A Siegmund dual for the line counting in a pruned ancestral selection graph

Ute Lenz

Goethe Universität Frankfurt

work in progress

joint with **Ellen Baake** (Bielefeld) and **Anton Wakolbinger** (Frankfurt)

1. Backward in time: ASG
2. pruned LD-ASG (with selection and mutation)
3. Forward in time: dual process

06.11.2014

PROBABILISTIC STRUCTURES IN EVOLUTION

DFG SPP 1590

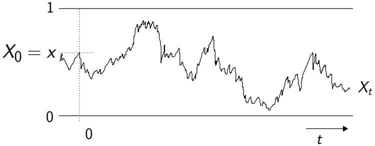
GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Universität Bielefeld

Wright-Fisher diffusion with mutation and selection

$$dX_t = \sqrt{2X_t(1-X_t)}\, dB_t \; + \; \Big[(1-X_t)\theta\nu_0 - X_t\theta\nu_1 + X_t(1-X_t)\sigma\Big]\, dt$$

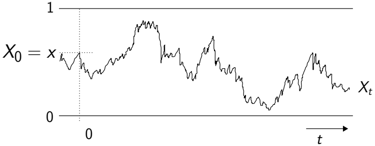
$X_t,\ 1-X_t$	...	frequency type 0 ('good'), type 1 ('bad')
θ	...	mutation rate
$\nu_0,\ \nu_1$	...	mutation probability to type 0, 1, $\nu_0 + \nu_1 = 1$
σ	...	selection coefficient

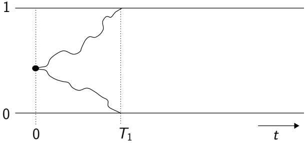


Wright-Fisher diffusion with mutation and selection

$$dX_t = \sqrt{2X_t(1-X_t)}\; dB_t \; + \; \Big[(1-X_t)\theta\nu_0 - X_t\theta\nu_1 + X_t(1-X_t)\sigma\Big]\; dt$$

$X_t, \; 1-X_t$	...	frequency type 0 ('good'), type 1 ('bad')
θ	...	mutation rate
$\nu_0, \; \nu_1$	...	mutation probability to type 0, 1, $\nu_0 + \nu_1 = 1$
σ	...	selection coefficient

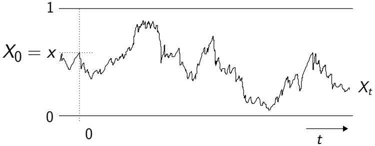


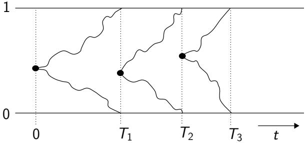


Wright-Fisher diffusion with mutation and selection

$$dX_t = \sqrt{2X_t(1-X_t)}\, dB_t \;+\; \Big[(1-X_t)\theta\nu_0 - X_t\theta\nu_1 + X_t(1-X_t)\sigma\Big]\, dt$$

$X_t,\ 1-X_t$	...	frequency type 0 ('good'), type 1 ('bad')
θ	...	mutation rate
$\nu_0,\ \nu_1$	...	mutation probability to type 0, 1, $\nu_0+\nu_1=1$
σ	...	selection coefficient

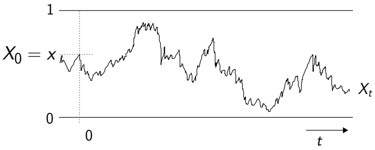


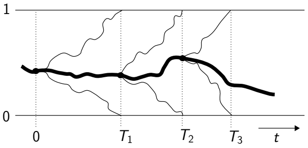


Wright-Fisher diffusion with mutation and selection

$$dX_t = \sqrt{2X_t(1-X_t)}\; dB_t \;+\; \Big[(1-X_t)\theta\nu_0 - X_t\theta\nu_1 + X_t(1-X_t)\sigma\Big]\; dt$$

$X_t,\ 1-X_t$	...	frequency type 0 ('good'), type 1 ('bad')
θ	...	mutation rate
$\nu_0,\ \nu_1$	...	mutation probability to type 0, 1, $\nu_0 + \nu_1 = 1$
σ	...	selection coefficient

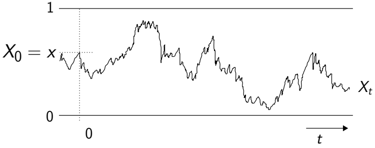


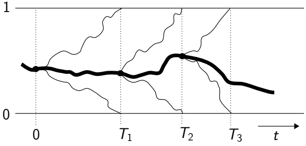


Wright-Fisher diffusion with mutation and selection

$$dX_t = \sqrt{2X_t(1-X_t)}\, dB_t \; + \; \Big[(1-X_t)\theta\nu_0 - X_t\theta\nu_1 + X_t(1-X_t)\sigma\Big]\, dt$$

$X_t,\ 1-X_t$	...	frequency type 0 ('good'), type 1 ('bad')
θ	...	mutation rate
$\nu_0,\ \nu_1$	...	mutation probability to type 0, 1, $\nu_0+\nu_1=1$
σ	...	selection coefficient



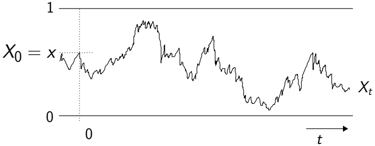


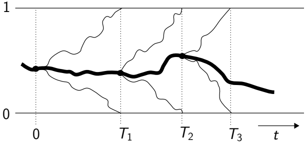
$$h(x) := \mathbb{P}\,(\text{'winner' at } t=0 \text{ is of type } 0 \mid X_0 = x)\}$$

Wright-Fisher diffusion with mutation and selection

$$dX_t = \sqrt{2X_t(1-X_t)}\, dB_t \;+\; \Big[(1-X_t)\theta\nu_0 - X_t\theta\nu_1 + X_t(1-X_t)\sigma\Big]\; dt$$

$X_t,\ 1-X_t$	...	frequency type 0 ('good'), type 1 ('bad')
θ	...	mutation rate
$\nu_0,\ \nu_1$	...	mutation probability to type 0, 1, $\nu_0+\nu_1=1$
σ	...	selection coefficient





$$h(x) := \mathbb{P}\,(\text{'winner' at } t=0 \text{ is of type } 0 \mid X_0 = x)\} = \sum_{n\geq 0} a_{n-1}x(1-x)^{n-1}$$

