Instituto de Matemáticas

Adrián González Casanova

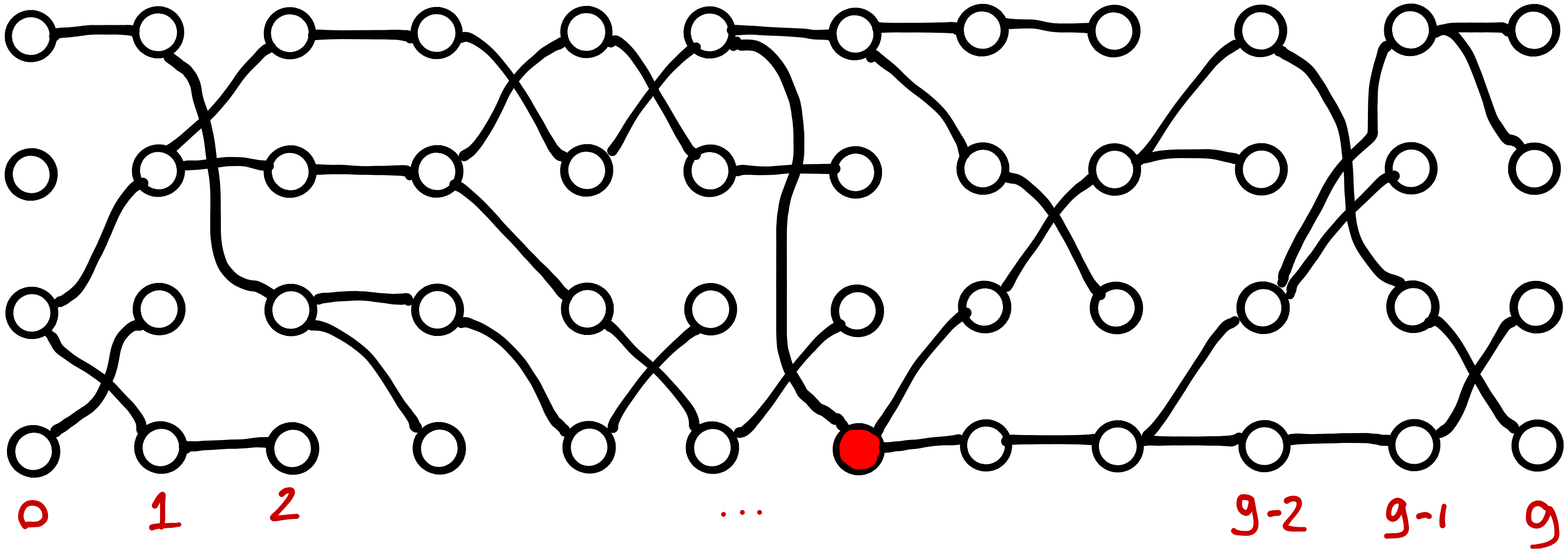
Institute of Mathematics

National Autonomous University of Mexico

UNAM

BIC Chile

The Wright Fisher Random Graph



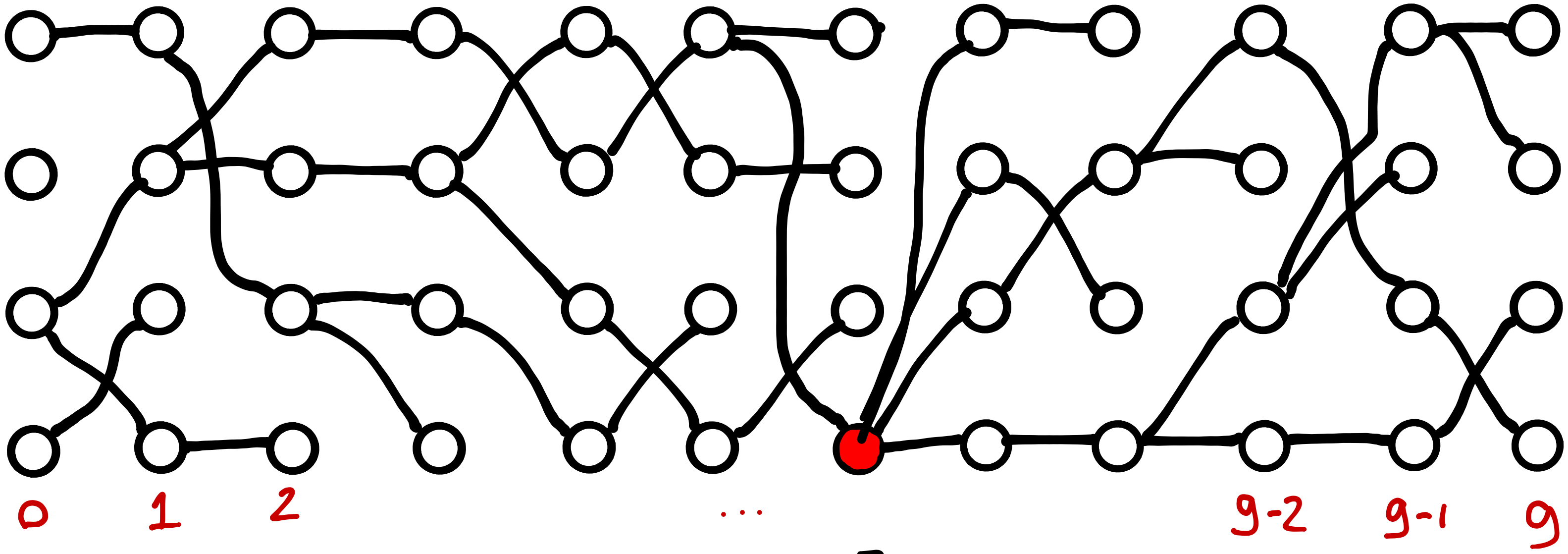
$$V = \mathbb{Z} \times [N]$$

Let $\{U_v\}_{v \in V}$ be a family of IID uniform random variables

the (random) set of edges is

$$E = \{ ((g,i),(g-1,U_{(g,i)})) : (g,i) \in V \}$$

The Paintbox Random Graph



$V = \mathbb{Z} \times [N]$

Let $\{W_v\}_{v \in V}$ be R.V. such that $\sum_{v \in gen(g)} W_v = 1$ and $W_v \overset{d}{=} W_{v'}$

