

Lines of descent in a Moran model with frequency-dependent selection and mutation.

Luigi Esercito

joint work with Ellen Baake & Sebastian Hummel

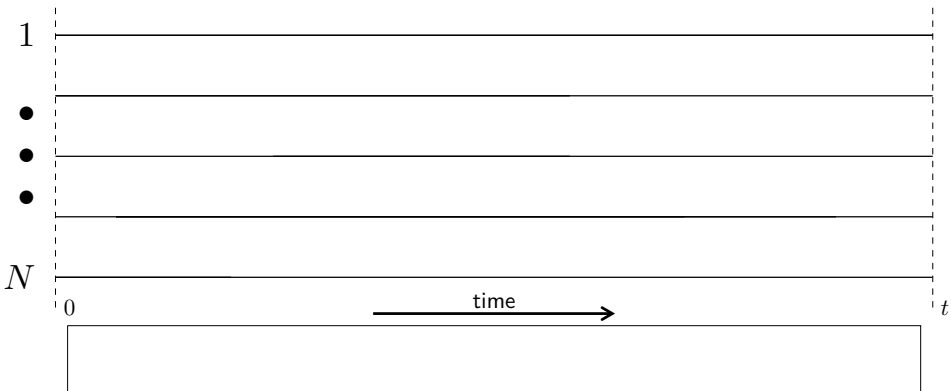
Bielefeld University



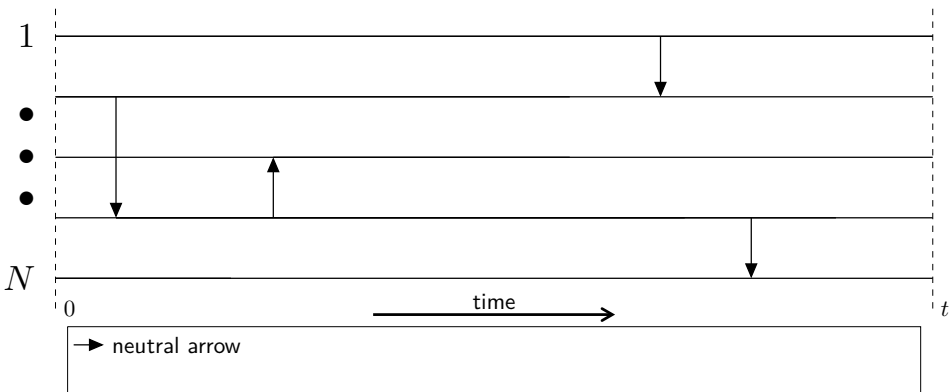
Moran model with non-linear dominance (DOM)

- Population composed by N haploid individuals, constant in size
- Continuous-time model
- Panmixia
- Two types: 0 (*fit*), 1 (*unfit*) (one locus of interest)
- *Selection*: type-1 individuals reproduce at rate 1, type-0 individuals reproduce at rate $1 + \sum_{m>0} \hat{s}_m \left(\frac{k}{N}\right)^{m-1}$, $\hat{s}_m \geq 0$
- *Mutation*: individuals experience a (*deleterious*) mutation to type 1 with rate $u\nu_1$, and a (*beneficial*) mutation to type 0 with rate $u\nu_0$.

DOM Moran model forward: untyped



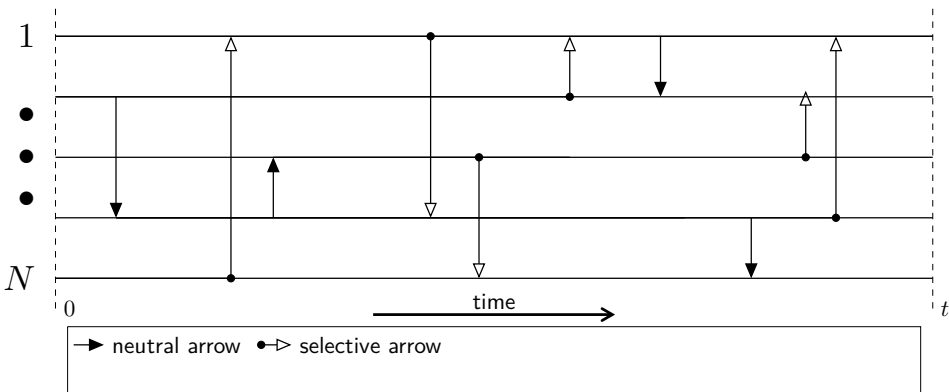
DOM Moran model forward: untyped



neutral arrows:

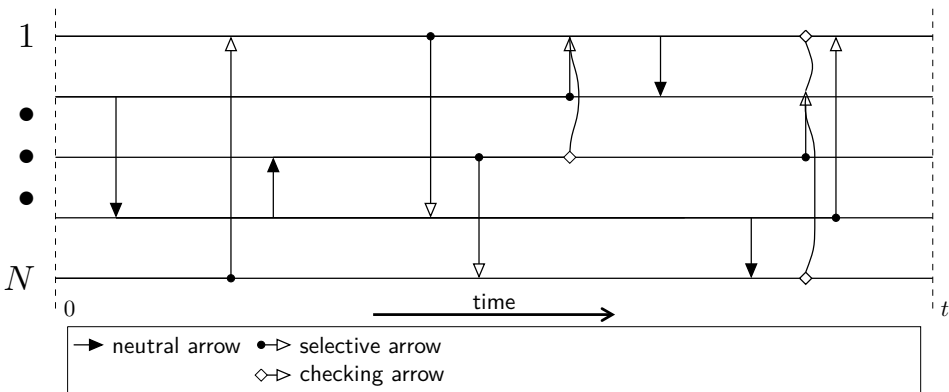
rate 1,

DOM Moran model forward: untyped



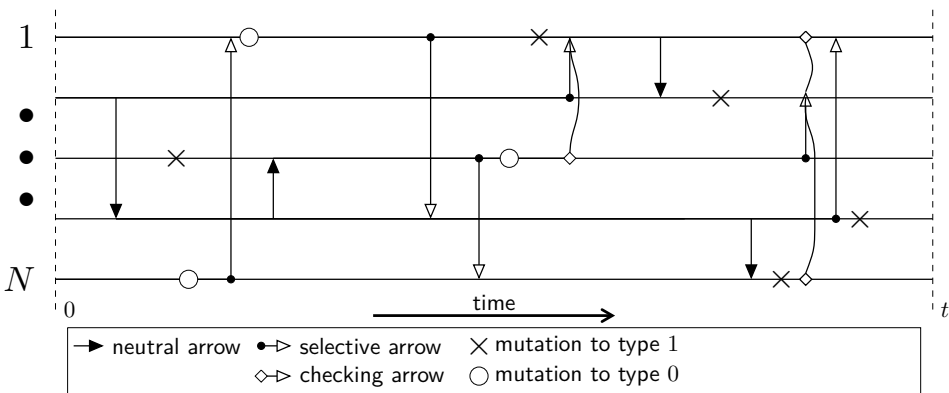
neutral arrows: rate 1,
selective arrows of order m : rate $\hat{s}_m^{(N)}$,

DOM Moran model forward: untyped



neutral arrows: rate 1,
 selective arrows of order m : rate $\hat{s}_m^{(N)}$,

DOM Moran model forward: untyped



neutral arrows:

rate 1,

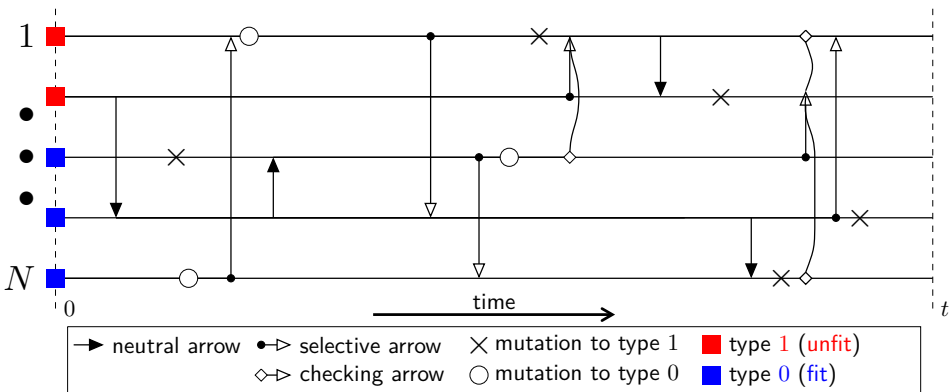
selective arrows of order m :

rate $\hat{s}_m^{(N)}$,

mutation to type 1 (resp. 0):

rate $u^{(N)}\nu_1$ (resp. $u^{(N)}\nu_0$), ($\nu_0 + \nu_1 = 1$).

DOM Moran model forward: untyped



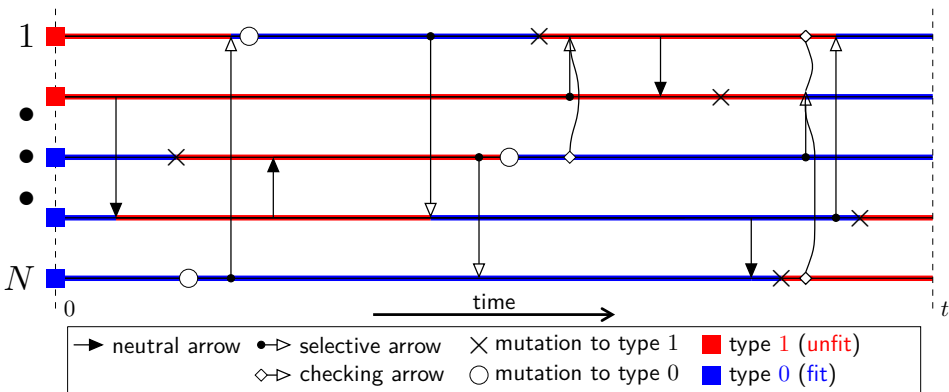
neutral arrows:

rate 1,

selective arrows of order m : rate $\hat{s}_m^{(N)}$,

mutation to type 1 (resp. 0): rate $u^{(N)}\nu_1$ (resp. $u^{(N)}\nu_0$), ($\nu_0 + \nu_1 = 1$).

DOM Moran model forward: untyped



neutral arrows:

rate 1,

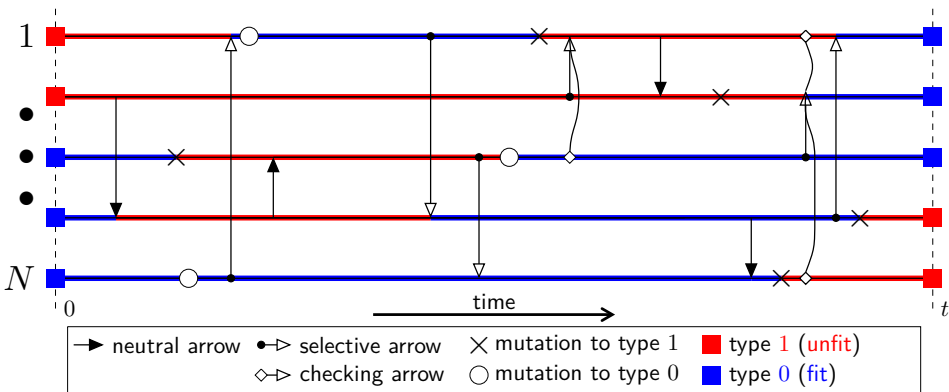
selective arrows of order m :

rate $\hat{s}_m^{(N)}$,

mutation to type 1 (resp. 0):

rate $u^{(N)}\nu_1$ (resp. $u^{(N)}\nu_0$), ($\nu_0 + \nu_1 = 1$).