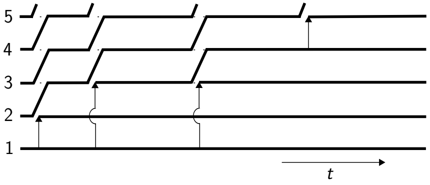
$$(n+1+\theta+\sigma)a_n = (n+1+\theta\nu_1)a_{n+1} + \sigma a_{n-1}$$ (Fearnhead 2002)

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

Lookdown construction with selection

Wright-Fisher diffusion

$$dX_t = \sqrt{2X_t(1-X_t)}\; dB_t$$



Lookdown model

Donnelly & Kurtz 1999

N individuals on levels $\{1, 2, \ldots\}$,
types 0 ('good'), 1 ('bad')

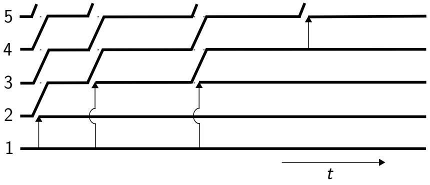
X_t^N frequency of type-0 individuals at time t

$\longrightarrow$ neutral reproduction, rate 2
(for all individuals towards higher levels)

Lookdown construction with selection

Wright-Fisher diffusion

$$dX_t = \sqrt{2X_t(1-X_t)}\; dB_t$$



Lookdown model

Donnelly & Kurtz 1999

N $(\to \infty)$ individuals on levels $\{1, 2, \ldots\}$,
types 0 ('good'), 1 ('bad')

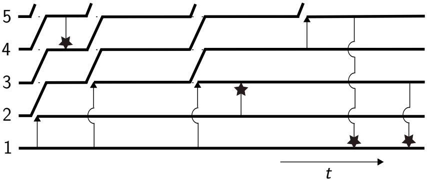
X_t^N $(\to X_t)$ frequency of type-0 individuals at time t

$\longrightarrow$ neutral reproduction, rate 2
(for all individuals towards higher levels)

Lookdown construction with selection

Wright-Fisher diffusion with selection

$$dX_t = \sqrt{2X_t(1-X_t)}\; dB_t \;+\; X_t(1-X_t)\sigma\; dt$$



Lookdown model

Donnelly & Kurtz 1999

N $(\to \infty)$ individuals on levels $\{1, 2, \ldots\}$,
types 0 ('good'), 1 ('bad')

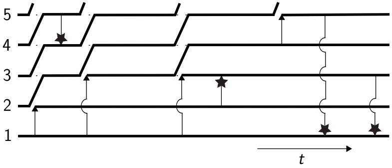
X_t^N $(\to X_t)$ frequency of type-0 individuals at time t

$\longrightarrow$ neutral reproduction, rate 2
(for all individuals towards higher levels)

$\longrightarrow\!\!\star$ selective reproduction, rate s_N (for 0 individuals)
$(Ns_N \to \sigma)$

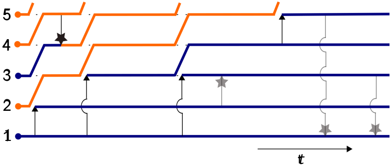
Lookdown construction with selection

assign types and let them evolve (forward in time)

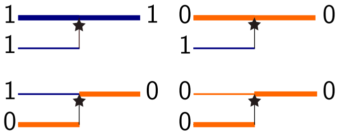


Lookdown construction with selection

assign types and let them evolve (forward in time)

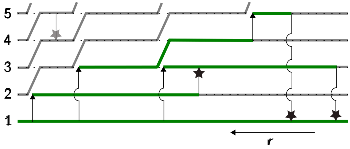


.... and keep in mind pecking order



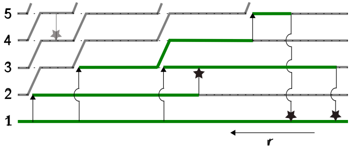
ASG

backward in time: trace back all potential ancestors

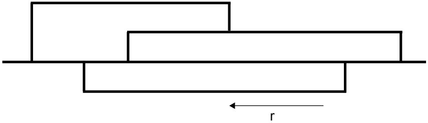


ASG

backward in time: trace back all potential ancestors

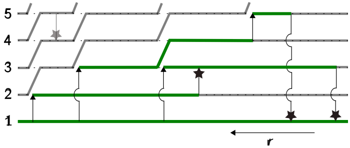


Ancestral Selection Graph (ASG) (Neuhauser and Krone 1997)

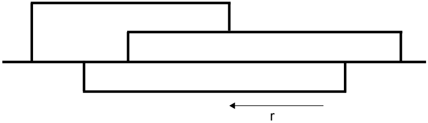


ASG

backward in time: trace back all potential ancestors



Ancestral Selection Graph (ASG) (Neuhauser and Krone 1997)

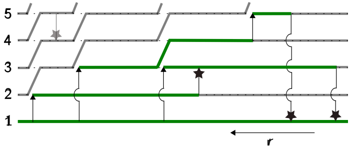


branching:
rate σ per line;

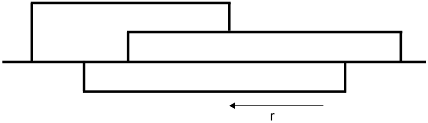
coalescence:
rate 2 per (unordered) pair

ASG

backward in time: trace back all potential ancestors



Ancestral Selection Graph (ASG) (Neuhauser and Krone 1997)



branching:
rate σ per line;

coalescence:
rate 2 per (unordered) pair

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

birth-death process with rates

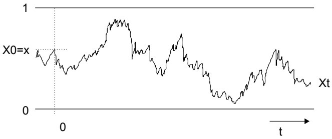
$$q_K(n, n-1) = n(n-1), \quad q_K(n, n+1) = n\sigma, \quad n \in \mathbb{N}$$

and (reversible) equilibrium distribution

$$\mathbb{P}(K_r = n) = \frac{\sigma^n}{(\exp(\sigma)-1)n!} \quad \text{(i.e., Poisson}(\sigma)\text{ conditioned to } \{1, 2, \ldots\})$$

Moment duality

Wright-Fisher diffusion with selection



forward in time

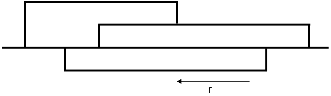
$(X_t)_{t\geq 0}$ frequency of type-0 individuals