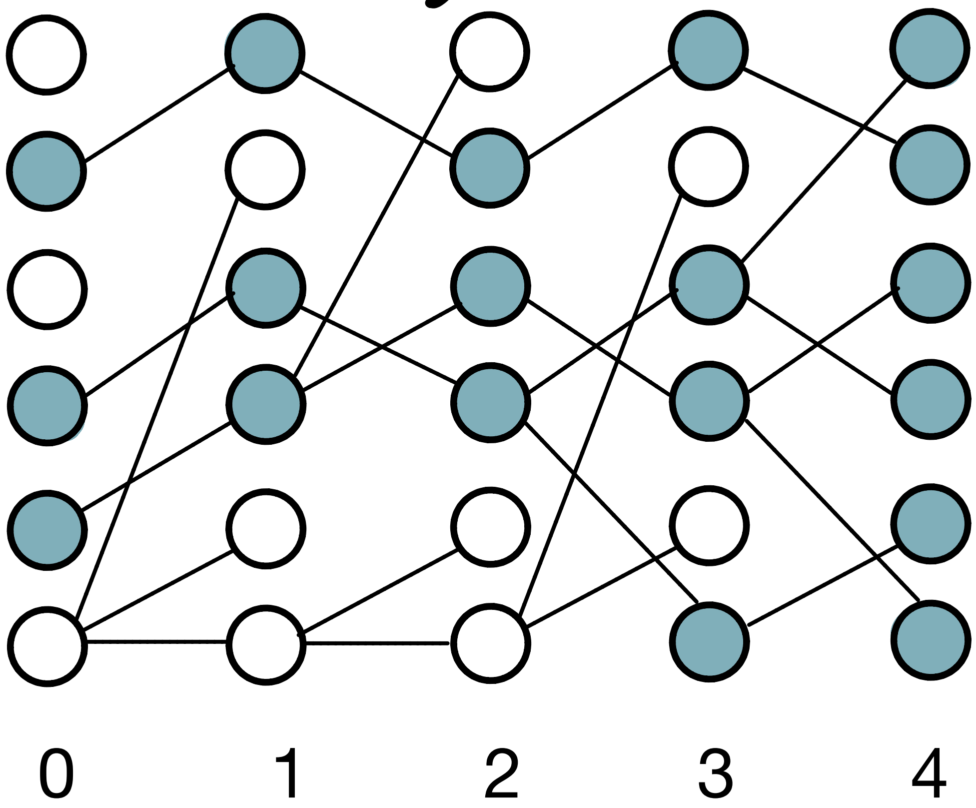
Let $\{U_v\}_{v \in V}$ be a family of uniform random variables such that

$$\mathbb{P}(U_{(i,g)} = j) = W_{(j,g-1)}$$

The (random) set of edges is

$$E = \{((g,i),(g-1,U_{(g,i)})) : (g,i) \in V\}$$

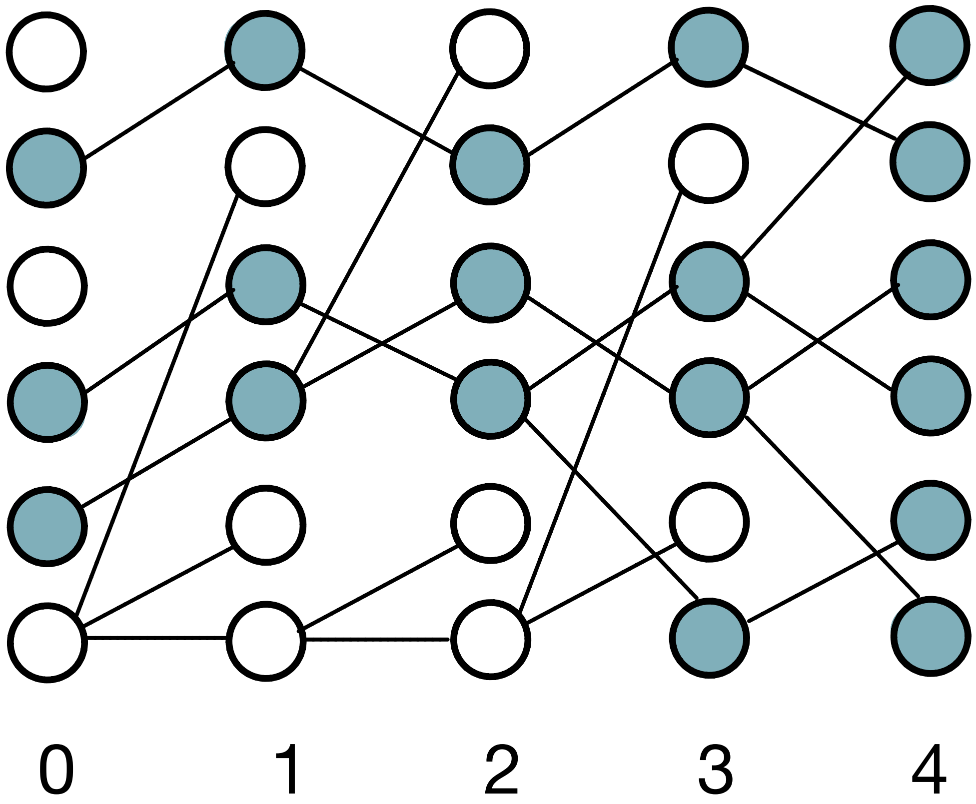
Wright-Fisher



$$X_g^N = \frac{\# \bullet \text{ in generation } g}{N}$$

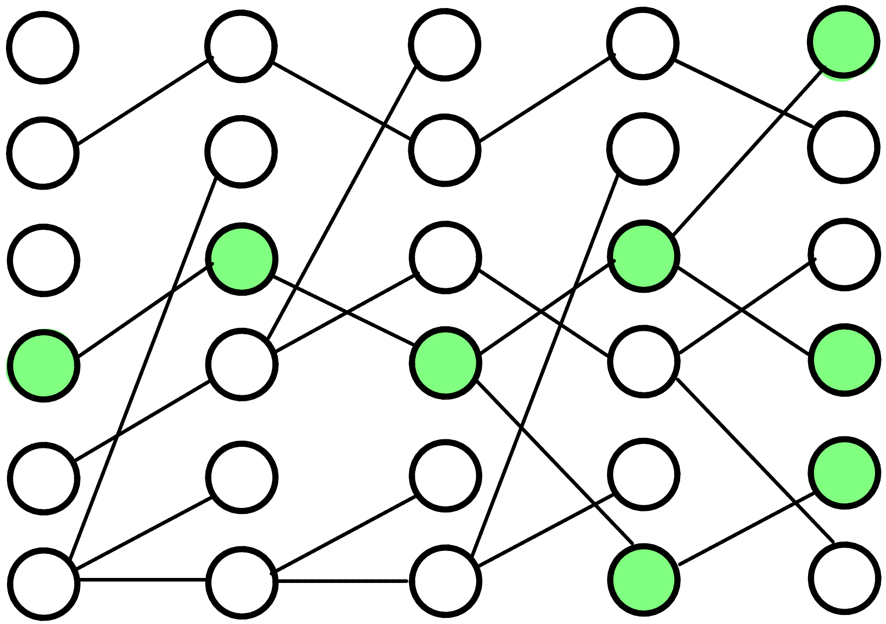
given X_g, $\quad X_{g+1} \sim \dfrac{\text{Binomial}(N, X_g)}{N}$