DOM Moran model forward: typed



neutral arrows:

rate 1,

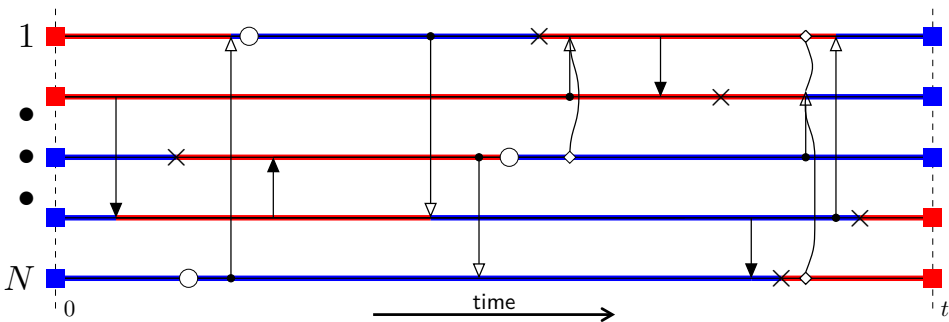
selective arrows of order m :

rate $\hat{s}_m^{(N)}$,

mutation to type 1 (resp. 0):

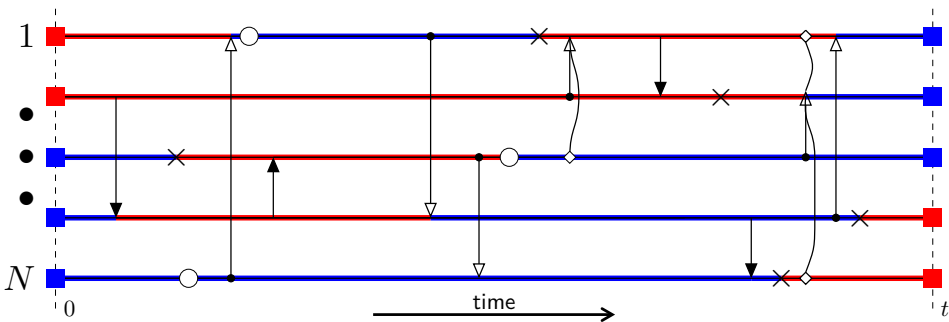
rate $u^{(N)}\nu_1$ (resp. $u^{(N)}\nu_0$), ($\nu_0 + \nu_1 = 1$).

DOM Moran model forward: typed



$\hat{Y}_t^{(N)}$:= proportion of individuals of type 1 at time t in DOM MoMo of size N

DOM Moran model forward: typed

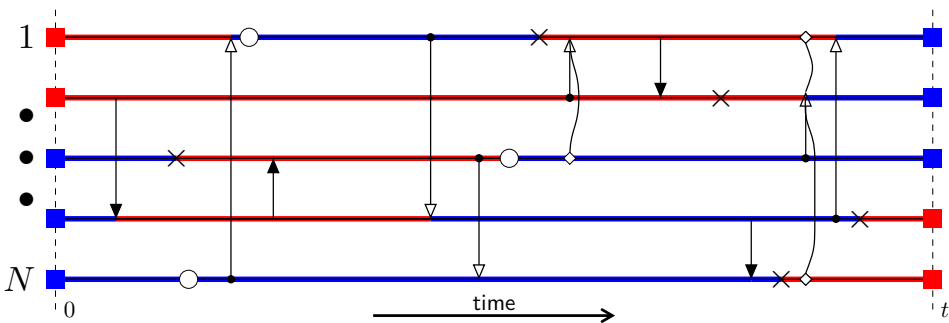


$\hat{Y}_t^{(N)}$:= proportion of individuals of type 1 at time t in DOM MoMo of size N

$$q_{\hat{Y}}^{(N)}\left(\frac{k}{N}, \frac{k+1}{N}\right) =$$

$$q_{\hat{Y}}^{(N)}\left(\frac{k}{N}, \frac{k-1}{N}\right) =$$

DOM Moran model forward: typed

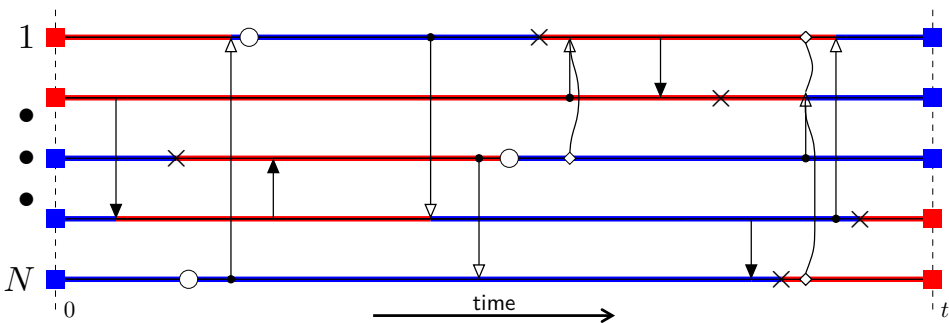


$\hat{Y}_t^{(N)}$:= proportion of individuals of type 1 at time t in DOM MoMo of size N

$$q_{\hat{Y}}^{(N)}\left(\frac{k}{N}, \frac{k+1}{N}\right) = k \frac{N-k}{N}$$

$$q_{\hat{Y}}^{(N)}\left(\frac{k}{N}, \frac{k-1}{N}\right) = (N-k) \frac{k}{N}$$

DOM Moran model forward: typed

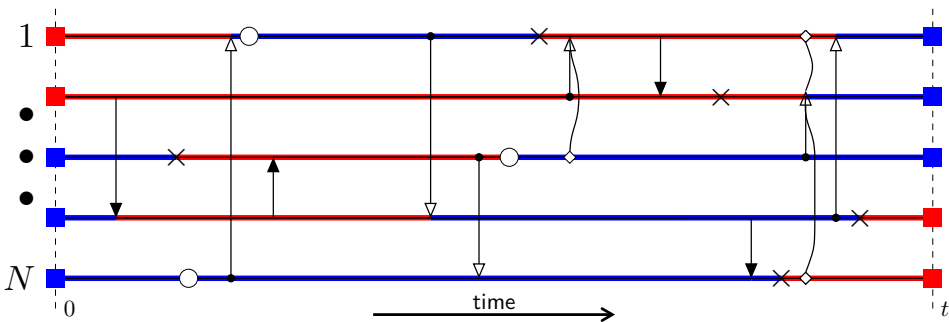


$\hat{Y}_t^{(N)}$:= proportion of individuals of type 1 at time t in DOM MoMo of size N

$$q_{\hat{Y}}^{(N)}\left(\frac{k}{N}, \frac{k+1}{N}\right) = k \frac{N-k}{N}$$

$$q_{\hat{Y}}^{(N)}\left(\frac{k}{N}, \frac{k-1}{N}\right) = (N-k) \frac{k}{N} \left(1 + \sum_{m>0} \hat{s}_m^{(N)}\right)$$

DOM Moran model forward: typed

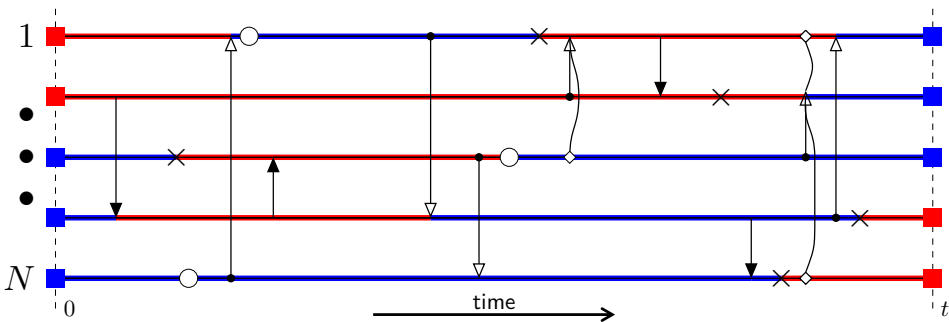


$\hat{Y}_t^{(N)}$:= proportion of individuals of type 1 at time t in DOM MoMo of size N

$$q_{\hat{Y}}^{(N)}\left(\frac{k}{N}, \frac{k+1}{N}\right) = k \frac{N-k}{N}$$

$$q_{\hat{Y}}^{(N)}\left(\frac{k}{N}, \frac{k-1}{N}\right) = (N-k) \frac{k}{N} \left(1 + \sum_{m>0} \hat{s}_m^{(N)} \left(\frac{k}{N}\right)^{m-1}\right)$$

DOM Moran model forward: typed



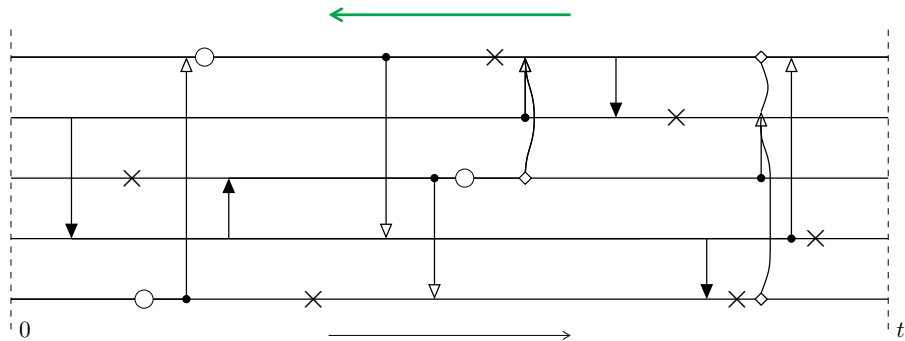
$\hat{Y}_t^{(N)}$:= proportion of individuals of type 1 at time t in DOM MoMo of size N

$$q_{\hat{Y}}^{(N)}\left(\frac{k}{N}, \frac{k+1}{N}\right) = k \frac{N-k}{N} + (N-k)u^{(N)}\nu_1$$

$$q_{\hat{Y}}^{(N)}\left(\frac{k}{N}, \frac{k-1}{N}\right) = (N-k) \frac{k}{N} \left(1 + \sum_{m>0} \hat{s}_m^{(N)} \left(\frac{k}{N}\right)^{m-1}\right) + k u^{(N)} \nu_0$$

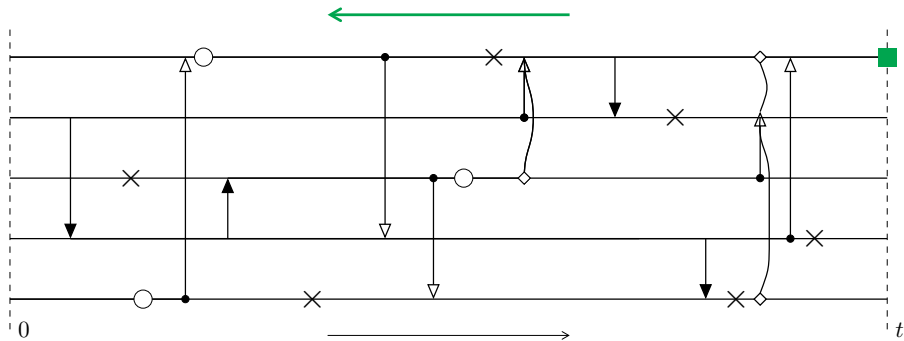
Ancestral selection graph (ASG)