$G_X f(x) =$

$x(1-x)f''(x) +$

$x(1-x)\sigma f'(x)$

ASG



backward in time

all potential ancestors of a sample

$(K_r)_{r=-t\geq 0}$ number of lines in ASG

$G_K f(n) =$

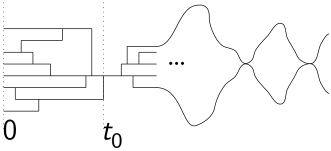
$n(n-1)\left[f(n-1)-f(n)\right] +$

$n\sigma\left[f(n+1)-f(n)\right]$

X and K are moment duals, i.e.

$$\mathbf{E}_x\left[(1-X_s)^n\right] = \mathbf{E}_n\left[(1-x)^{K_s}\right] \quad \forall\ s\geq 0,\ x\in[0,1],\ n\in\mathbb{N}_0$$

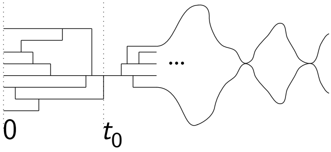
ASG



equilibrium ASG has bottlenecks

$\rightsquigarrow$ type 0 is ancestral iff at least one potential ancestor at $t = 0$ is of type 0 (Mano 2009; Pokalyuk and Pfaffelhuber 2013)

ASG



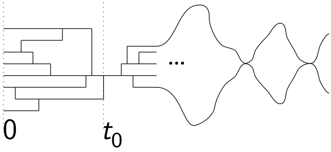
equilibrium ASG has bottlenecks

$\rightsquigarrow$ type 0 is ancestral iff at least one potential ancestor at $t = 0$ is of type 0 (Mano 2009; Pokalyuk and Pfaffelhuber 2013)

assign types by drawing iid according to X_0

$\rightsquigarrow$ $\mathbb{P}(\text{at least one potential ancestor is of type } 0 \mid X_0 = x) = 1 - (1-x)^{K_0}$, so

ASG



equilibrium ASG has bottlenecks

$\rightsquigarrow$ type 0 is ancestral iff at least one potential ancestor at $t = 0$ is of type 0 (Mano 2009; Pokalyuk and Pfaffelhuber 2013)

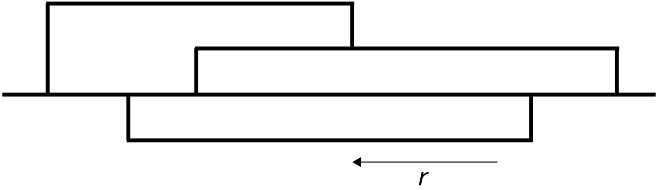
assign types by drawing iid according to X_0

$\rightsquigarrow$ $\mathbb{P}$(at least one potential ancestor is of type 0 $\mid X_0 = x) = 1 - (1-x)^{K_0}$, so

$$\begin{aligned} h(x) &= \sum_{n\geq 1} \left(1-(1-x)^n\right)\left[\mathbb{P}(K_0 \geq n) - \mathbb{P}(K_0 \geq n+1)\right] \\ &= \sum_{n\geq 1} \mathbb{P}(K_0 \geq n) x (1-x)^{n-1} \end{aligned}$$

ASG with mutations

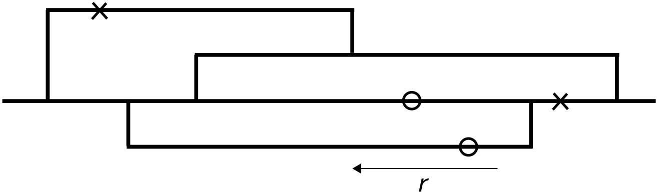
step 1: backward, without types



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

ASG with mutations

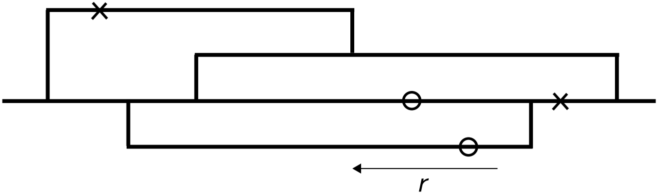
step 1: backward, without types



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

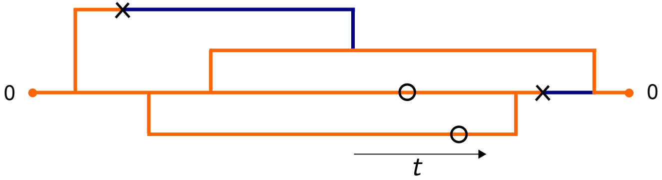
ASG with mutations

step 1: backward, without types



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

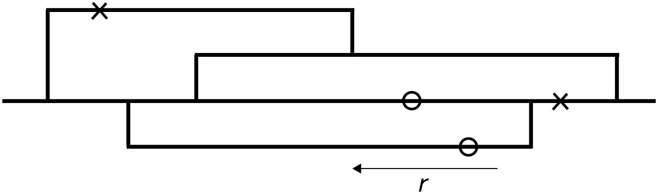
step 2: forward, with types



(resolve branching events)

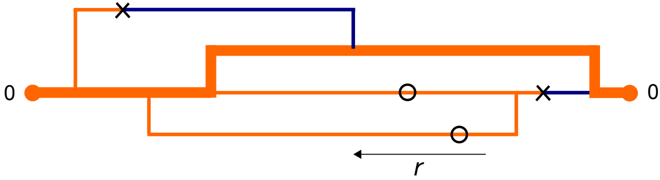
ASG with mutations

step 1: backward, without types



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

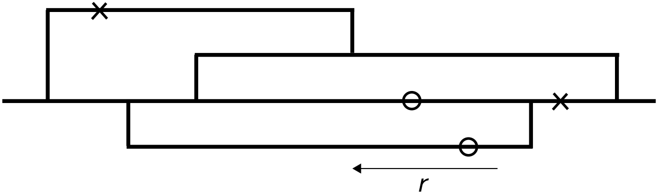
step 2: forward, with types



(identify ancestral line)

