved $\rightsquigarrow$ assign types iid according to X_0

$$h(x) = \mathbb{P}\,(\text{level 0 is 'good'} \mid X_0 = x) = x$$

Lookdown construction ($\sigma = 0$)

Donnelly & Kurtz 1999:



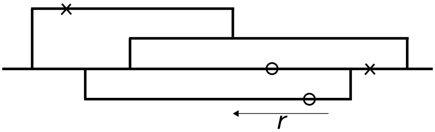
lines on levels $\in \mathbb{N}_{\geq 0}$:

- immortal line on level 0 at all times
- exchangeability preserved
  $\rightsquigarrow$ assign types iid according to X_0

$$h(x) = \mathbb{P}\,(\text{level 0 is 'good'} \mid X_0 = x) = x$$

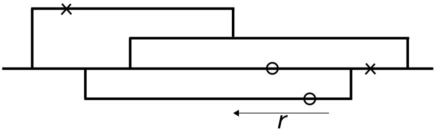
Lookdown-like representation of ASG ??

Ordering the ASG



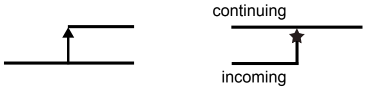
ASG

Ordering the ASG



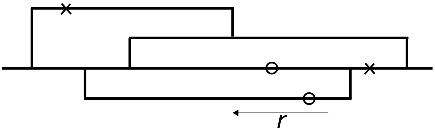
ASG

ordering convention:



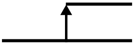
coalescence branching

Ordering the ASG

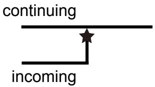


ASG

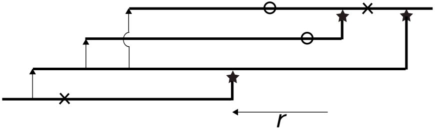
ordering convention:



coalescence



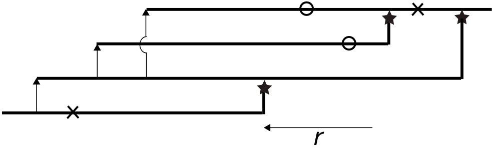
branching



ordered ASG

The Lookdown ASG

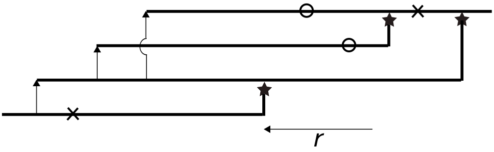
1. Construction from a given realisation of the ordered ASG



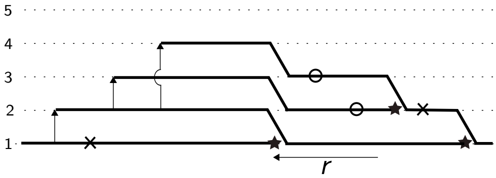
ordered ASG

The Lookdown ASG

1. Construction from a given realisation of the ordered ASG



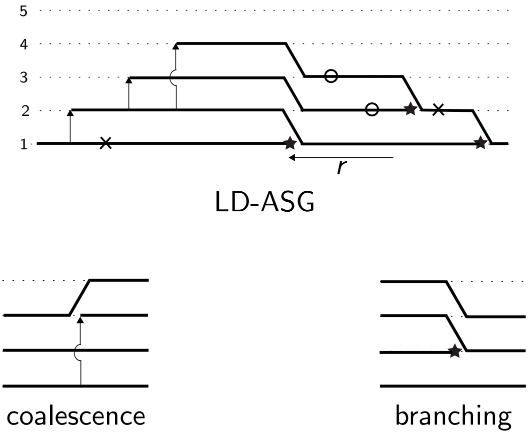
ordered ASG



LD-ASG

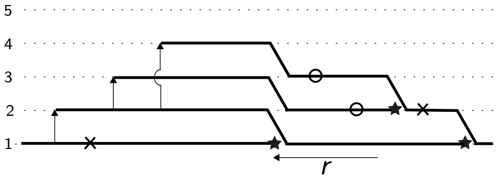
The Lookdown ASG

1. Construction from a given realisation of the ordered ASG



The Lookdown ASG

2. Markovian dynamics backward in time



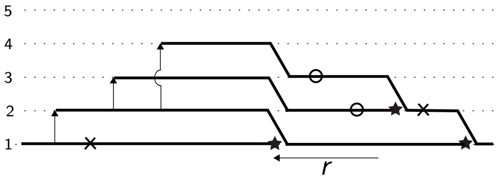
LD-ASG

independent Poisson point processes:

- branching: stars at rate σ on every occupied level
- coalescence: arrows at rate 2 for every ordered pair (i, j) of occupied levels $(j > i)$
- mutation: circles and crosses at rates $\theta\nu_0$ and $\theta\nu_1$

The Lookdown ASG

2. Markovian dynamics backward in time



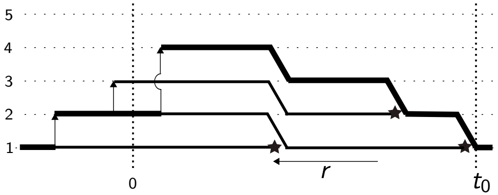
LD-ASG

K_r highest occupied level (= number of lines) at time r
($(K_r)_{r \in \mathbb{R}}$ birth-death process)

The immune line

Definition

At any given time, the immune line is the line that will be immortal if all lines at that time are of type 1.



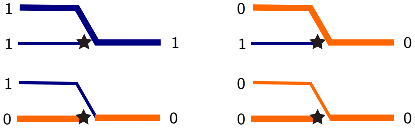
for $\theta = 0$:

- starts at bottleneck
- moves up at branching events $\rightsquigarrow$ follows continuing branch!
- follows coalescence events downwards

LD-ASG with types ($\theta = 0$)

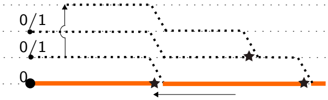
assign types (at $t = 0$ iid according to X_0)

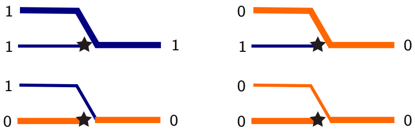
$\rightsquigarrow$ type and level of immortal line?



LD-ASG with types ($\theta = 0$)

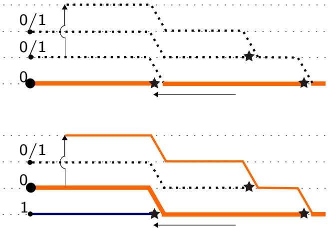
assign types (at $t = 0$ iid according to X_0)
$\rightsquigarrow$ type and level of immortal line?

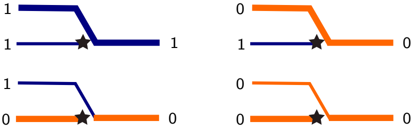




LD-ASG with types ($\theta = 0$)

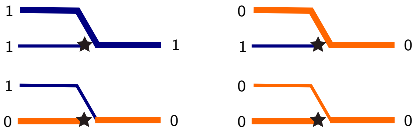
assign types (at $t = 0$ iid according to X_0)
$\rightsquigarrow$ type and level of immortal line?

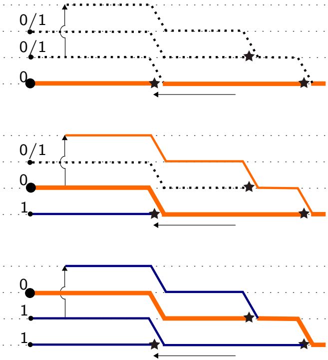




LD-ASG with types ($\theta = 0$)

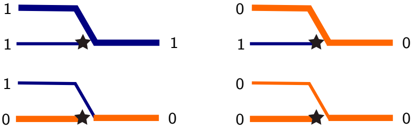
assign types (at $t = 0$ iid according to X_0)
$\rightsquigarrow$ type and level of immortal line?

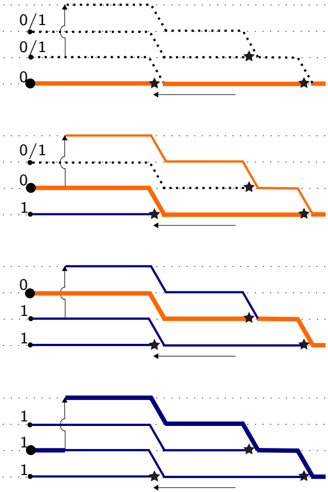




LD-ASG with types ($\theta = 0$)

assign types (at $t = 0$ iid according to X_0)
$\rightsquigarrow$ type and level of immortal line?