Krone and Neuhauser 1997



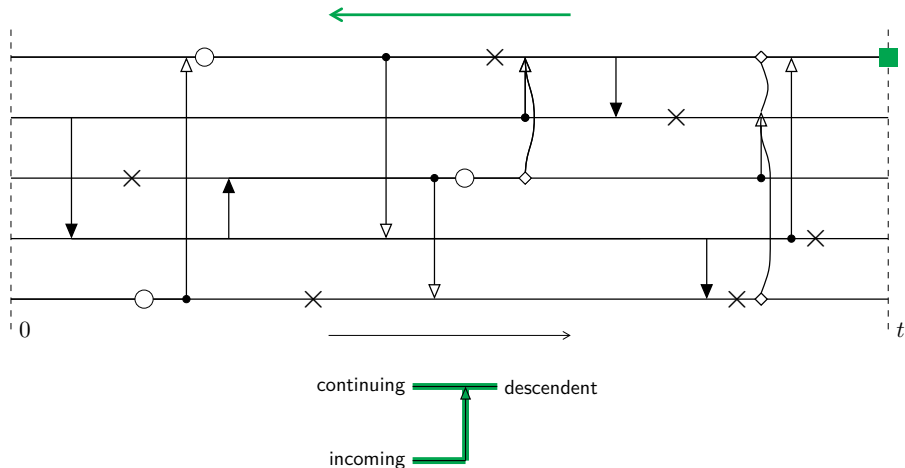
Ancestral selection graph (ASG)

Krone and Neuhauser 1997



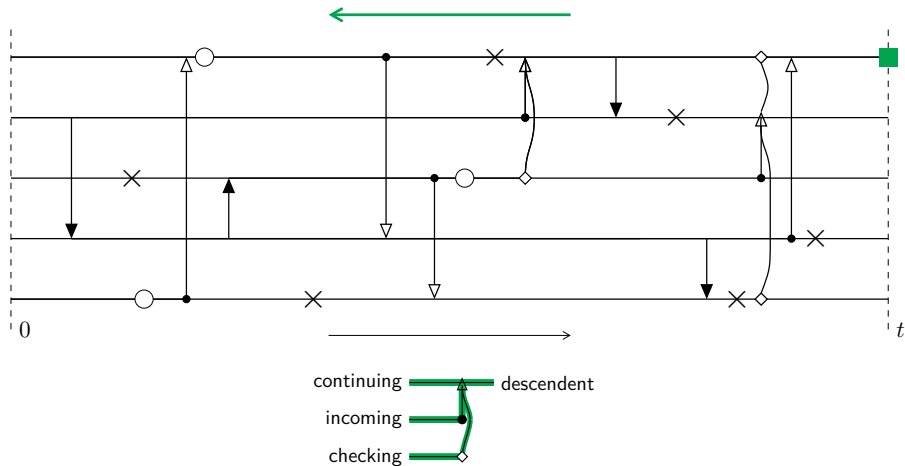
Ancestral selection graph (ASG)

Krone and Neuhauser 1997



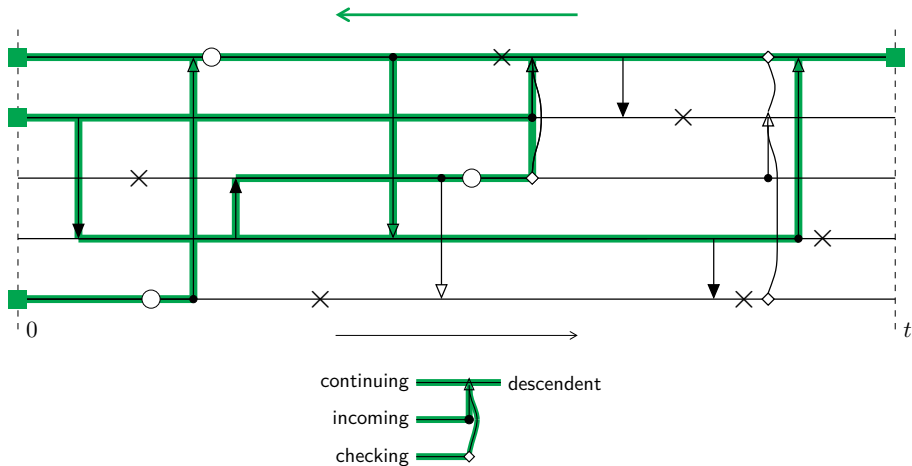
Ancestral selection graph (ASG)

Krone and Neuhauser 1997



Ancestral selection graph (ASG)

Krone and Neuhauser 1997

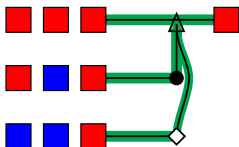
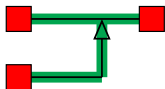


- Problem #1 ancestral lines don't all have the same role: influencers vs potential ancestors

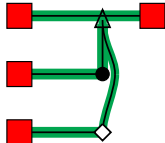
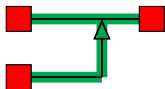
- Problem #1 ancestral lines don't all have the same role: influencers vs potential ancestors
- Problem #2 coalescences complicate the situation

- Problem #1 ancestral lines don't all have the same role: influencers vs potential ancestors
- Problem #2 coalescences complicate the situation
- Idea: find some more "symmetric" rule for selection

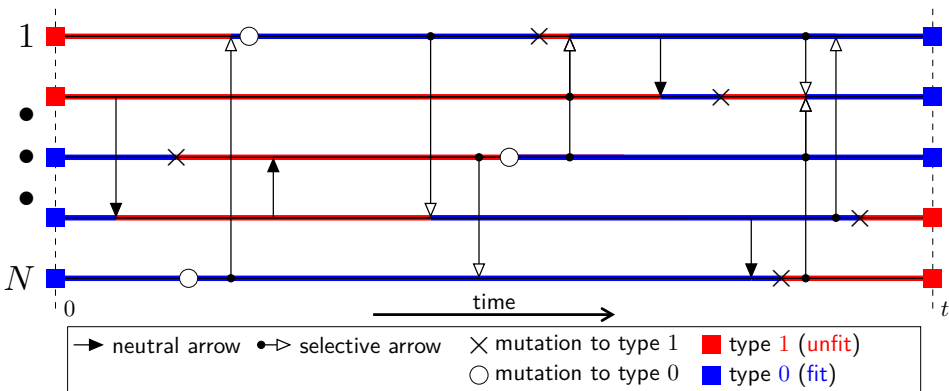
- Problem #1 ancestral lines don't all have the same role: influencers vs potential ancestors
- Problem #2 coalescences complicate the situation
- Idea: find some more "symmetric" rule for selection



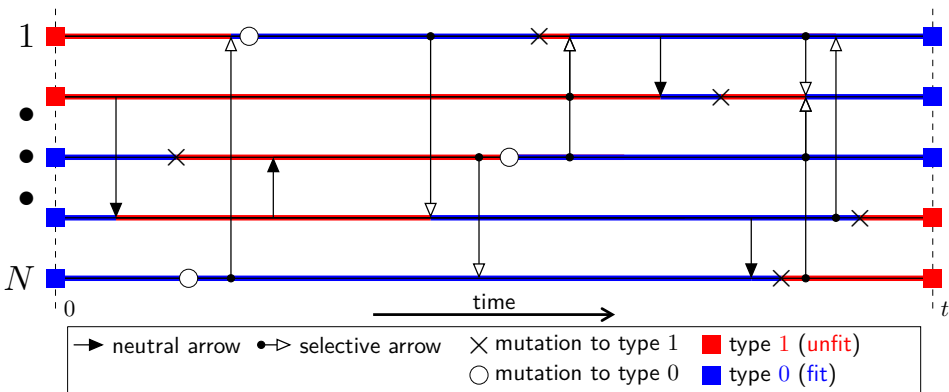
- **Problem #1** ancestral lines don't all have the same role: influencers vs potential ancestors
- **Problem #2** coalescences complicate the situation
- **Idea:** find some more "symmetric" rule for selection
→ "fittest type wins"!



Fittest-type-wins Moran model (FTW)



Fittest-type-wins Moran model (FTW)



neutral arrows:

rate 1,

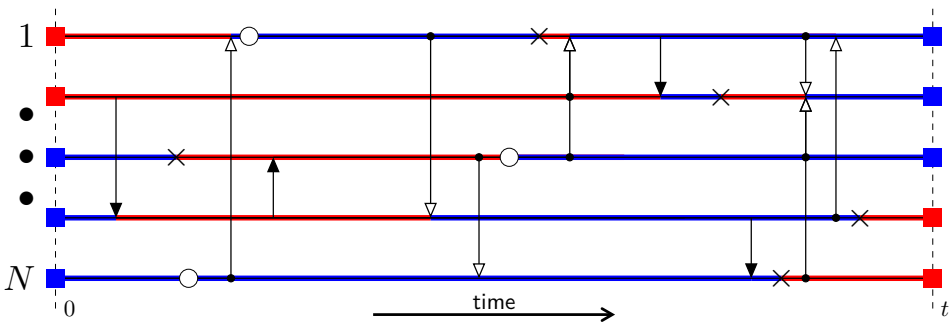
selective arrows of order m :

rate $s_m^{(N)}$,

mutation to type 1 (resp. 0):

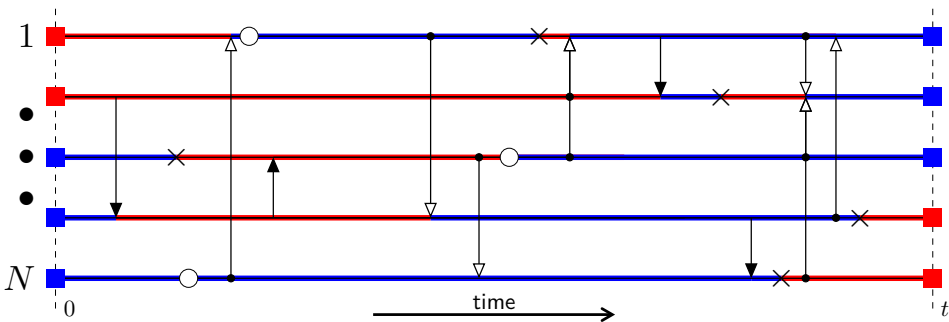
rate $u^{(N)}\nu_1$ (resp. $u^{(N)}\nu_0$), ($\nu_0 + \nu_1 = 1$).