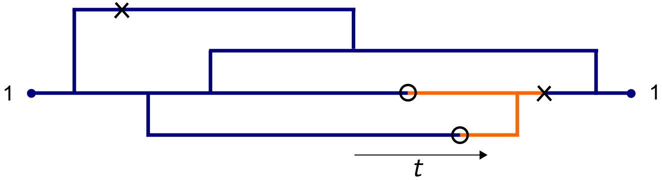
ASG with mutations

step 1: backward, without types



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

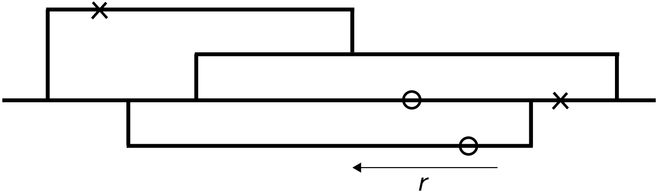
step 2: forward, with types



(resolve branching events)

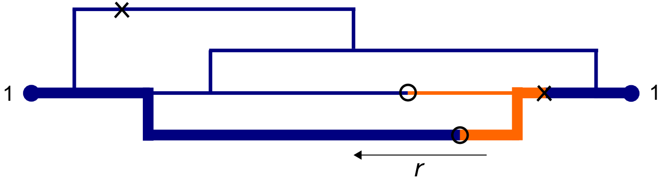
ASG with mutations

step 1: backward, without types



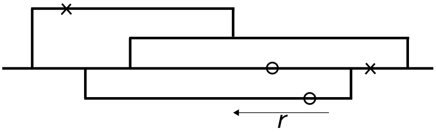
(mutation: rates $\theta\nu_0, \theta\nu_1$ per line)

step 2: forward, with types



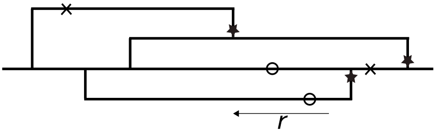
(identify ancestral line)

Ordering the ASG



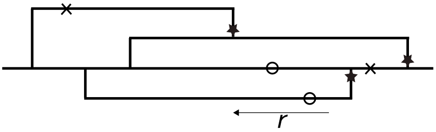
ASG

Ordering the ASG



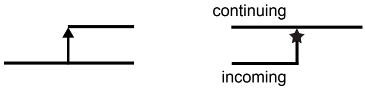
ASG

Ordering the ASG



ASG

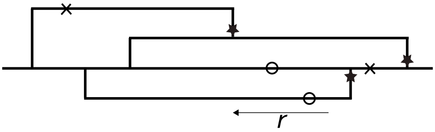
ordering convention:



coalescence

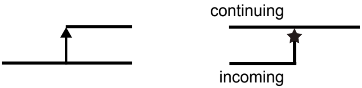
branching

Ordering the ASG

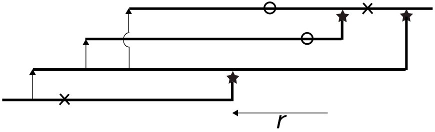


ASG

ordering convention:

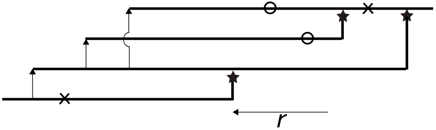


coalescence branching



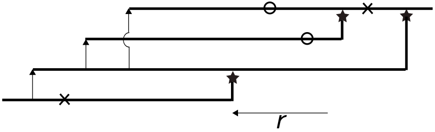
ordered ASG

The Lookdown ASG



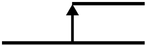
ordered ASG

The Lookdown ASG



ordered ASG

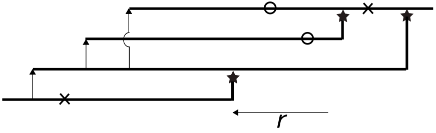
elements:



coalescence

branching

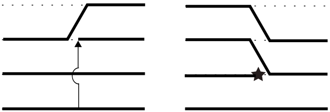
The Lookdown ASG



ordered ASG

elements:

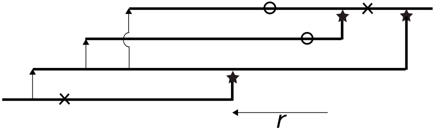
Lookdown elements:



coalescence

branching

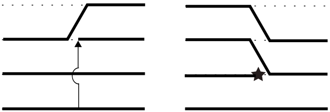
The Lookdown ASG



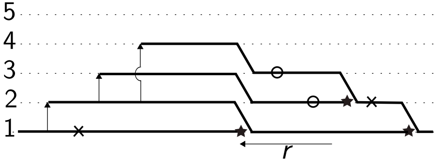
ordered ASG

elements:

Lookdown elements:



coalescence branching

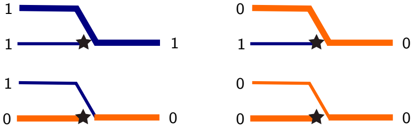


LD-ASG

LD-ASG with types

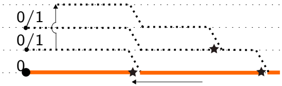
assign types (at $t = 0$ iid according to X_0)
$\rightsquigarrow$ type and level of immortal line?

branching: pecking order

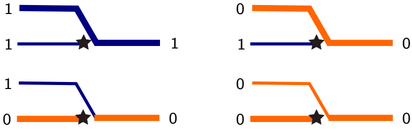


LD-ASG with types

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

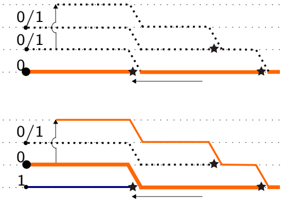


branching: pecking order

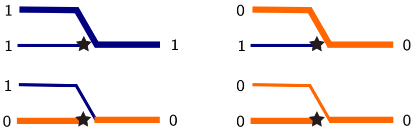


LD-ASG with types

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

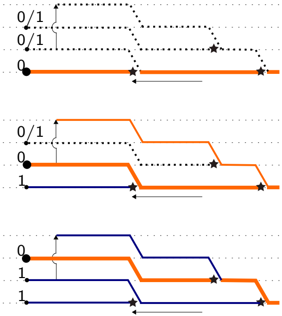


branching: pecking order

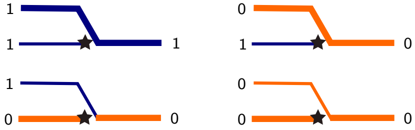


LD-ASG with types

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



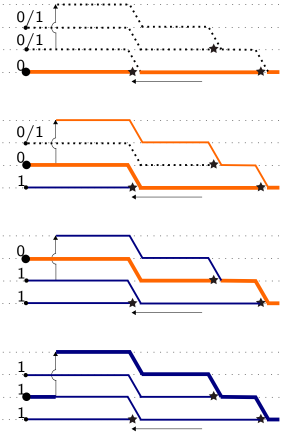
branching: pecking order



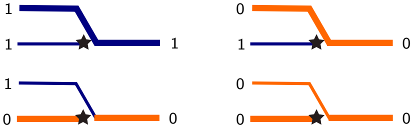
LD-ASG with types

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

$\rightsquigarrow$ type and level of immortal line?