Level and type of immortal line ($\theta = 0$)

Proposition (Level and type of immortal line)

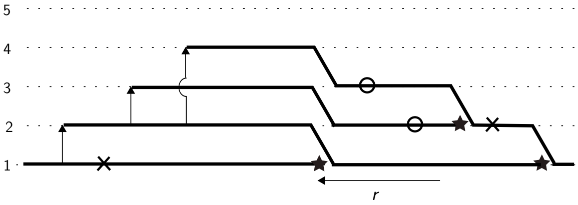
For $\theta = 0$, the level of the immortal line in the equilibrium LD-ASG with types assigned at $t = 0$ is either the lowest type-0 level or, if all lines at $t = 0$ are of type 1, it is the level of the immune line. Therefore,

$$\begin{aligned} h(x) &= \mathbb{P}(\text{immortal line has type 0 at } t = 0 \mid X_0 = x) \\ &= \sum_{n \geq 1} \mathbb{P}(K_0 \geq n)(1-x)^{n-1} x \\ &= \frac{1-\exp(-\sigma x)}{1-\exp(-\sigma)}. \qquad \text{(Kimura 1962)} \end{aligned}$$

$\theta > 0$?????

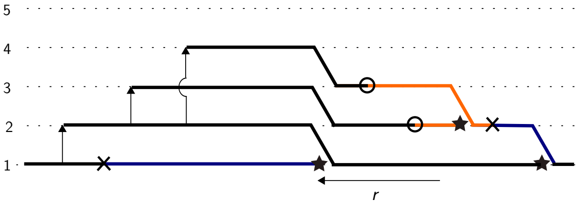
LD-ASG with mutations

LD-ASG:



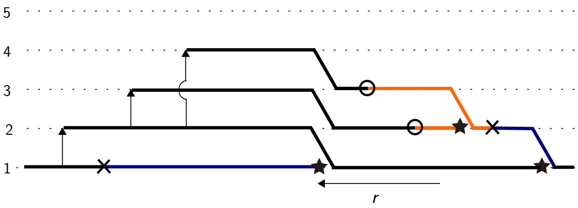
LD-ASG with mutations

LD-ASG:

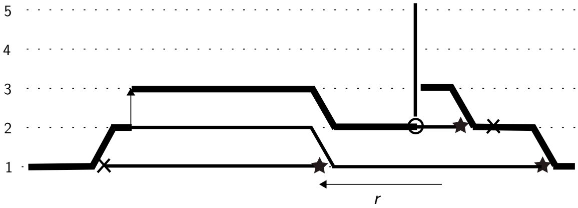


LD-ASG with mutations

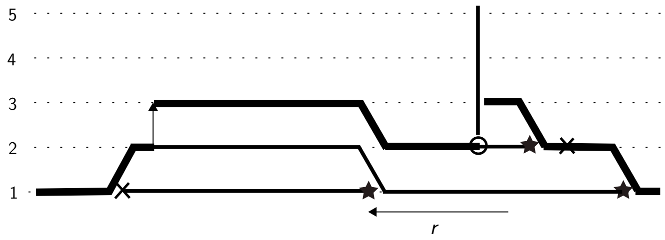
LD-ASG:

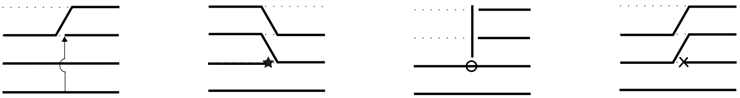


pruned LD-ASG:



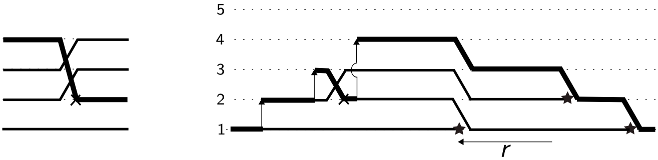
The pruned LD-ASG





immune line: jumps to levels of circles

The pruned LD-ASG

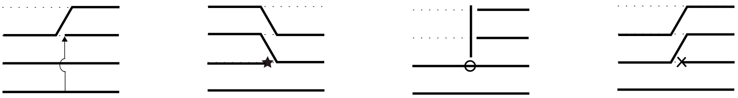


relocation to top on deleterious mutation

The pruned LD-ASG

L_r = highest occupied level (= number of lines) at time r

$(L_r)_{r\in\mathbb{R}}$ Markov chain in continuous time with rates



$q_L^{\uparrow}(n,n-1) = n(n-1)$ $\quad$ $q_L^{*}(n,n+1) = n\sigma$ $\quad$ $q_L^{\circ}(n,n-\ell) = \theta\nu_0$ $\quad$ $q_L^{\times}(n,n-1) = (n-1)\theta\nu_1$

stationary distribution ($r\to\infty$):