Fittest-type-wins Moran model (FTW)



$Y_t^{(N)}$:= proportion of individuals of type 1 at time t in FTW MoMo of size N

Fittest-type-wins Moran model (FTW)

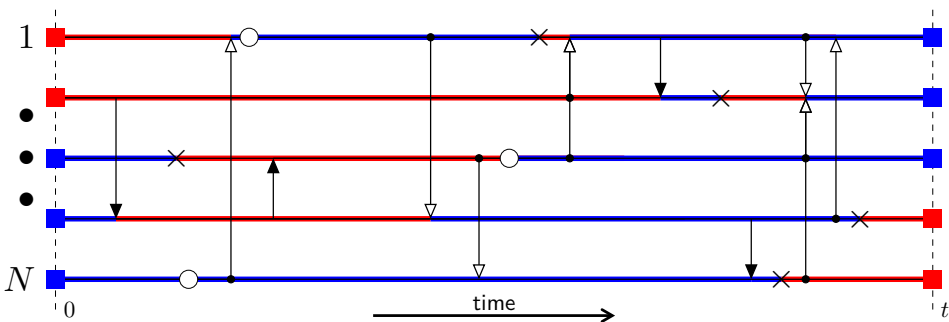


$Y_t^{(N)}$:= proportion of individuals of type 1 at time t in FTW MoMo of size N

$$q_Y^{(N)}\left(\frac{k}{N}, \frac{k+1}{N}\right) =$$

$$q_Y^{(N)}\left(\frac{k}{N}, \frac{k-1}{N}\right) =$$

Fittest-type-wins Moran model (FTW)

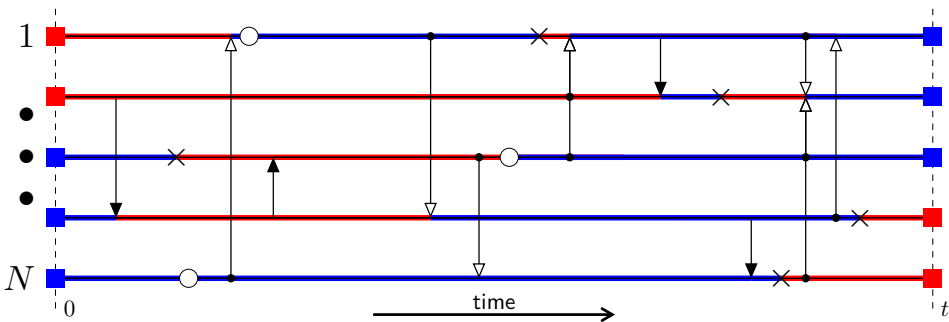


$Y_t^{(N)}$:= proportion of individuals of type 1 at time t in FTW MoMo of size N

$$q_Y^{(N)}\left(\frac{k}{N}, \frac{k+1}{N}\right) = k \frac{N-k}{N}$$

$$q_Y^{(N)}\left(\frac{k}{N}, \frac{k-1}{N}\right) = (N-k) \frac{k}{N}$$

Fittest-type-wins Moran model (FTW)

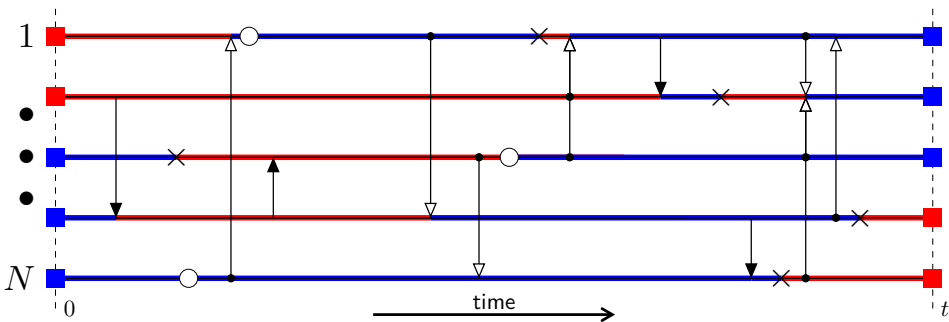


$Y_t^{(N)}$:= proportion of individuals of type 1 at time t in FTW MoMo of size N

$$q_Y^{(N)}\left(\frac{k}{N}, \frac{k+1}{N}\right) = k \frac{N-k}{N}$$

$$q_Y^{(N)}\left(\frac{k}{N}, \frac{k-1}{N}\right) = (N-k) \frac{k}{N} + k \sum_{m>0} s_m^{(N)}$$

Fittest-type-wins Moran model (FTW)

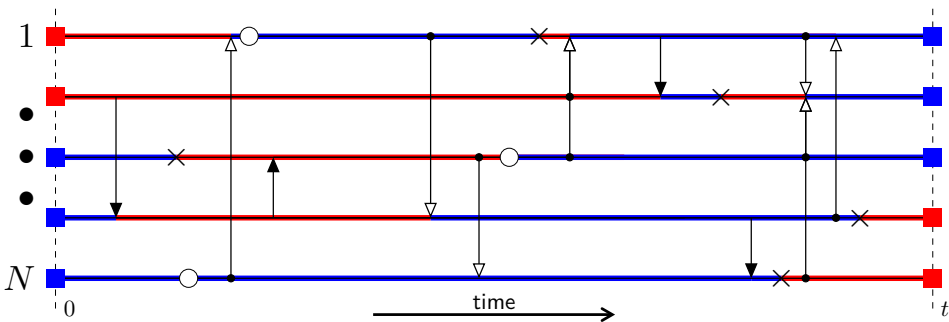


$Y_t^{(N)}$:= proportion of individuals of type 1 at time t in FTW MoMo of size N

$$q_Y^{(N)}\left(\frac{k}{N}, \frac{k+1}{N}\right) = k \frac{N-k}{N}$$

$$q_Y^{(N)}\left(\frac{k}{N}, \frac{k-1}{N}\right) = (N-k) \frac{k}{N} + k \sum_{m>0} s_m^{(N)} \left(1 - \left(\frac{k}{N}\right)^m\right)$$

Fittest-type-wins Moran model (FTW)



$Y_t^{(N)}$:= proportion of individuals of type 1 at time t in FTW MoMo of size N

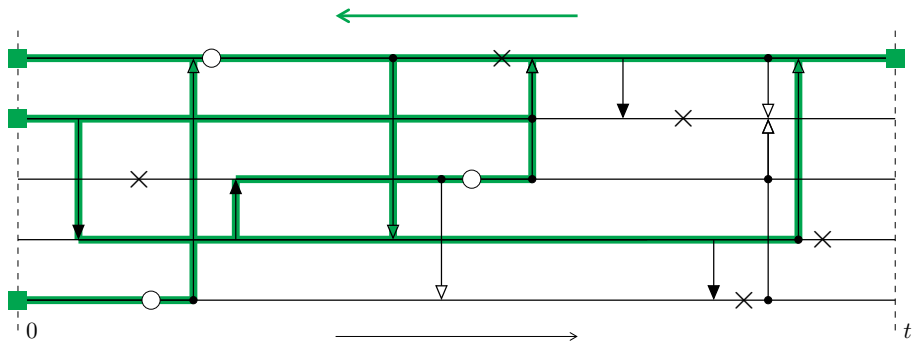
$$q_Y^{(N)}\left(\frac{k}{N}, \frac{k+1}{N}\right) = k \frac{N-k}{N} + (N-k)u^{(N)}\nu_1$$

$$q_Y^{(N)}\left(\frac{k}{N}, \frac{k-1}{N}\right) = (N-k)\frac{k}{N} + k \sum_{m>0} s_m^{(N)} \left(1 - \left(\frac{k}{N}\right)^m\right) + ku^{(N)}\nu_0$$

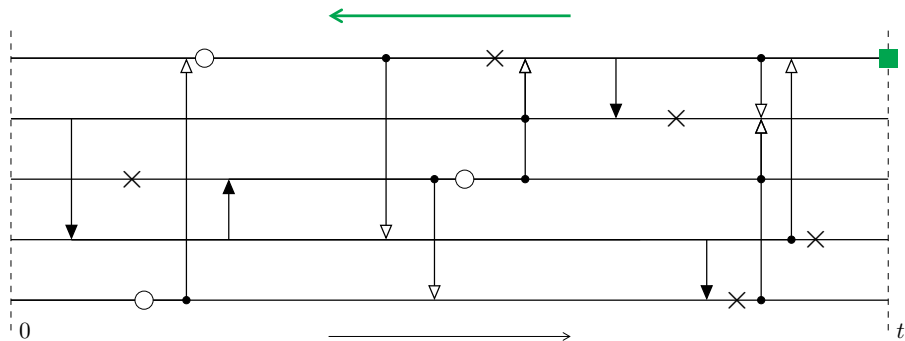
Fact

The birth-death processes \hat{Y} and Y (of DOM and FTW, respectively) are identical in distribution provided $\hat{Y}_0 = Y_0$ and $s_m = \hat{s}_m - \hat{s}_{m+1}$ for $m > 0$ (or, equivalently, $\hat{s}_m = \sum_{n \geq m} s_n$).

- **Question:** what is the stationary behaviour of the type-frequency of the population?



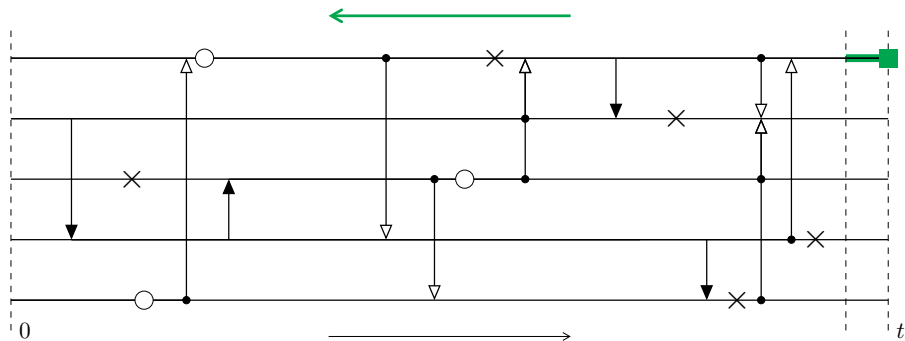
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = 1$$

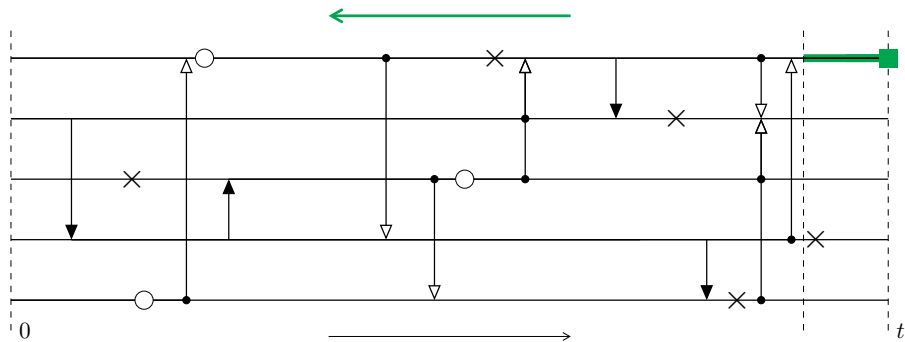
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = 1$$