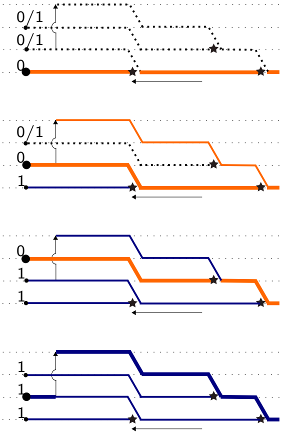


branching: pecking order

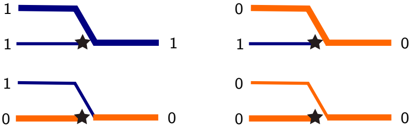


LD-ASG with types

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



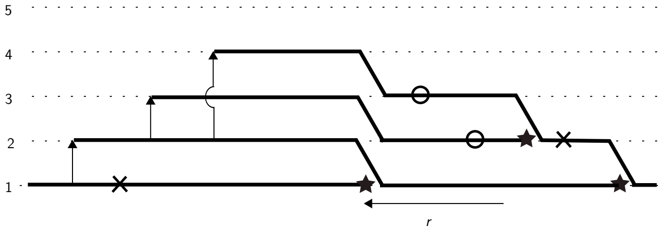
branching: pecking order



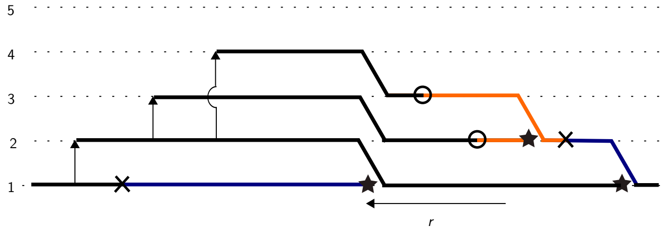
immune line
(:=immortal if all lines are of type 1)

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

LD-ASG with mutations



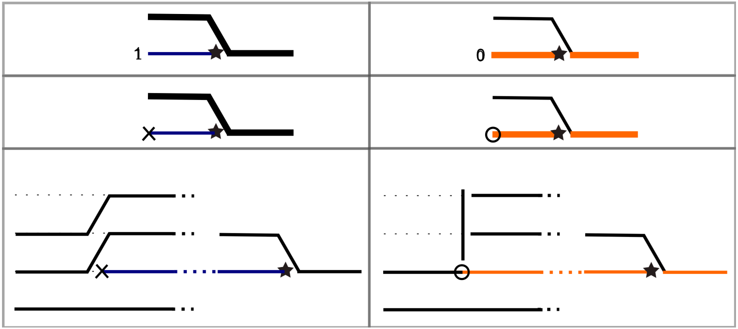
LD-ASG with mutations



some potential ancestors drop out
$\rightarrow$ erase lines and get a pruned LD-ASG

The pruned LD-ASG

Pruning procedure



mutation reduces number of potential ancestors

- (×) erase lines that carry the deleterious mutation
- (○) erase all lines above a beneficial mutation

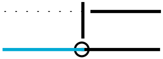
The pruned LD-ASG

Pruning procedure

But: immune line cannot be erased
(immune line at time t is immortal
if all lines at time t are assigned type 1)

- is not erased when hit by a cross
  but lines are rearranged

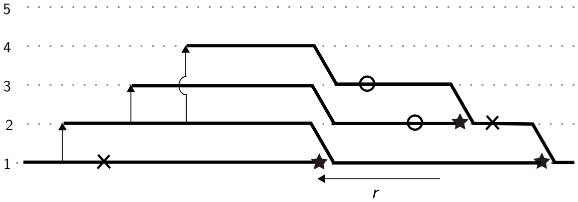
- jumps to the level of the circle at
  a beneficial mutation



The pruned LD-ASG

Construction from a given realisation of the LD-ASG

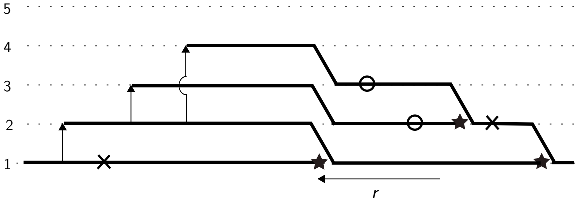
LD-ASG:



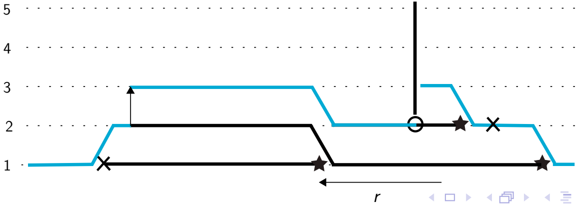
The pruned LD-ASG

Construction from a given realisation of the LD-ASG

LD-ASG:

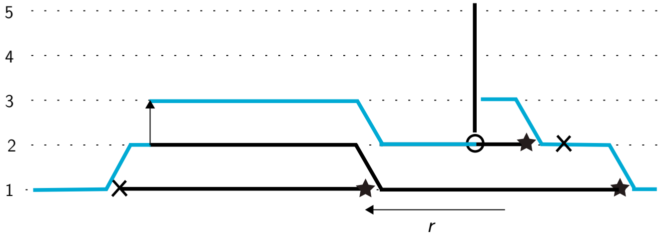


pruned LD-ASG:



The pruned LD-ASG

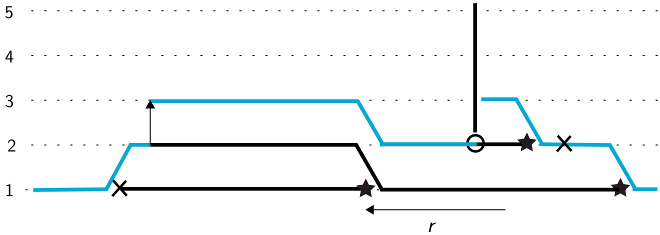
Markovian dynamics backward in time



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

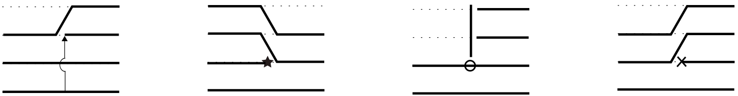
The pruned LD-ASG

Markovian dynamics backward in time



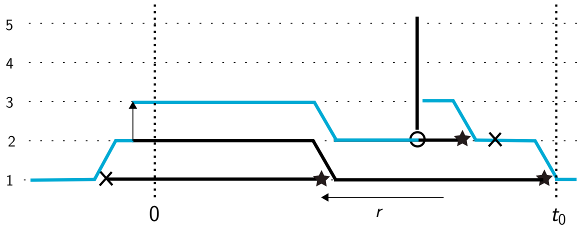
$L_r :=$ highest occupied level ($=$ number of lines) at time r

$(L_r)_{r\in\mathbb{R}}$ jump process with rates



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

The pruned LD-ASG



realisation of pruned equilibrium LD-ASG

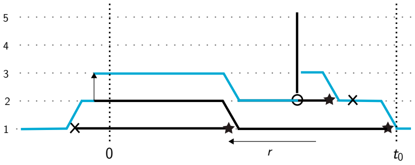
no mutations except on the immune line

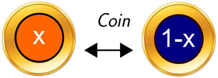
assign types at $t = 0$

Lemma (Type of the immortal line)

The immortal line at $t = 0$ in the pruned equilibrium LD-ASG with types assigned at $t = 0$ is of type 0 iff there is at least one line of type 0.

Common ancestor type distribution





$X_r :=$ frequency of type-0 individuals at time r

$I^k :=$ type of line on level $k \in \mathbb{N}_0$

$$
\begin{aligned}
h(x) &= \mathbb{P}\left(\text{immortal line has type 0 at } t = 0 \mid X_0 = x\right) \\
&= \mathbb{P}\left(\text{at least one line of type 0 at } t = 0 \mid X_0 = x\right) \\
&= \sum_{n \geq 1} \mathbb{P}\left(I^n = 0, I^1 = \ldots = I^{n-1} = 1, L_0 \geq n \mid X_0 = x\right) \\
&= \sum_{n \geq 1} x(1-x)^{n-1} \cdot \mathbb{P}(L_0 \geq n)
\end{aligned}
$$

Common ancestor type distribution

Theorem (main result)

1. $h(x) = \mathbb{P}\,(\text{immortal line has type } 0 \mid x)$ is the probability of at least one success when tossing L^{equ} times a coin with success probability x,

$$h(x) = \sum_{n \geq 1} x(1-x)^{n-1} \cdot a_{n-1}$$

2. The Fearnhead coefficients (a_n) are the tail probabilities of the stationary distribution of L, $a_{n-1} = \mathbb{P}(L^{equ} \geq n)$, determined by

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

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

Common ancestor type distribution

Theorem (main result)

1. $h(x) = \mathbb{P}$ (immortal line has type $0 \mid x$) is the probability of at least one success when tossing L^{equ} times a coin with success probability x,

$$h(x) = \sum_{n \geq 1} x(1-x)^{n-1} \cdot a_{n-1} \quad \checkmark$$

2. The Fearnhead coefficients (a_n) are the tail probabilities of the stationary distribution of L, $a_{n-1} = \mathbb{P}(L^{equ} \geq n)$, determined by

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

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

Common ancestor type distribution

Theorem (main result)

1. $h(x) = \mathbb{P}\,(\text{immortal line has type } 0 \mid x)$ is the probability of at least one success when tossing L^{equ} times a coin with success probability x,

$$h(x) = \sum_{n \geq 1} x(1-x)^{n-1} \cdot a_{n-1} \quad \checkmark$$

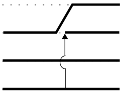
2. The Fearnhead coefficients (a_n) are the tail probabilities of the stationary distribution of L, $a_{n-1} = \mathbb{P}(L^{equ} \geq n)$, determined by

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

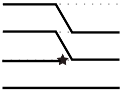
$$1 = a_0 \geq a_1 \geq \cdots \, , \qquad \lim_{n\to\infty} a_{n+1}/a_n = 0 \; .$$

The processes L and D

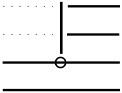
transition rates of L backward in time



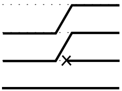
$q_L^{\uparrow}(\ell, \ell - 1)$
$= \ell(\ell - 1)$



$q_L^{*}(\ell, \ell + 1)$
$= \ell\sigma$



$q_L^{\circ}(\ell, \ell - m)$
$= \theta\nu_0$

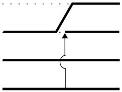


$q_L^{\times}(\ell, \ell - 1)$
$= (\ell - 1)\theta\nu_1$

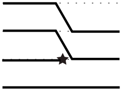
ignore on immune line

The processes L and D

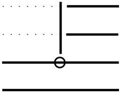
transition rates of L backward in time



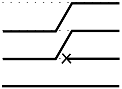
$q_L^{\uparrow}(\ell, \ell-1) = \ell(\ell-1)$



$q_L^{*}(\ell, \ell+1) = \ell\sigma$



$q_L^{\circ}(\ell, \ell-m) = \theta\nu_0$



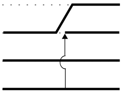
$q_L^{\times}(\ell, \ell-1) = (\ell-1)\theta\nu_1$

ignore on immune line

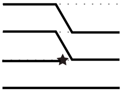
"transposed" rates forward in time

The processes L and D

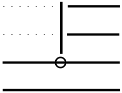
transition rates of L backward in time



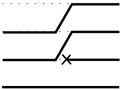
$q_L^{\uparrow}(\ell, \ell-1)$
$= \ell(\ell-1)$



$q_L^{*}(\ell, \ell+1)$
$= \ell\sigma$



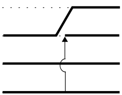
$q_L^{\circ}(\ell, \ell-m)$
$= \theta\nu_0$