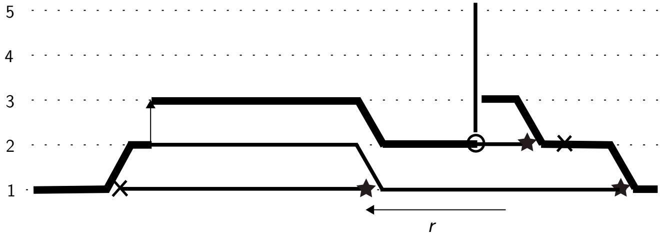
$$\rho_n := \mathbb{P}(L=n), \quad a_n := \mathbb{P}(L>n), \quad n=0,1,2,\dots$$

given via recursion (Fearnhead/Taylor)

$$(n+1+\theta+\sigma)a_n = (n+1+\theta\nu_1)a_{n+1} + \sigma a_{n-1},$$

$$a_0 = 1, \quad \lim_{n\to\infty} a_{n+1}/a_n = 0.$$

The pruned LD-ASG with types

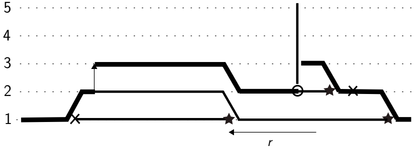


assign types: type and level of immortal line?

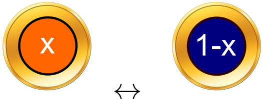
all lines untyped (except immune line),
and arranged according to pecking order

$\rightsquigarrow$ results for $\theta = 0$ carry over!

Level and type of immortal line ($\theta = 0$)



Coin



$$h(x) = \mathbb{P}\,(\text{immortal line has type 0 at } t = 0 \mid X_0 = x)$$
$$= \sum_{n \geq 1} \mathbb{P}(L \geq n)(1-x)^{n-1}x$$

$$\mathbb{P}(L \geq n) = \mathbb{P}(\text{level } n \text{ is occupied}),$$

$$1 = \mathbb{P}(L \geq 1) \geq \mathbb{P}(L \geq 2) \geq \mathbb{P}(L \geq 3) \ldots$$

$\rightsquigarrow$ bias towards type 0.

Level and type of immortal line ($\theta > 0$)

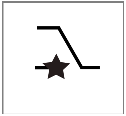
Theorem

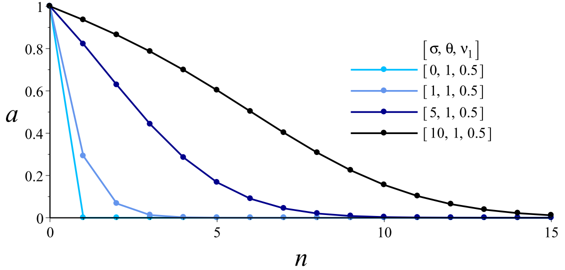
1. The level of the immortal line in the LD-ASG with types assigned at $t = 0$ is either the lowest type-0 level or, if all lines at $t = 0$ are of type 1, it is the level of the immune line.
2. $h(x) = \mathbb{P}\,(\text{immortal line has type } 0 \mid x)$ is the probability of at least one success when tossing L times a coin with success probability x,

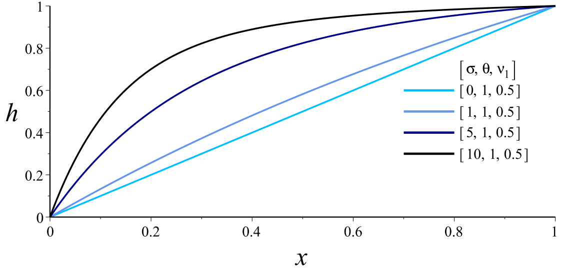
$$h(x) = \sum_{n \geq 1} \mathbb{P}(L \geq n)(1-x)^{n-1} x \ .$$

Some pictures: $a_n := \mathbb{P}(L > n)$ and $h(x)$

$\sigma = 0, 1, 5, 10, \quad \theta = 1, \quad \nu_1 = 0.5$







Conclusion

Pruned LD-ASG to identify ancestral individual and obtain its type distribution

Key ingredients:

- equilibrium ASG (without types)
- ordering of lines
- LD-ASG
- pruned LD-ASG (still without types)
- assign types

TPB 2015