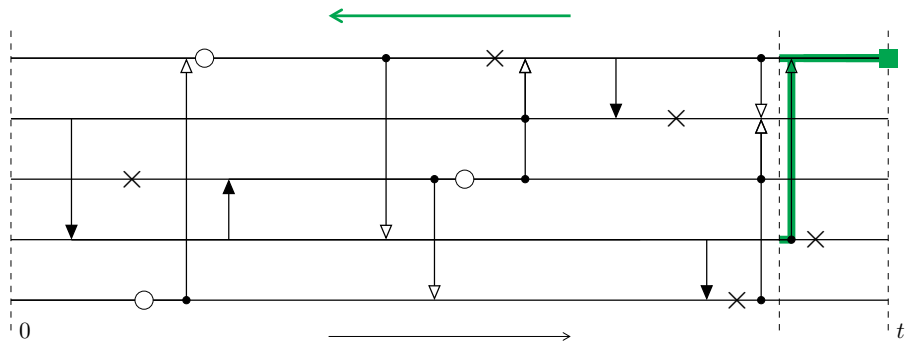
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = 1$$

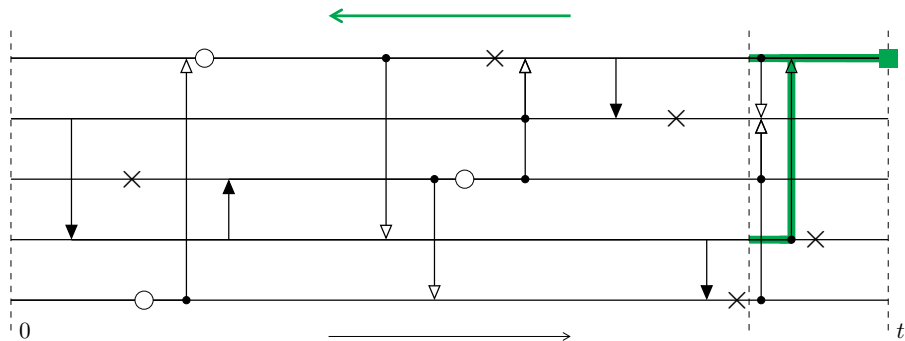
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = 2$$

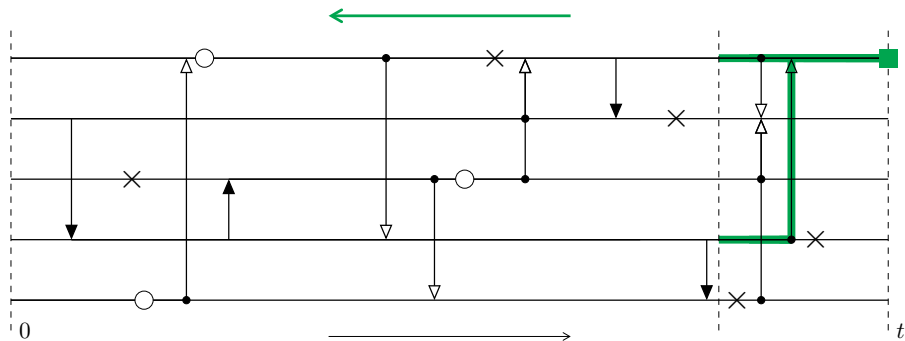
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = 2$$

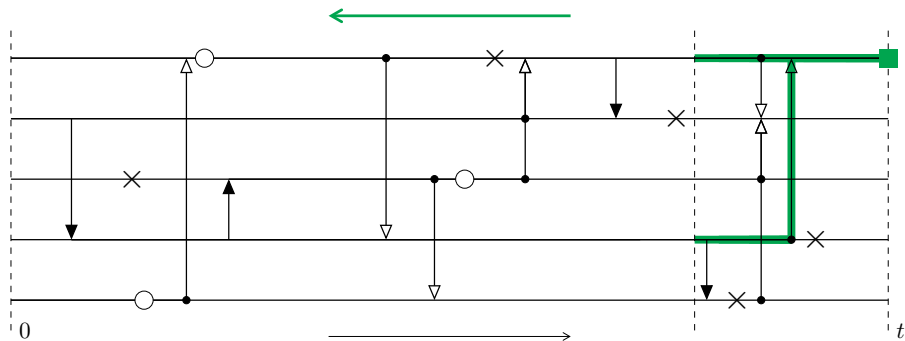
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = 2$$

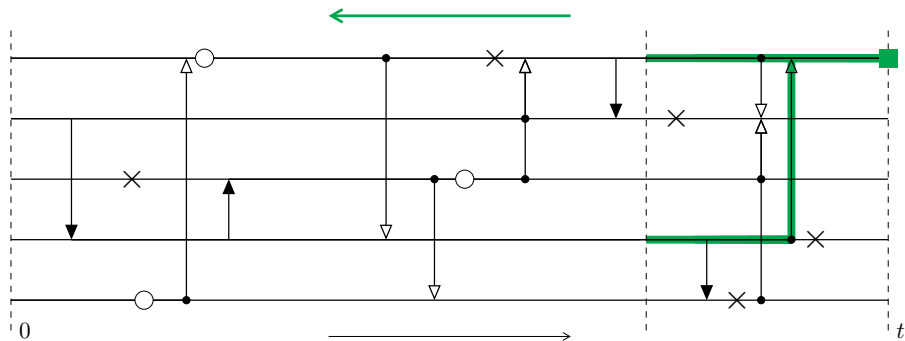
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = 2$$

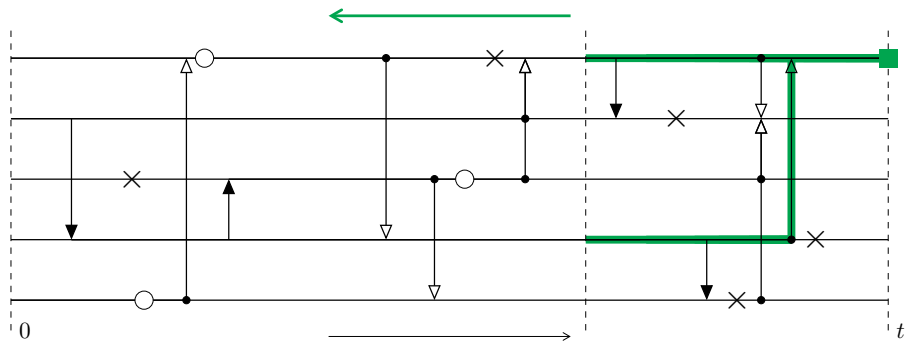
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = 2$$

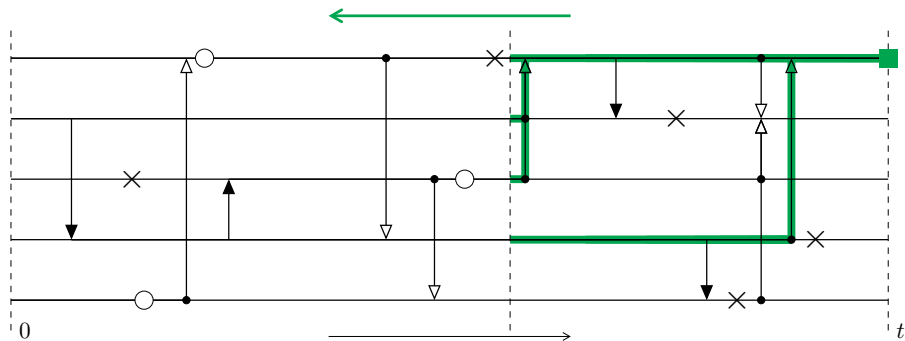
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = 2$$

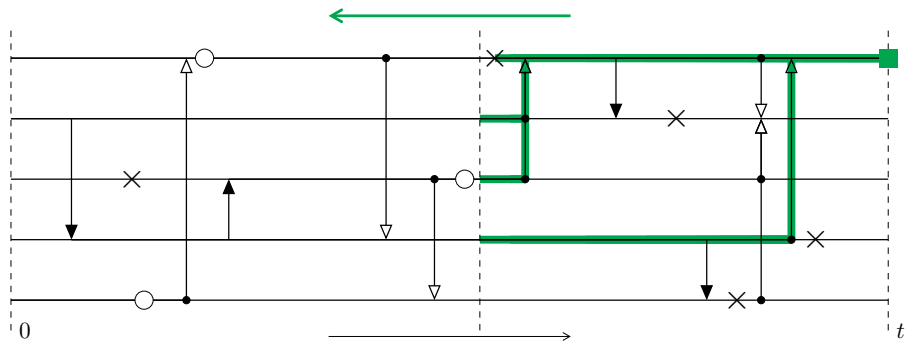
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = 4$$

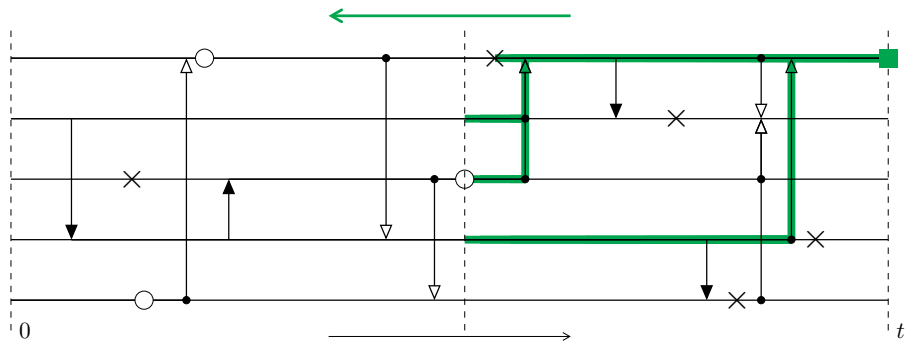
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = 3$$

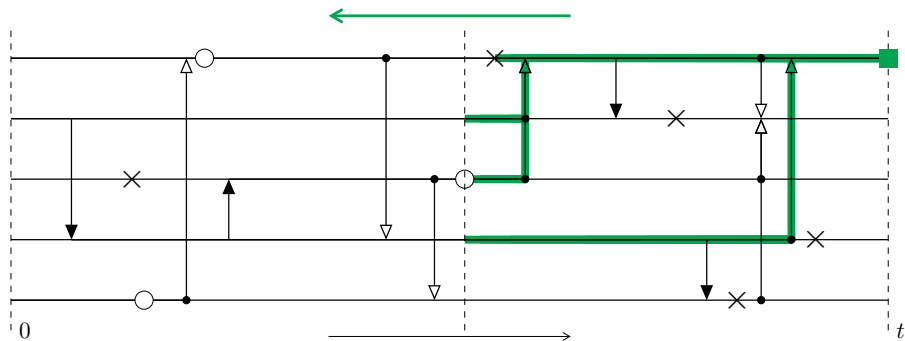
killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r =$$

killed ASG (kASG)



$R_r := \#$ of lines at backwards time r in the kASG

$$R_r = \Delta$$

Factorial-moment duality

- R is a continuous-time Markov chain.