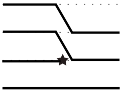
$q_L^{\times}(\ell, \ell-1)$
$= (\ell-1)\theta\nu_1$

ignore on immune line

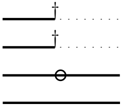
"transposed" rates forward in time



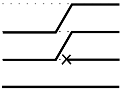
$q_D^{\uparrow}(n, n+1)$
$= n(n-1)$



$q_D^{*}(n, n-1)$
$= (n-1)\sigma$



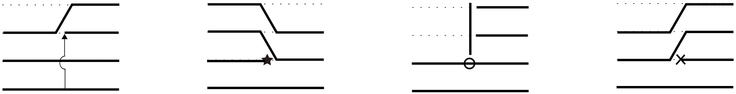
$q_D^{\circ}(n, \dagger)$
$= (n-1)\theta\nu_0$



$q_D^{\times}(n, n+1)$
$= (n-1)\theta\nu_1$

The processes L and D

transition rates of L backward in time



$q_L^{\uparrow}(\ell, \ell-1) = \ell(\ell-1)$

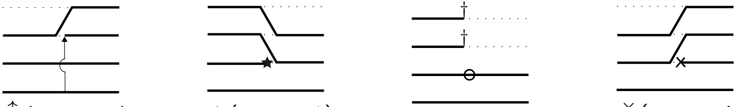
$q_L^{*}(\ell, \ell+1) = \ell\sigma$

$q_L^{\circ}(\ell, \ell-m) = \theta\nu_0$

$q_L^{\times}(\ell, \ell-1) = (\ell-1)\theta\nu_1$

ignore on immune line

"transposed" rates forward in time



$q_D^{\uparrow}(n, n+1) = n(n-1)$

$q_D^{*}(n, n-1) = (n-1)\sigma$

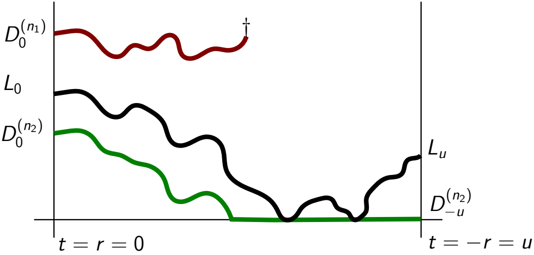
$q_D^{\circ}(n, \dagger) = (n-1)\theta\nu_0$

$q_D^{\times}(n, n+1) = (n-1)\theta\nu_1$

$\rightsquigarrow$ forward in time process $(D_t)_{t\geq 0}$:

- birth and death process
- absorbing states 1, $\dagger$, and ∞

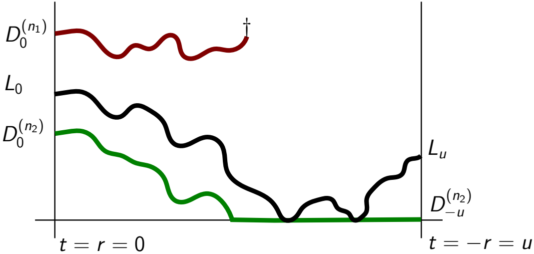
The processes L and D



Heuristics:

$$\mathbf{P}(L_0 \geq n) = \mathbf{P}(D^{(n)} \text{ eventually reaches level 1 forward in time})$$

The processes L and D

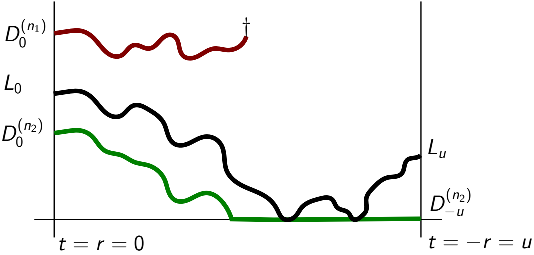


Heuristics:

$$\mathbf{P}(L_0 \geq n) = \mathbf{P}(D^{(n)} \text{ eventually reaches level 1 forward in time})$$

$\rightsquigarrow$ tail probabilities of L = absorption probabilities of D in 1

The processes L and D



Indeed,

Theorem (Siegmund duality)

L and D are Siegmund duals, i.e.

$$\mathbf{E}_\ell \left[\mathbf{I}_{\{L_s \geq n\}} \right] = \mathbf{E}_n \left[\mathbf{I}_{\{\ell \geq D_s\}} \right] \quad \forall s \geq 0, \ell \in \mathbb{N}_0, n \in \mathbb{N}_0 \cup \{\dagger\}$$

Proof:
Check calculation for the generators with duality function $\mathbf{I}_{\{L \geq D\}}$.

Siegmund duality

$$\mathbf{E}_\ell\left[\mathbf{I}_{\{L_s\geq n\}}\right] = \mathbf{E}_n\left[\mathbf{I}_{\{\ell\geq D_s\}}\right] \quad \forall \ell\in\mathbb{N}_0, n\in\mathbb{N}_0\cup\{\dagger\}$$
$$\Leftrightarrow \quad \mathbf{P}_\ell(L_s\geq n) = \mathbf{P}_n(\ell\geq D_s)$$

take $\ell = 1, s\to\infty$ (equilibrium situation)
(start L with one line in the far future)

$$\mathbf{P}(L^{equ}\geq n) = \mathbf{P}_n(1\geq D_\infty)$$

$$a_{n-1} = \mathbf{P}(L^{equ}\geq n) = \mathbf{P}_n(D_\infty = 1)$$

First step decomposition of the event 'absorption in one' yields

$$[n(n-1) + (n-1)\theta\nu_1 + (n-1)\sigma + (n-1)\theta\nu_0]\, a_{n-1}$$
$$= [n(n-1) + (n-1)\theta\nu_1]\, a_n \; + \; (n-1)\sigma a_{n-2} \; + \; (n-1)\theta\nu_0 a_\dagger$$

$$\Leftrightarrow [n+1+\theta+\sigma]\, a_n = [n+1+\theta\nu_1]\, a_{n+1} + \sigma a_{n-1}$$