Paintbox



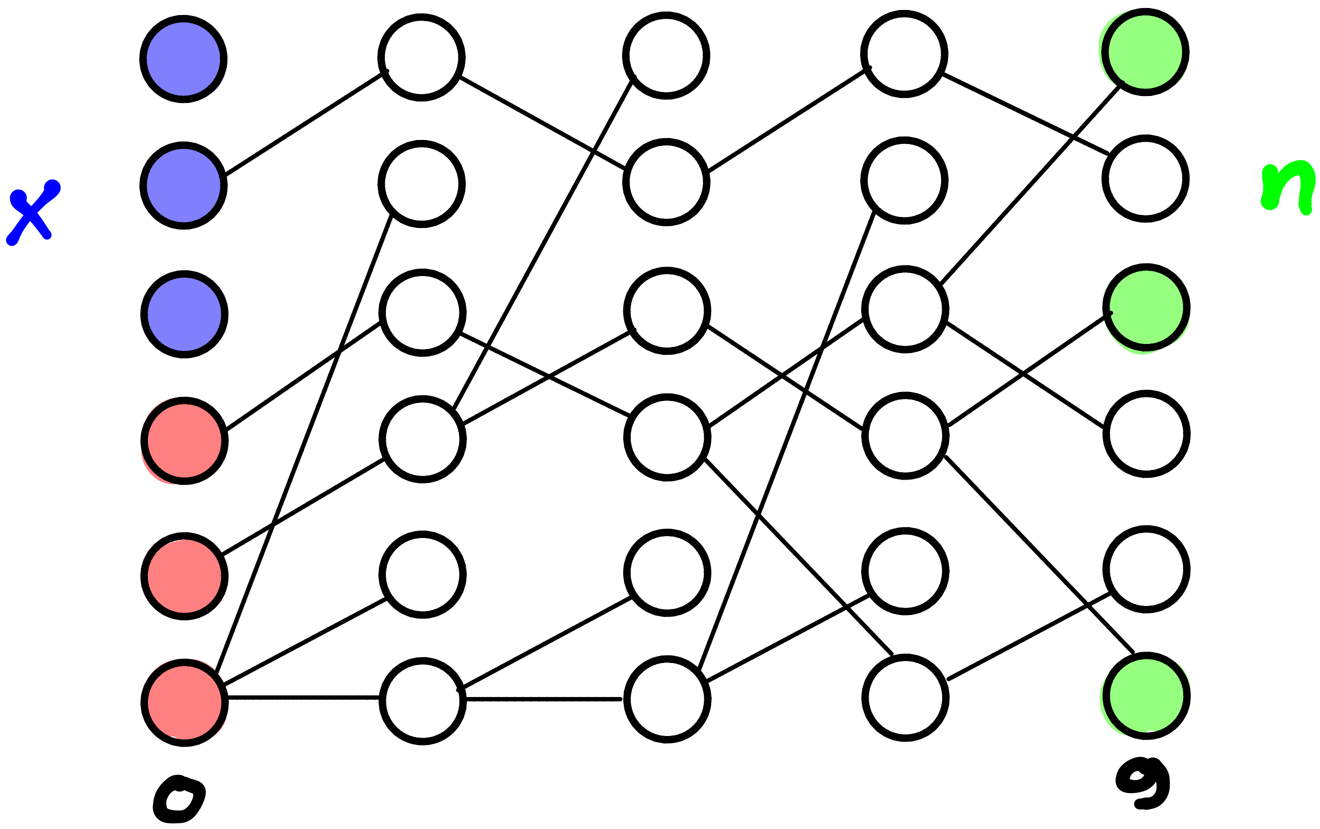
$$X_g^N = \frac{\# \bullet \text{ in generation } g}{N}$$

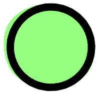
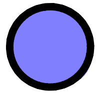
given $P(g) = \sum_{i=1}^{N} w_{(g,i)} \mathbb{1}\{\tau(g,i) = \bullet\}$, $X_{g+1} \sim \dfrac{\text{Binomial}(N, P(g))}{N}$

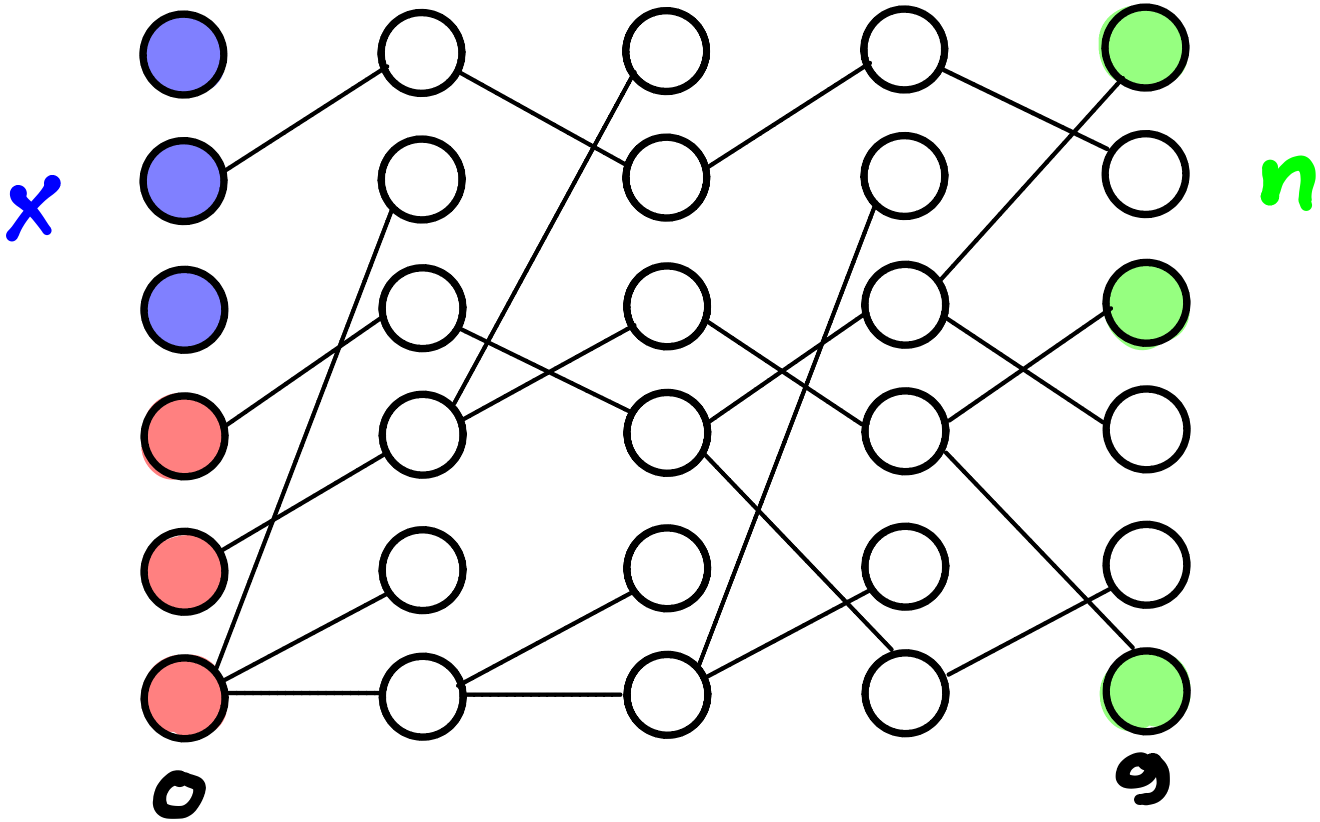


$A_g^N = \#$ ● in gen. g

What is the Probability that all are ? (P)