Theorem (Factorial-moment duality)

Let Y be the line-counting process of the FTW MoMo and R the line-counting process of the k ASG. Then $\forall t \geq 0$, $n \in \mathbb{N}_0 \cup \Delta$, and $k \in [N]_0$,

$$\mathbb{E}\left[\frac{Y_t^n}{N^n} \mid Y_0 = k\right] = \mathbb{E}\left[\frac{k^{R_t}}{N^{R_t}} \mid R_0 = n\right], \quad (k^\Delta/N^\Delta := 0, \forall k)$$

That is, Y and R are dual w.r.t. the duality function

$$H_F(k, n) := \frac{k^n}{N^n}.$$

$$\mathbb{E}\left[\frac{Y_t^n}{N^n} \mid Y_0 = k\right] = \mathbb{E}\left[\frac{k^{R_t}}{N^{R_t}} \mid R_0 = n\right]$$

- If $u = 0$, Y absorbs almost surely in $\{0, N\}$, while R converges to a stationary distribution π_R .
- If $u > 0$, Y converges to a stationary distribution π_Y , while R absorbs almost surely in $\{0, \Delta\}$.

$$\mathbb{E}\left[\frac{Y_t^n}{N^n} \mid Y_0 = k\right] = \mathbb{E}\left[\frac{k^{R_t}}{N^{R_t}} \mid R_0 = n\right]$$

- If $u = 0$, Y absorbs almost surely in $\{0, N\}$, while R converges to a stationary distribution π_R .
- If $u > 0$, Y converges to a stationary distribution π_Y , while R absorbs almost surely in $\{0, \Delta\}$.

Corollary (Representation of absorption probabilities)

Suppose $u = 0$. For $k \in [N]_0$,

$$\mathbb{P}\left(\lim_{t \rightarrow \infty} Y_t = N \mid Y_0 = k\right) = \mathbb{E}\left[\frac{k^{R_\infty}}{N^{R_\infty}}\right].$$

Suppose $u > 0$. For $n \in [N]_0 \cup \{\Delta\}$,

$$\mathbb{P}\left(\lim_{t \rightarrow \infty} R_t = 0 \mid R_0 = n\right) = \mathbb{E}\left[\frac{Y_\infty^n}{N^n}\right].$$

Conclusions

In short:

- Introduced two separate generalization of the MoMo with genic selection and mutation (DOM, FTW)

Conclusions

In short:

- Introduced two separate generalization of the MoMo with genic selection and mutation (DOM, FTW)
- Shown that the two models are equivalent when studying the behaviour of the type distributions

Conclusions

In short:

- Introduced two separate generalization of the MoMo with genic selection and mutation (DOM, FTW)
- Shown that the two models are equivalent when studying the behaviour of the type distributions
- Introduced the kASG to find a representation of the forward type distribution through duality

Further results

Further results:

- Introduced the pLD-ASG to find representation of the ancestral type distribution and the common ancestral type distribution

Further results

Further results:

- Introduced the pLD-ASG to find representation of the ancestral type distribution and the common ancestral type distribution
- Factorial moment dual, related to the forward MoMo for the line-counting process of the pld-ASG

Further results:

- Introduced the pLD-ASG to find representation of the ancestral type distribution and the common ancestral type distribution
- Factorial moment dual, related to the forward MoMo for the line-counting process of the pld-ASG
- Convergence of FTW MoMo and line-counting processes for $N \rightarrow \infty$ in the diffusion limit, with relative dualities

Further results

Further results:

- Introduced the pLD-ASG to find representation of the ancestral type distribution and the common ancestral type distribution
- Factorial moment dual, related to the forward MoMo for the line-counting process of the pld-ASG
- Convergence of FTW MoMo and line-counting processes for $N \rightarrow \infty$ in the diffusion limit, with relative dualities

[1] Baake, Ellen, Luigi Esercito, and Sebastian Hummel. "Lines of descent in a Moran model with frequency-dependent selection and mutation." *arXiv preprint arXiv:2011.08888* (2020).

Thank you for your attention!

Ancestral type distribution

Definition (Ancestral type distribution)

Let Γ be a random leaf colouring, and consider an ASG in the time interval $[0, r]$ for some $r > 0$. The *ancestral type distribution* at backward time r , conditional on a leaf-colouring with k type-1 individuals, is defined as

$$h_r(k) := \mathbb{P}(\text{the ancestor at time } r \text{ of an individual is of type } 1 \mid |\Gamma| = k),$$

where $|\Gamma| = \sum_{i \in [N]} \Gamma(i)$ and $k \in [N]_0$.

Ancestral type distribution

Definition (Ancestral type distribution)

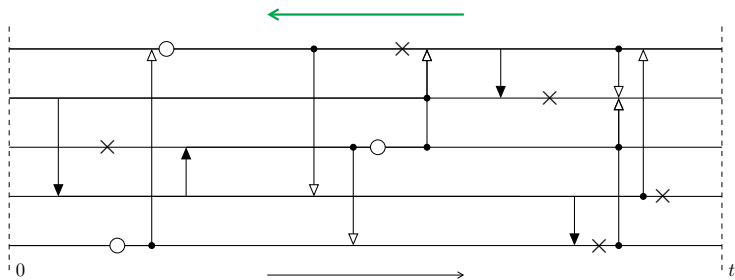
Let Γ be a random leaf colouring, and consider an ASG in the time interval $[0, r]$ for some $r > 0$. The *ancestral type distribution* at backward time r , conditional on a leaf-colouring with k type-1 individuals, is defined as

$$h_r(k) := \mathbb{P}(\text{the ancestor at time } r \text{ of an individual is of type 1} \mid |\Gamma| = k),$$

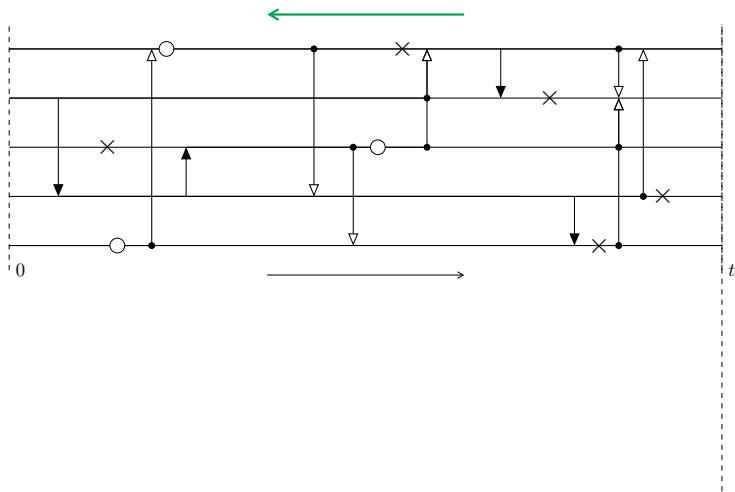
where $|\Gamma| = \sum_{i \in [N]} \Gamma(i)$ and $k \in [N]_0$.

- **Question:** what can we say about this distribution?

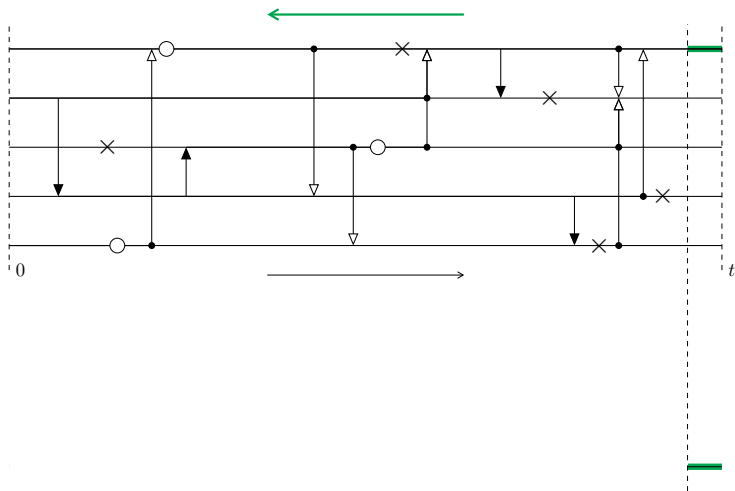
pruned lookahead ASG (pLD ASG)



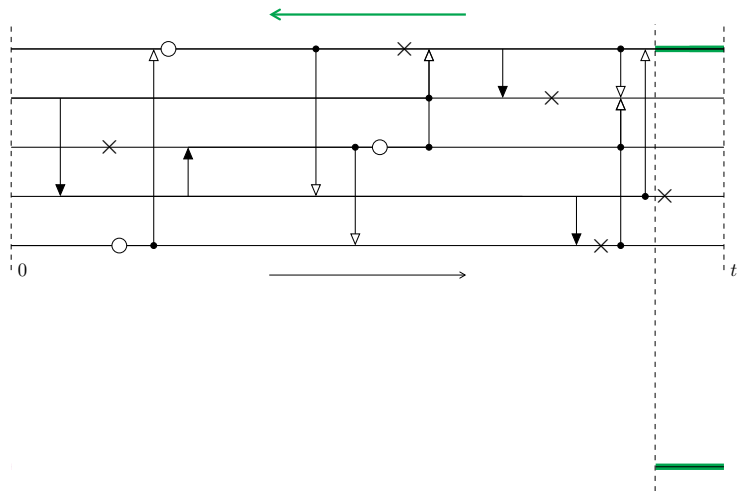
pruned lookahead ASG (pLD ASG)



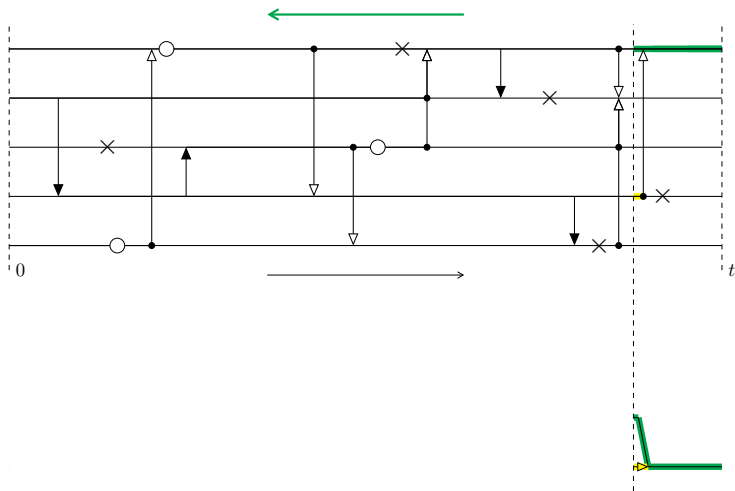
pruned lookahead ASG (pLD ASG)



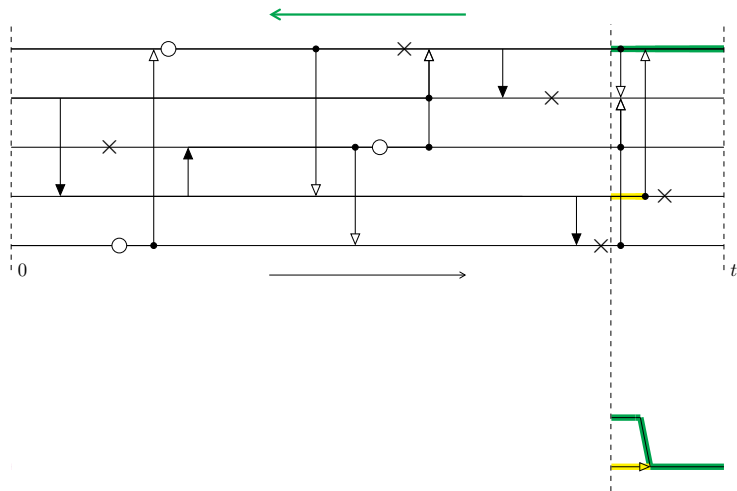
pruned lookahead ASG (pLD ASG)