Siegmund duality

$$\mathbf{E}_\ell\left[\mathbf{I}_{\{L_s \geq n\}}\right] = \mathbf{E}_n\left[\mathbf{I}_{\{\ell \geq D_s\}}\right] \quad \forall \ell \in \mathbb{N}_0, n \in \mathbb{N}_0 \cup \{\dagger\}$$

$$\Leftrightarrow \quad \mathbf{P}_\ell(L_s \geq n) = \mathbf{P}_n(\ell \geq D_s)$$

take $\ell = 1, s \to \infty$ (equilibrium situation)
(start L with one line in the far future)

$$\mathbf{P}(L^{equ} \geq n) = \mathbf{P}_n(1 \geq D_\infty)$$

$$a_{n-1} = \mathbf{P}(L^{equ} \geq n) = \mathbf{P}_n(D_\infty = 1)$$

First step decomposition of the event 'absorption in one' yields

$$[n(n-1) + (n-1)\theta\nu_1 + (n-1)\sigma + (n-1)\theta\nu_0]\, a_{n-1}$$
$$= [n(n-1) + (n-1)\theta\nu_1]\, a_n \; + \; (n-1)\sigma a_{n-2} \; + \; (n-1)\theta\nu_0 a_\dagger$$

$$\Leftrightarrow [n + 1 + \theta + \sigma]\, a_n = [n + 1 + \theta\nu_1]\, a_{n+1} + \sigma a_{n-1}$$

Siegmund duality

$$\mathbf{E}_\ell\left[\mathbf{I}_{\{L_s \geq n\}}\right] = \mathbf{E}_n\left[\mathbf{I}_{\{\ell \geq D_s\}}\right] \quad \forall \ell \in \mathbb{N}_0, n \in \mathbb{N}_0 \cup \{\dagger\}$$
$$\Leftrightarrow \quad \mathbf{P}_\ell(L_s \geq n) = \mathbf{P}_n(\ell \geq D_s)$$

take $\ell = 1, s \to \infty$ (equilibrium situation)
(start L with one line in the far future)

$$\mathbf{P}(L^{equ} \geq n) = \mathbf{P}_n(1 \geq D_\infty)$$

$$a_{n-1} = \mathbf{P}(L^{equ} \geq n) = \mathbf{P}_n(D_\infty = 1)$$

First step decomposition of the event 'absorption in one' yields

$$[n(n-1) + (n-1)\theta\nu_1 + (n-1)\sigma + (n-1)\theta\nu_0]\, a_{n-1}$$
$$= [n(n-1) + (n-1)\theta\nu_1]\, a_n \;+\; (n-1)\sigma a_{n-2} \;+\; (n-1)\theta\nu_0 a_\dagger$$

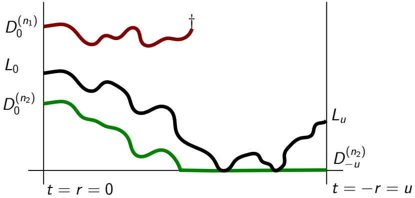
$$\Leftrightarrow [n+1+\theta+\sigma]\, a_n = [n+1+\theta\nu_1]\, a_{n+1} + \sigma a_{n-1} \;\checkmark$$

Siegmund duality

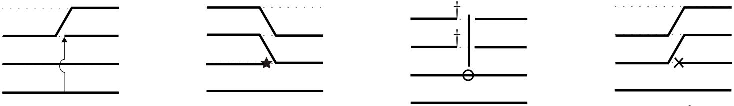
Siegmund duals are strongly pathwise duals:

We can find a version of the processes L and D, defined on a common probability space, such that $\forall \ell, n \in \mathbb{N}, s \geq 0$

$$\mathbf{1}_{\{L_s^{(\ell)} \geq n\}} = \mathbf{1}_{\{(\ell \geq D_s^{(n)}\}}$$



graphical representation of the underlying auxiliary process:



rate 2 per pair of levels

rate σ per level

rate $\theta\nu_0$ per level

rate $\theta\nu_1$ per level

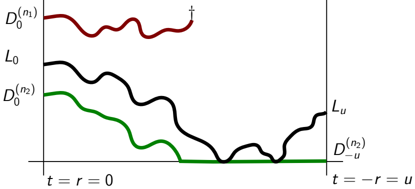
ignore on immune line

Siegmund duality

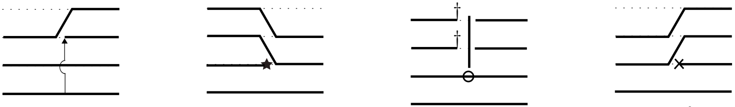
Siegmund duals are strongly pathwise duals:

We can find a version of the processes L and D, defined on a common probability space, such that $\forall \ell, n \in \mathbb{N}, s \geq 0$

$$\mathbf{1}_{\{L_s^{(\ell)} \geq n\}} = \mathbf{1}_{\{(\ell \geq D_s^{(n)}\}}$$



graphical representation of the underlying auxiliary process:



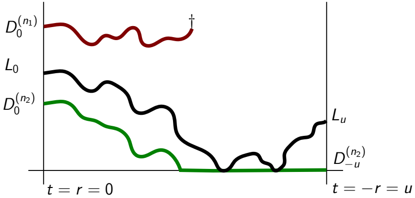
rate 2 per pair of levels | rate σ per level | rate $\theta\nu_0$ per level | rate $\theta\nu_1$ per level

~~ignore on immune line~~

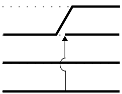
Siegmund duals are strongly pathwise duals:

We can find a version of the processes L and D, defined on a common probability space, such that $\forall \ell, n \in \mathbb{N}, s \geq 0$

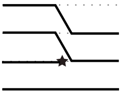
$$\mathbf{1}_{\left\{L_s^{(\ell)} \geq n\right\}} = \mathbf{1}_{\left\{(\ell \geq D_s^{(n)}\right\}}$$



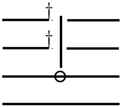
graphical representation of the underlying auxiliary process:



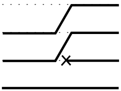
rate 2
per pair of levels



rate σ
per level



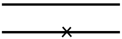
rate $\theta\nu_0$
per level



rate $\theta\nu_1$
per level

~~ignore on immune line~~

ignore on level 1

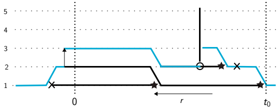


Conclusion

Two-type Wright-Fisher diffusion with mutation and selection

Backward in time:

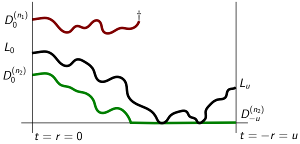
Transparent graphical method to identify all potential ancestors: pruned equilibrium LD-ASG (L top level)



arXiv:1409.0642

Forward in time:

Siegmund dual process D yields tail probabilities of L



Thank you!