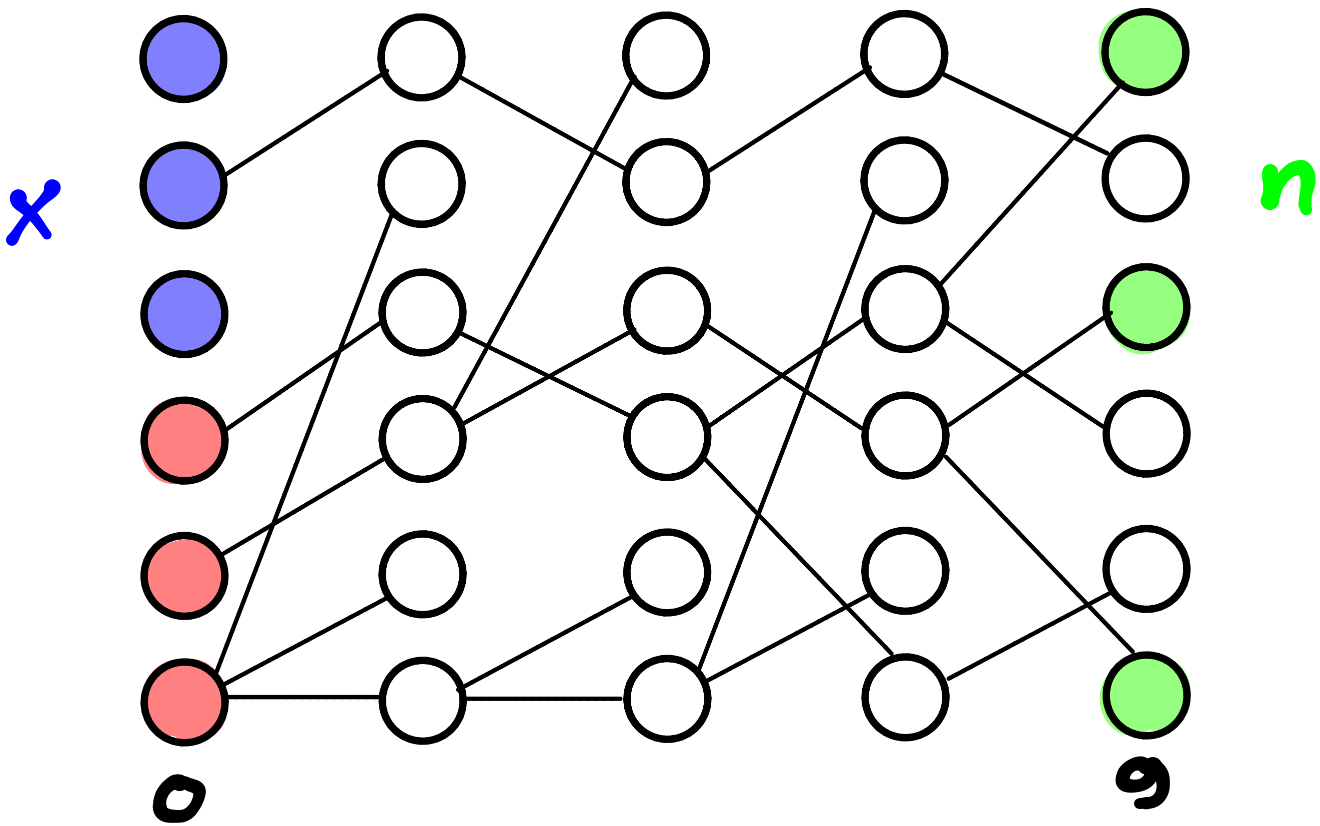


What is the Probability that all  are ? (P)

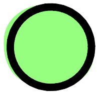
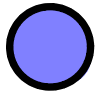


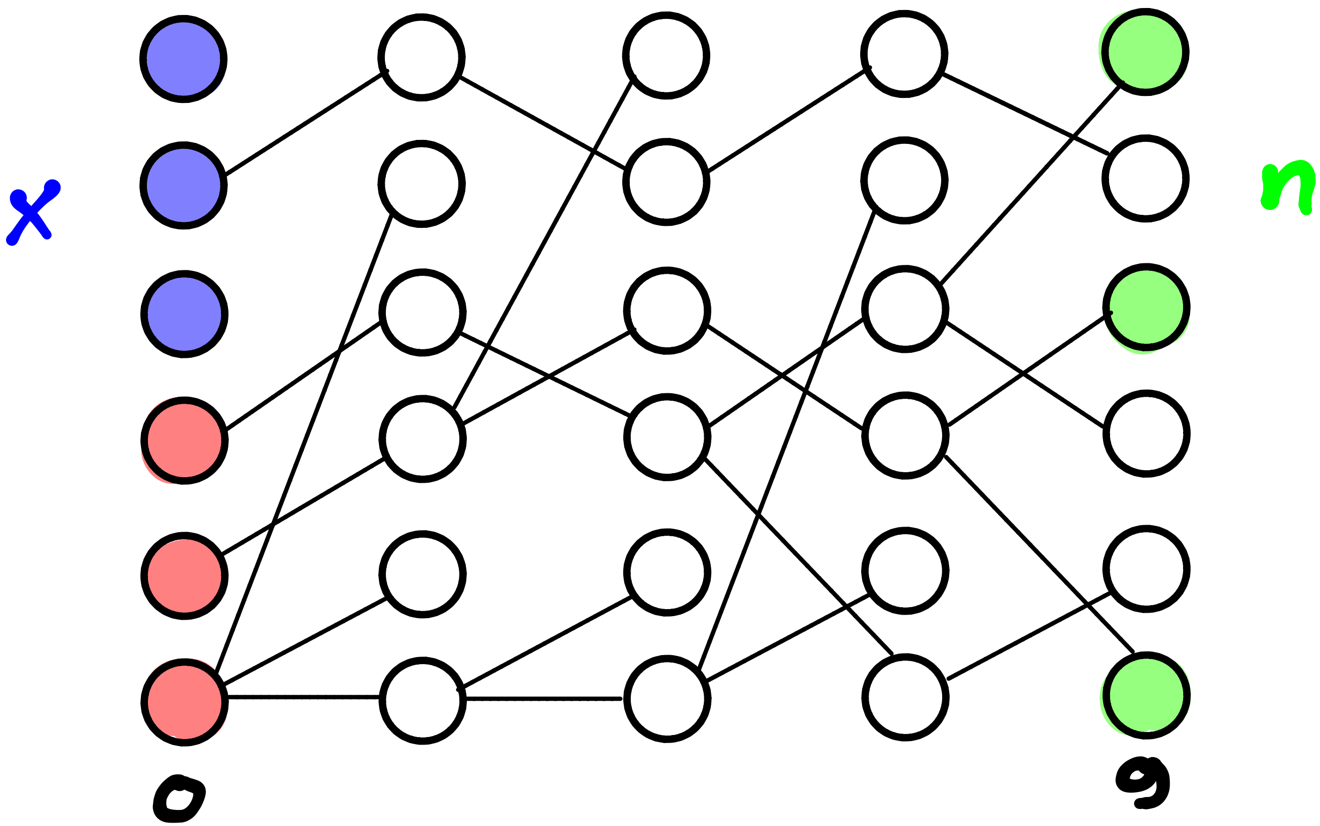
• Given $NX_9 = y$, $P = \frac{n}{N} \cdot \frac{n-1}{N-1} \cdot \ldots \cdot \frac{n-y}{N-y}$

What is the Probability that all are ? (P)



·· Given $A_9 = m$, $P = X \cdot \frac{NX-1}{N-1} \cdot \ldots \cdot \frac{NX-m}{X-m}$