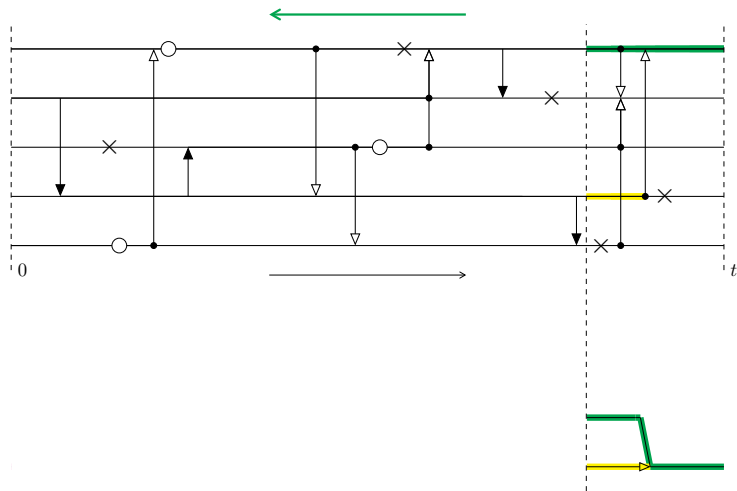
pruned lookahead ASG (pLD ASG)



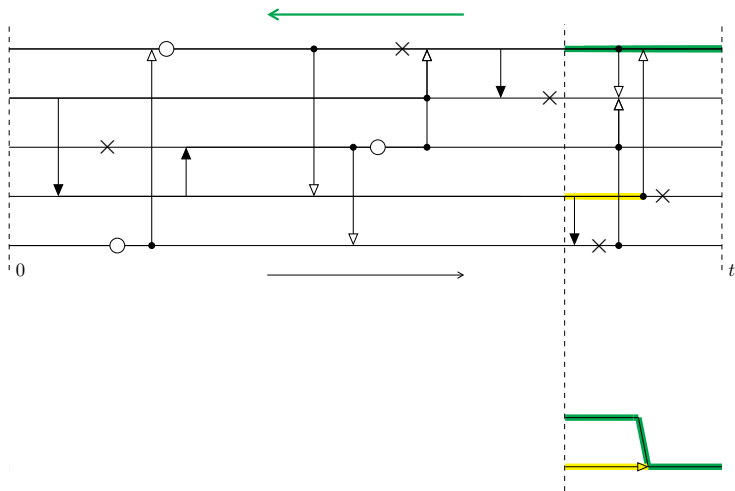
pruned lookahead ASG (pLD ASG)



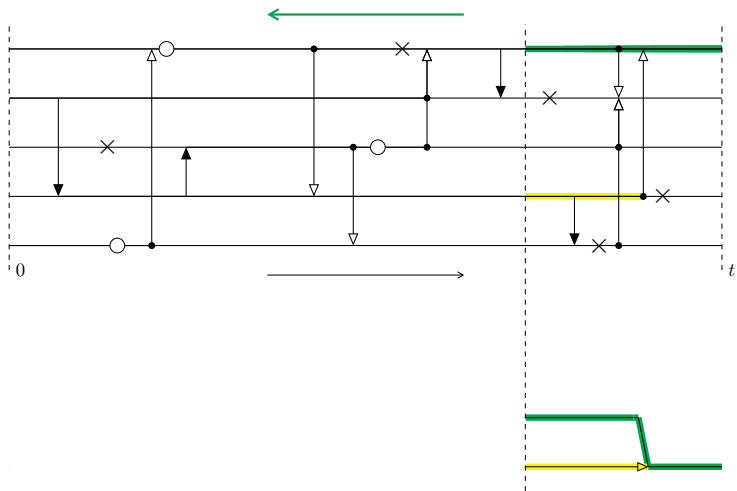
pruned lookahead ASG (pLD ASG)



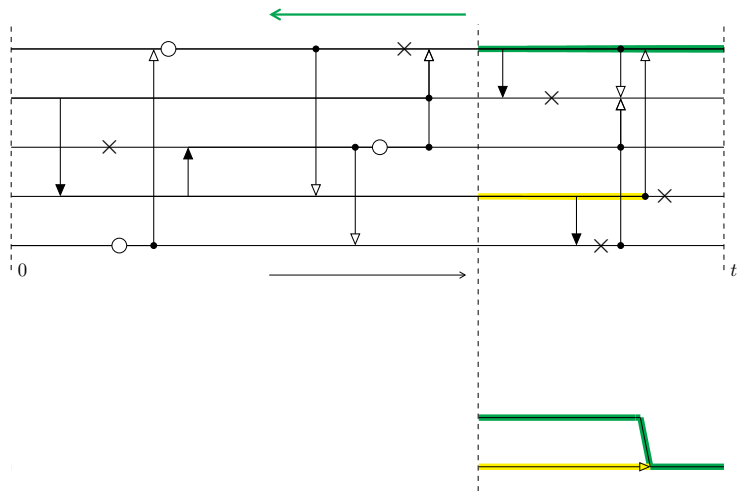
pruned lookahead ASG (pLD ASG)



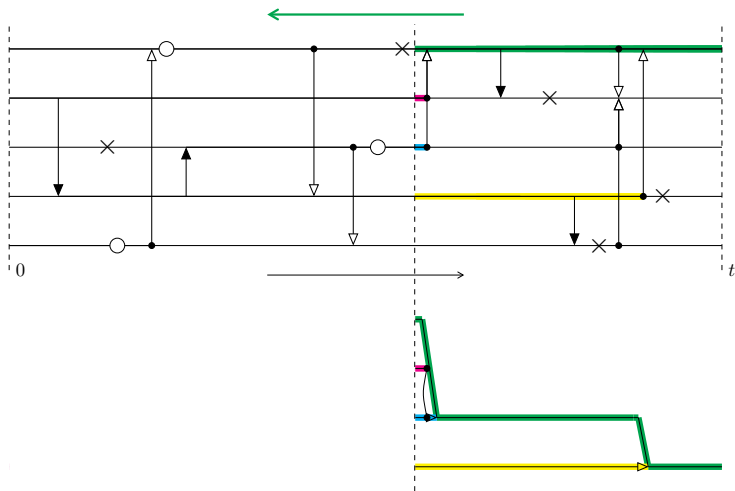
pruned lookahead ASG (pLD ASG)



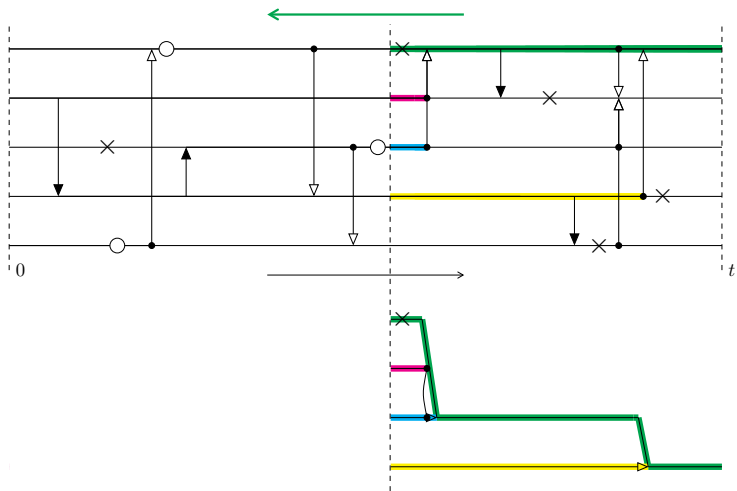
pruned lookahead ASG (pLD ASG)



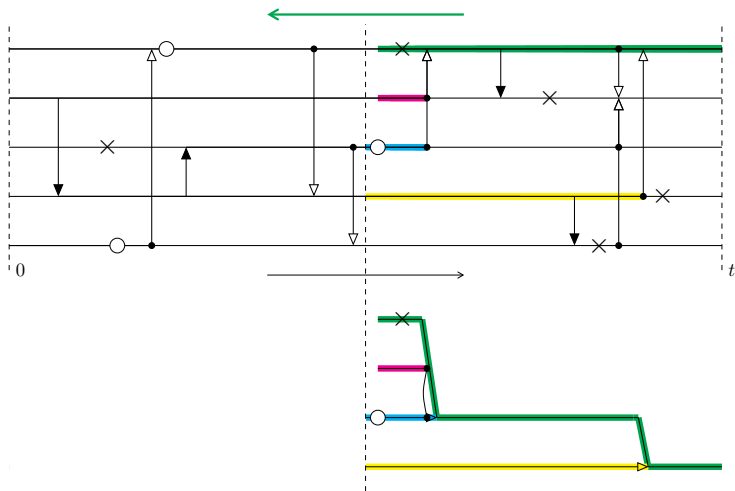
pruned lookahead ASG (pLD ASG)



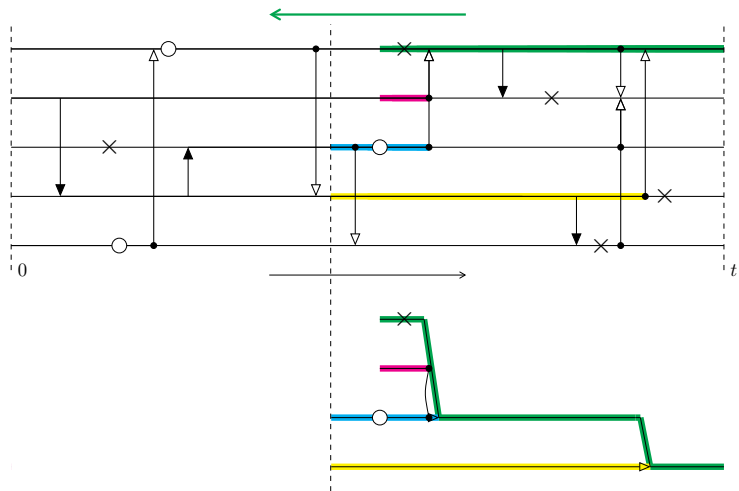
pruned lookahead ASG (pLD ASG)



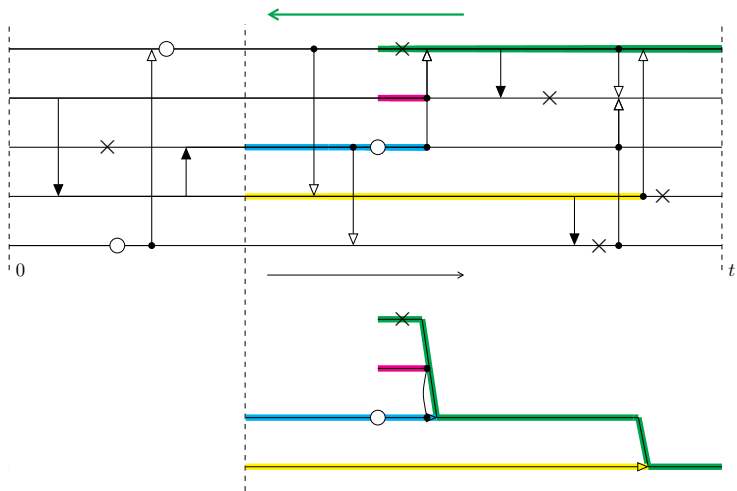
pruned lookahead ASG (pLD ASG)



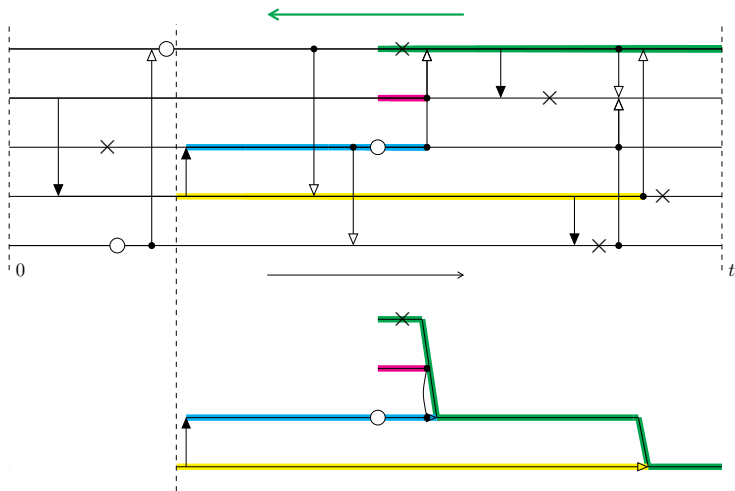
pruned lookahead ASG (pLD ASG)



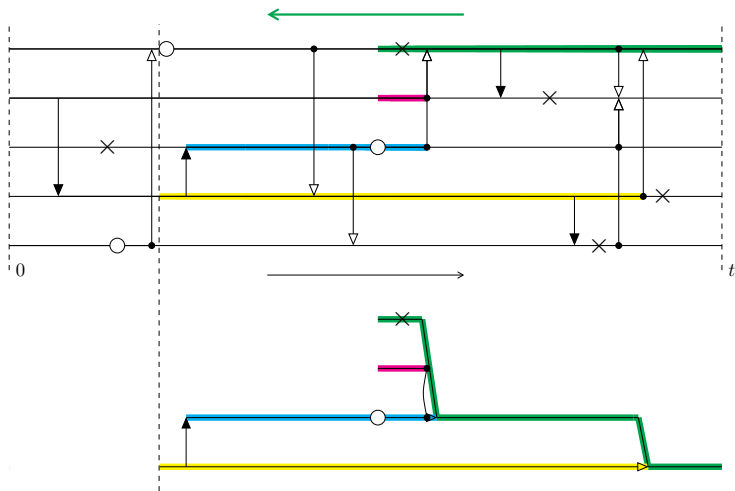
pruned lookahead ASG (pLD ASG)



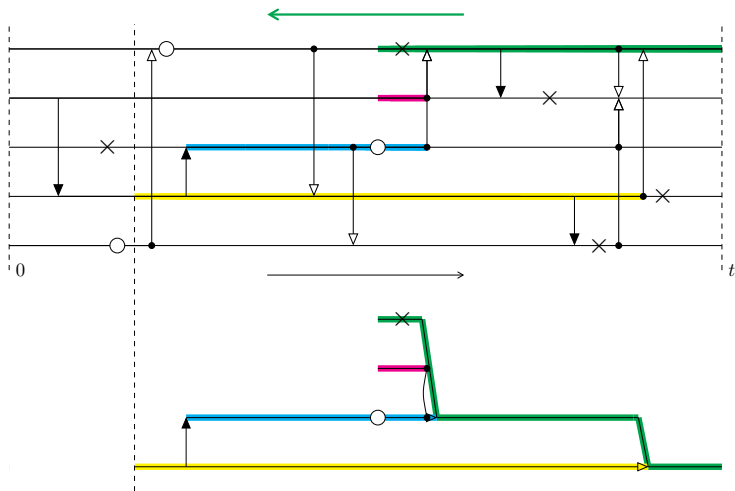
pruned lookahead ASG (pLD ASG)



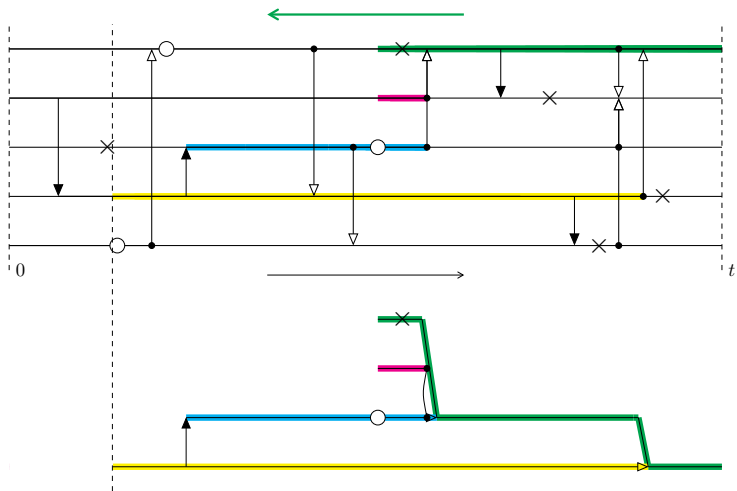
pruned lookahead ASG (pLD ASG)



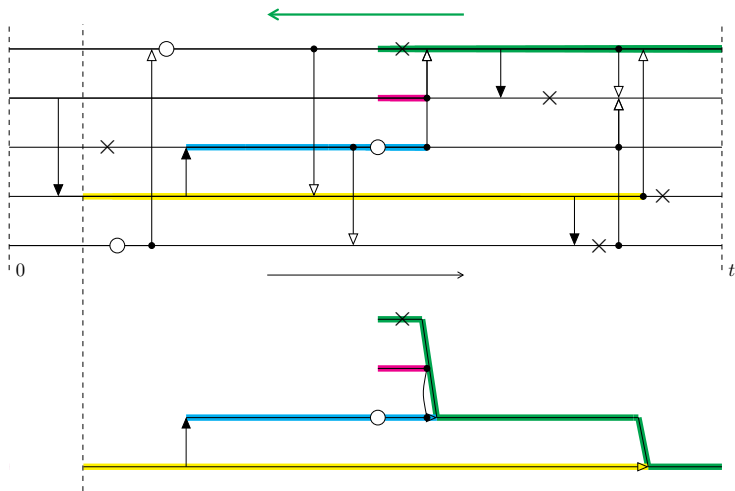
pruned lookahead ASG (pLD ASG)



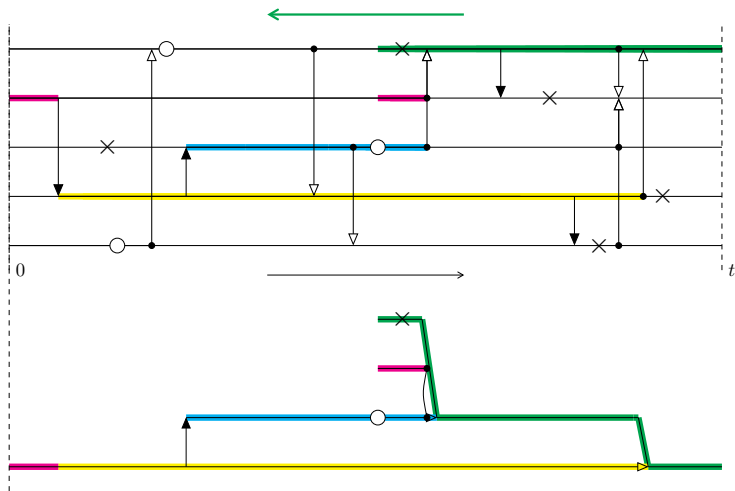
pruned lookahead ASG (pLD ASG)



pruned lookahead ASG (pLD ASG)

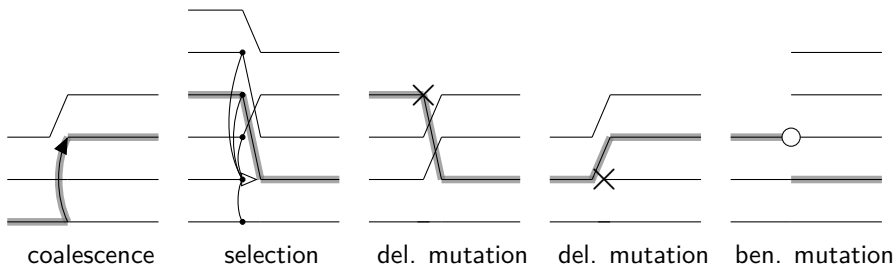


pruned lookahead ASG (pLD ASG)

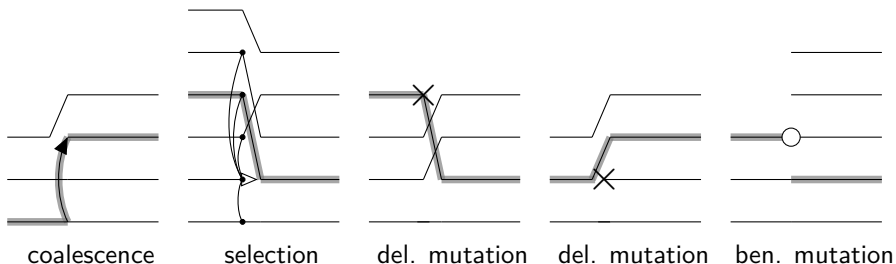


- The pLD-ASG is described by the *Level-process* $\ell = (\ell_r)$ and the *immune-line process* $i = (i_r)$

- Transitions:



- The pLD-ASG is described by the *Level-process* $\ell = (\ell_r)$ and the *immune-line process* $i = (i_r)$
- Transitions:



Theorem

The level of the ancestral line at time r is either the lowest finite level that has type 0 at time r ; or, if all finite levels are of type 1, it is $i(r)$, the level of the immune line at time r . In particular, the ancestral site is of type 1 at time r if and only if all lines at finite levels are of type 1 at time r .

- The line-counting process $L = (L_r)_r$ of the pLD-ASG is a continuous-time Markov chain.

- The line-counting process $L = (L_r)_r$ of the pLD-ASG is a continuous-time Markov chain.

Theorem (Representation of ancestral type distribution)

We have

$$h_r(k) = \mathbb{E}_1 \left[\frac{k^{L_r}}{N^{L_r}} \right], \quad k \in [N]_0.$$

Furthermore, $h_\infty(k) = \lim_{r \rightarrow \infty} h_r(k)$ exists and is given by

$$h_\infty(k) = \mathbb{E} \left[\frac{k^{L_\infty}}{N^{L_\infty}} \right], \quad k \in [N]_0.$$

- The line-counting process $L = (L_r)_r$ of the pLD-ASG is a continuous-time Markov chain.

Theorem (Representation of ancestral type distribution)

We have

$$h_r(k) = \mathbb{E}_1 \left[\frac{k^{L_r}}{N^{L_r}} \right], \quad k \in [N]_0.$$

Furthermore, $h_\infty(k) = \lim_{r \rightarrow \infty} h_r(k)$ exists and is given by

$$h_\infty(k) = \mathbb{E} \left[\frac{k^{L_\infty}}{N^{L_\infty}} \right], \quad k \in [N]_0.$$

- h_∞ is the *common ancestor type distribution*.