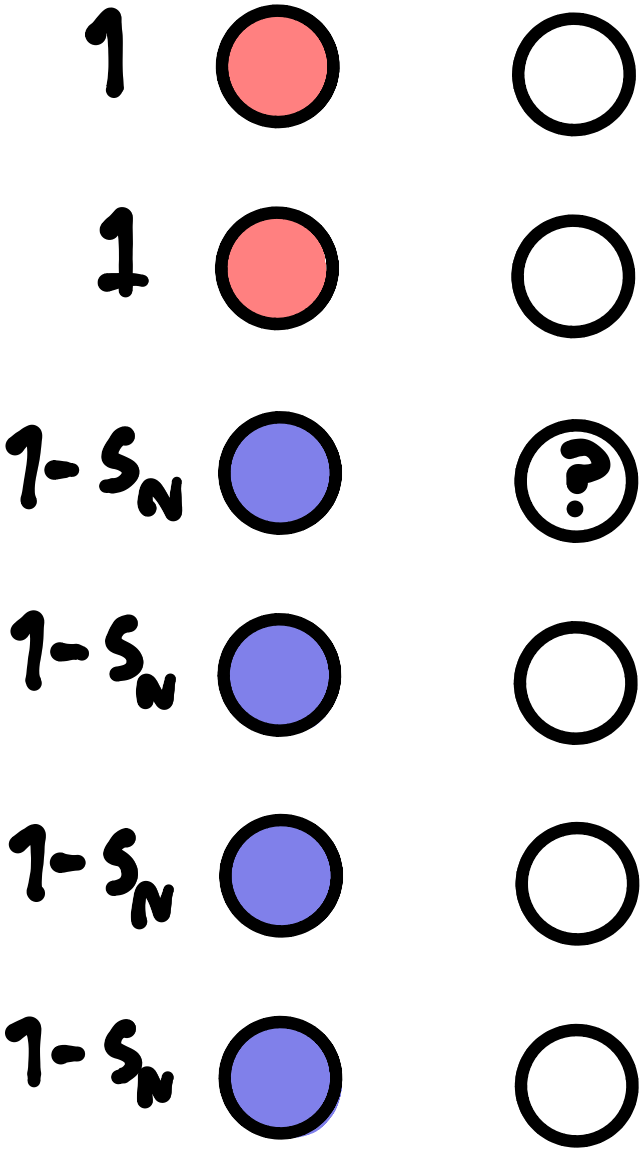
What is the Probability that all  are ? (P)



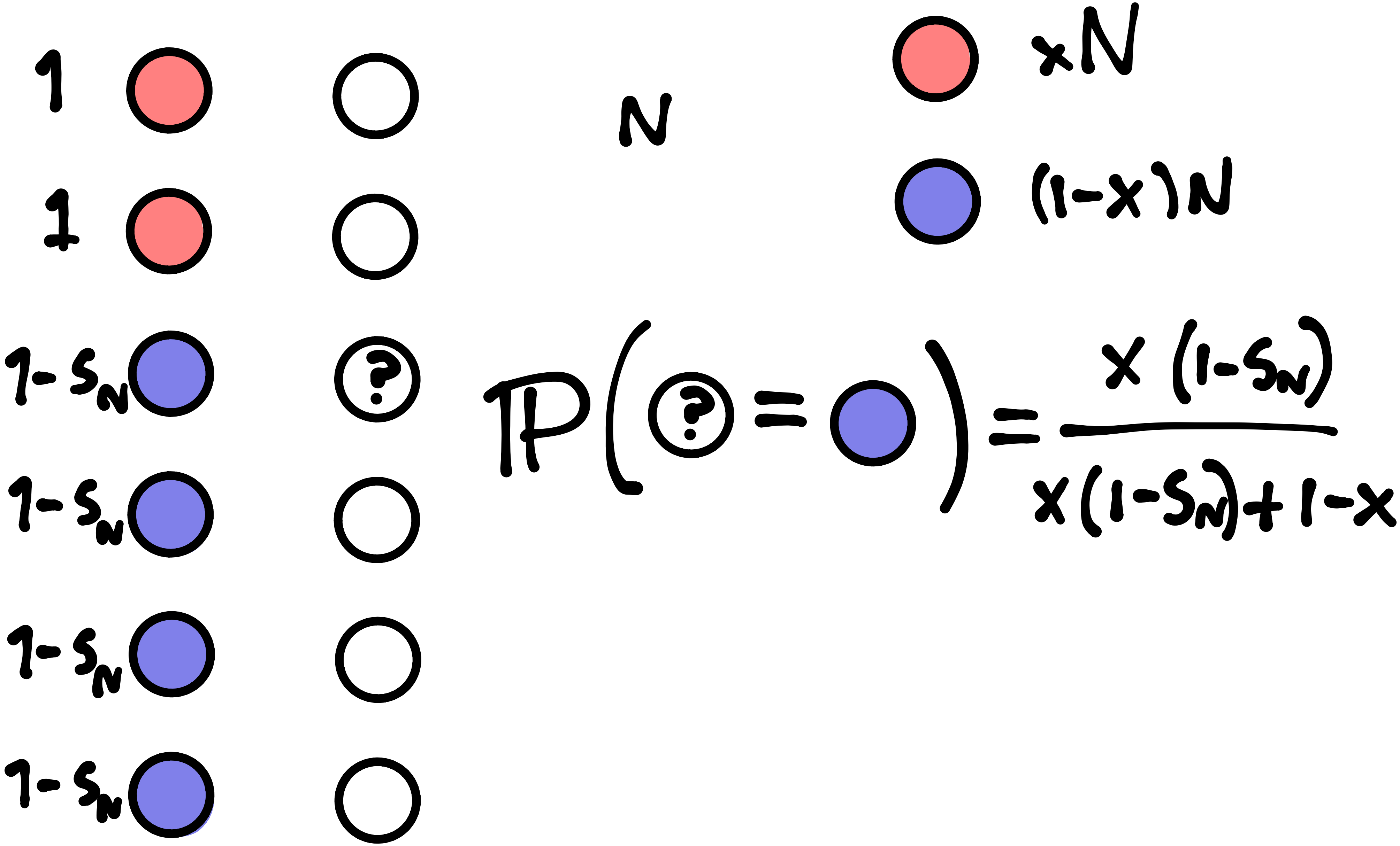
$$\#_n \left[x \cdot \frac{Nx-1}{N-1} \cdots \frac{Nx-A_g}{N-A_g} \right] = \#_x \left[\frac{n}{N} \cdot \frac{n-1}{N-1} \cdot \ldots \cdot \frac{n-NX_g}{N-NX_g} \right]$$

In the Wright Fisher model

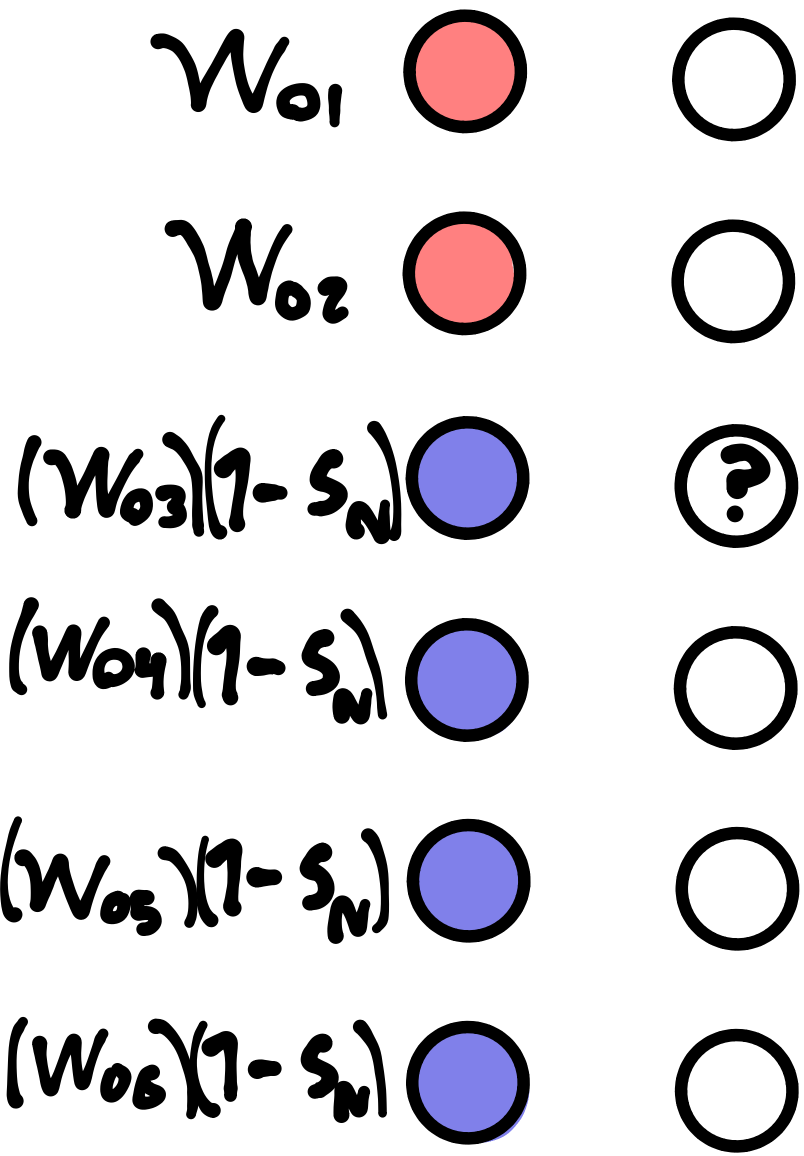
Let $s_N \leq 1$



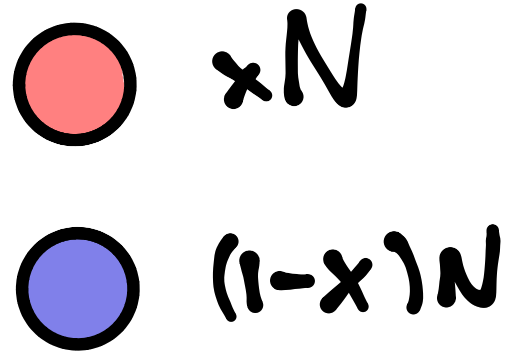
In the Wright Fisher model



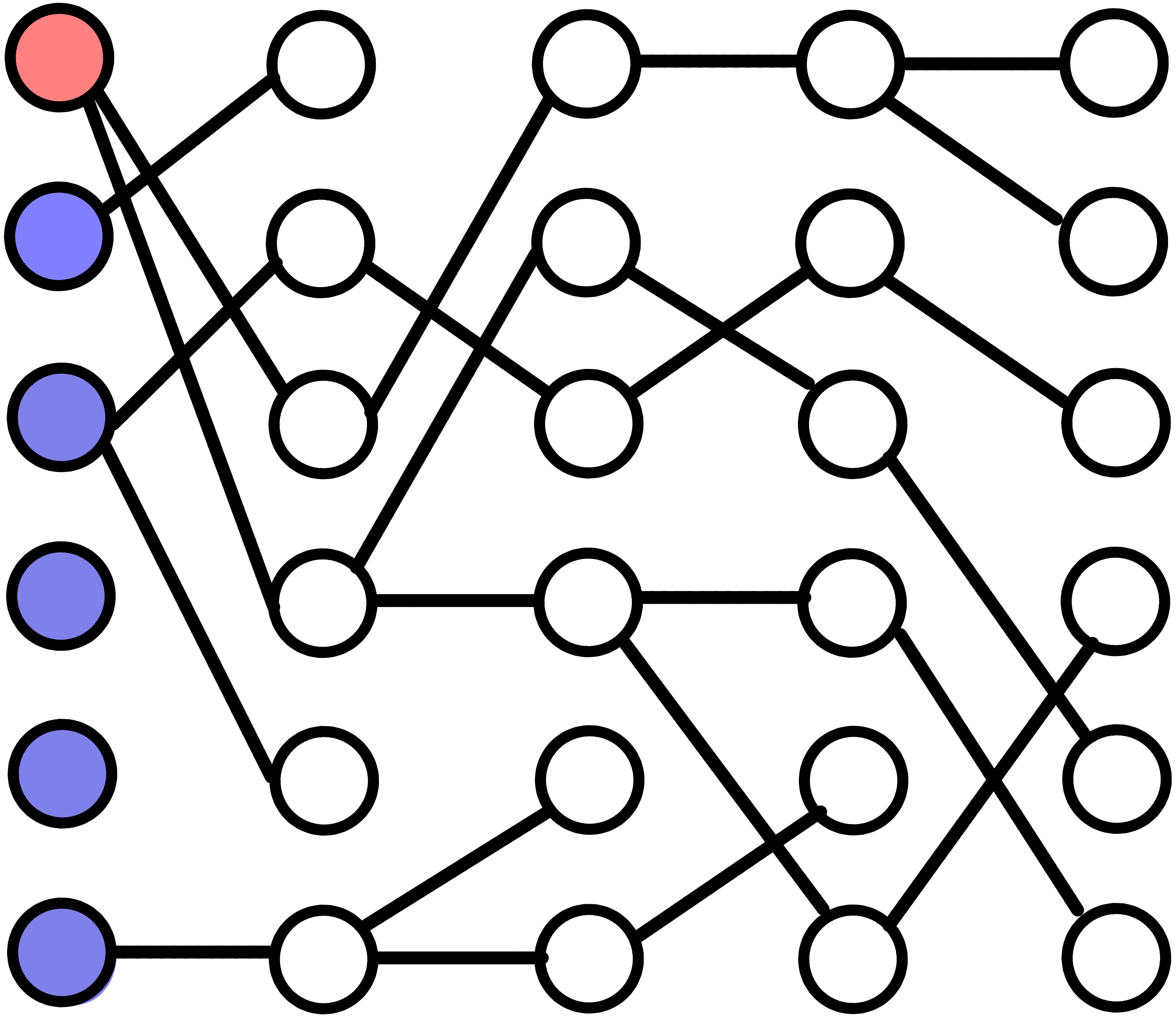
In the Paintbox model

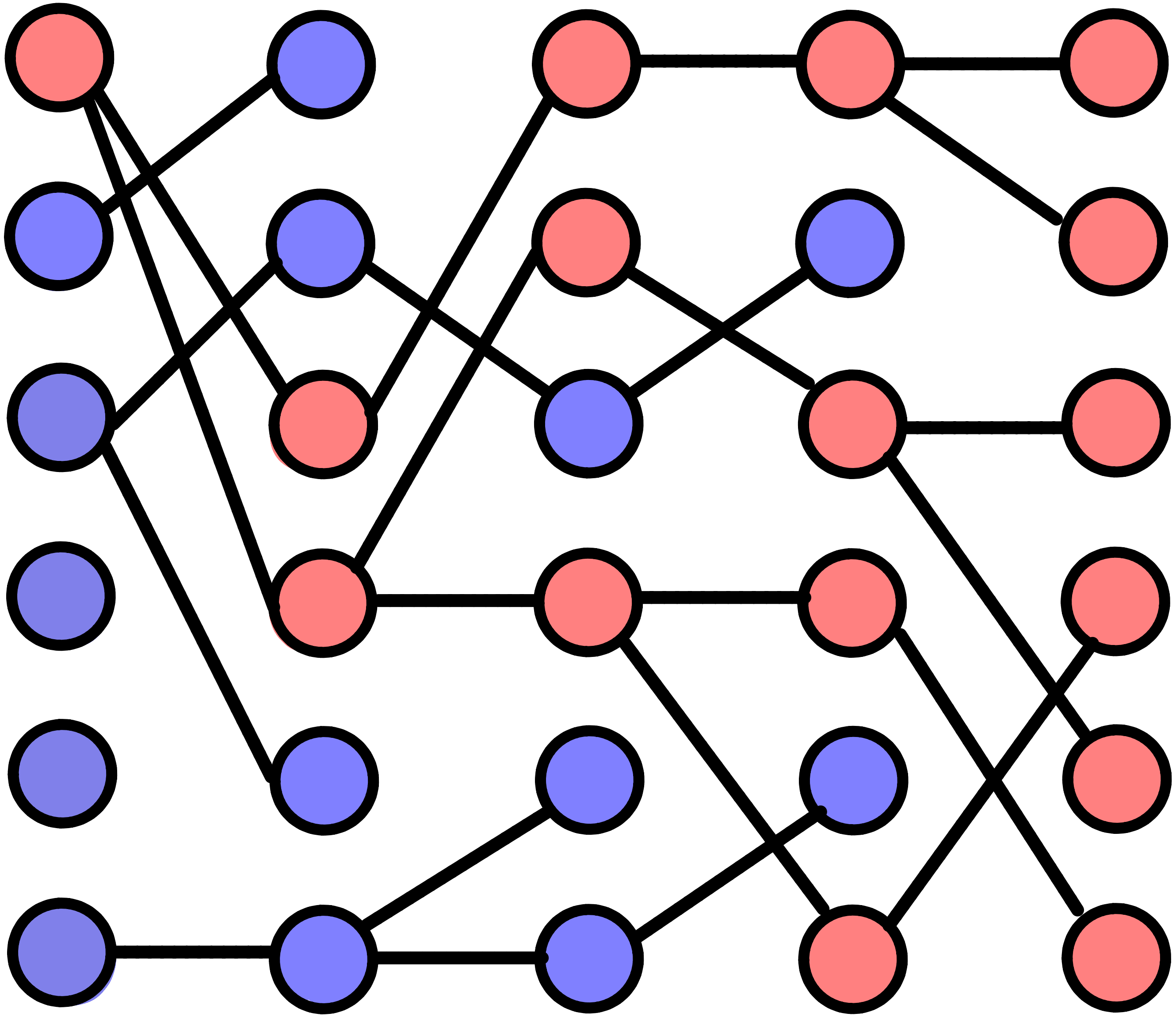


N



$$\mathbb{P}(? = \bullet) = \frac{\sum_{i=\lfloor xN \rfloor}^{N} w_{0i}\,(1-s_N)}{\sum_{i=1}^{N} w_{0i} - s_N \sum_{i=\lfloor xN \rfloor}^{N} w_{0i}}$$





$$X_g = \frac{\# \bullet \text{ in generation } g}{N}$$

Now take $X_0 = xN$

for some $x \in [0,1]$

and $s \in [0,1)$

- If $s_N = s$ $\quad (X_g^N) \Rightarrow (x_g)$ where

$$x_g = \frac{(1-s)x_g}{(1-s)x_g + 1 - x_g}$$

Now take $X_0 = xN$
for some $x \in [0,1]$
and $s \in [0,1)$

If $s_N = \dfrac{s}{N^\beta}$, $\beta \in (0,1)$

$$\left(X^N_{\lfloor N^\beta t \rfloor}\right) \Rightarrow \left(X(t)\right)$$

$$\frac{dX(t)}{dt} = -s\,X(t)\left(1 - X(t)\right)$$

Now take $X_0 = xN$

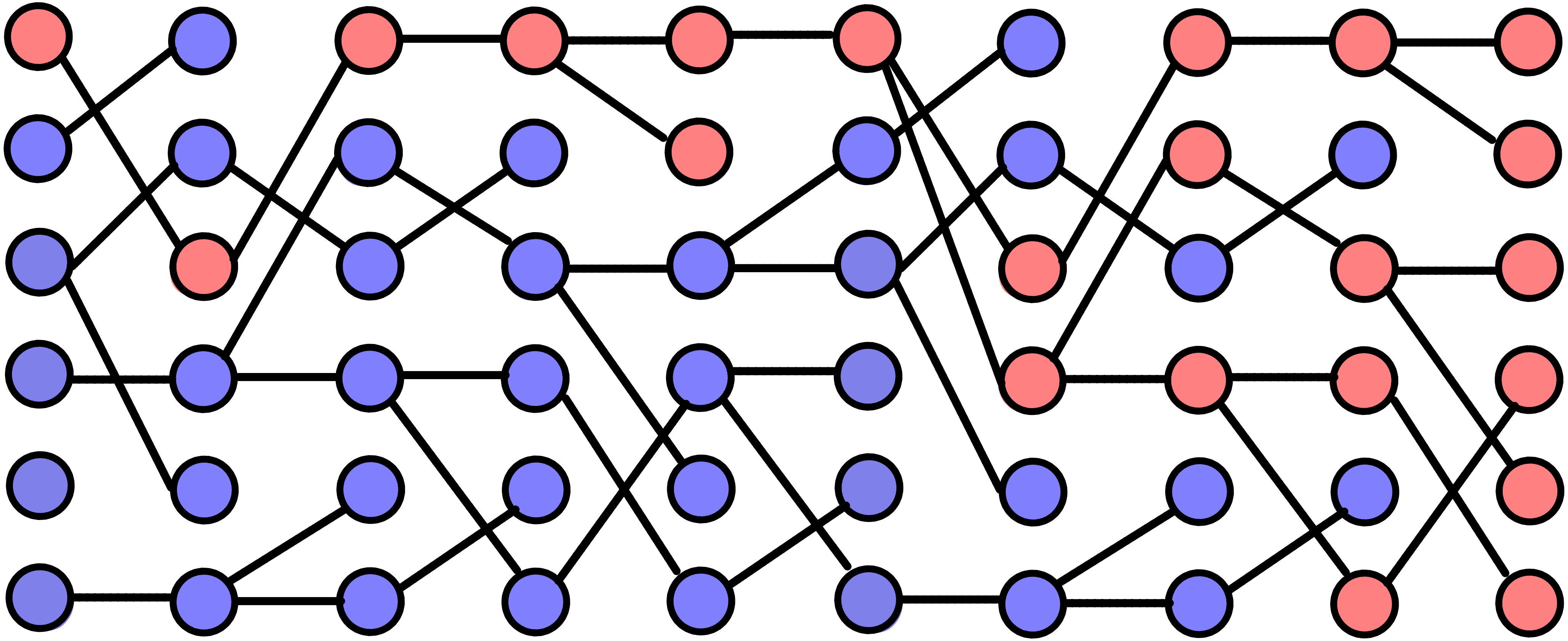
for some $x \in [0,1]$

and $s \in [0,1)$

... If $s_N = \frac{s}{N}$, $(X^N_{\lfloor tN \rfloor}) \Rightarrow (X_t)$

$$dX_t = -sX_t(1-X_t)dt + \sqrt{X_t(1-X_t)}\,B_t$$

Fixation $= \{\exists\, g : X_g = 1\}$



fix $s>0$

Haldane ~ 27

$s_N = s$

$$\mathbb{P}^N_{1/N}(\text{fix.}) = 2s + o(N^{-1})$$

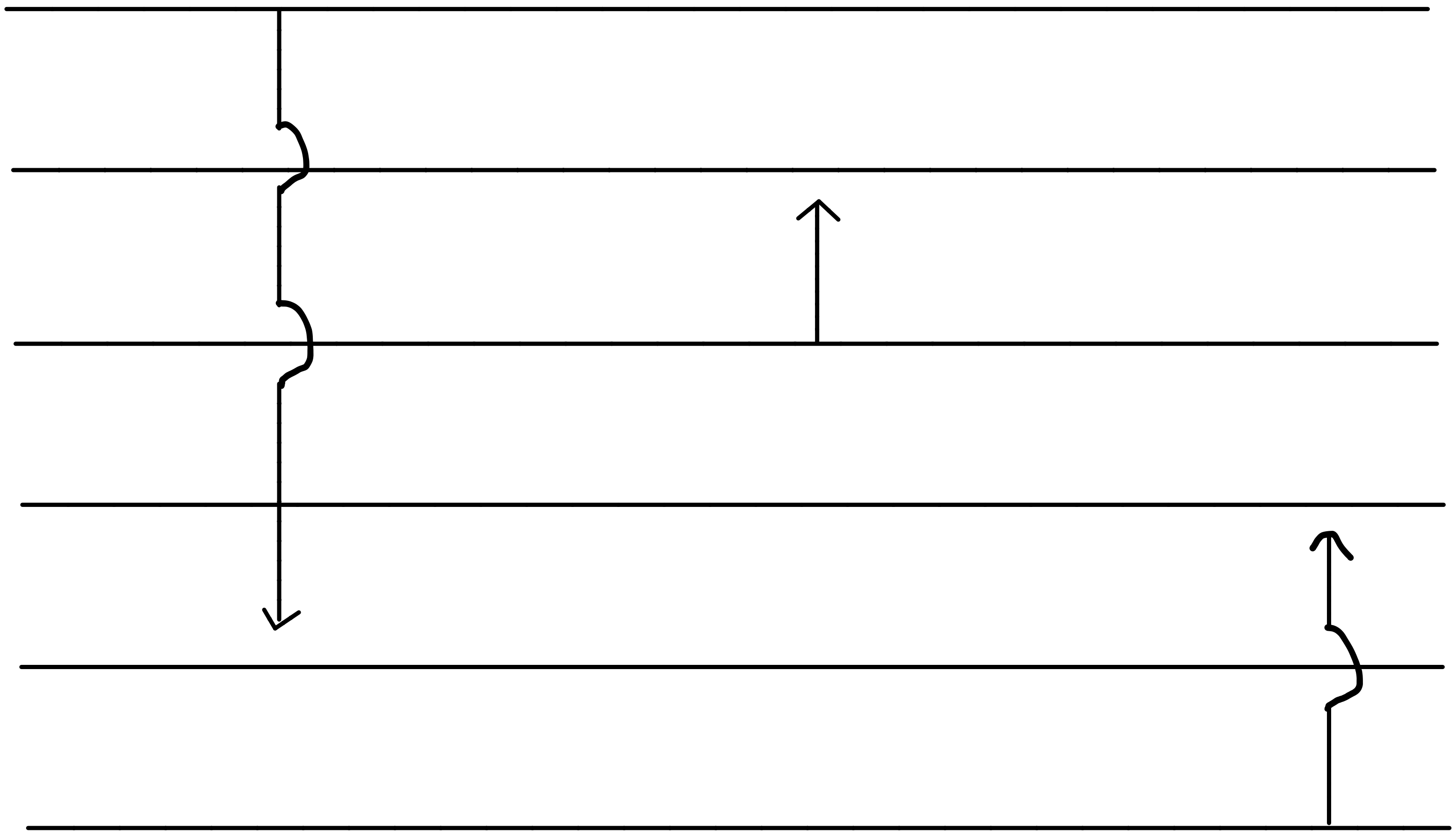
Kimura ~ 62

$s_N = s/N$

$$\mathbb{P}^N_{1/N}(\text{fix.}) = \frac{2s}{1-e^{-2s}} + o(1)$$

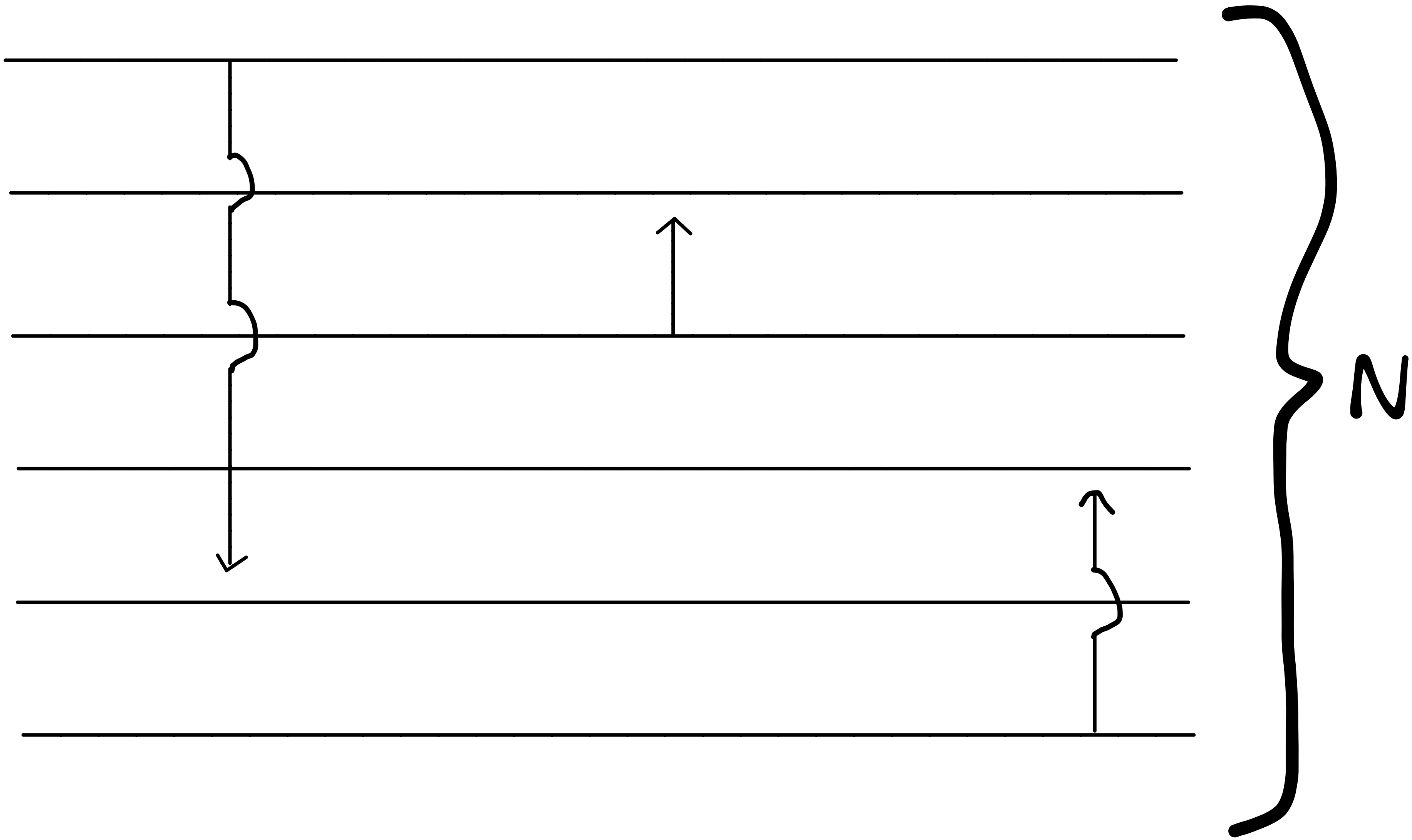
See also Gillespie 74

The Moran model



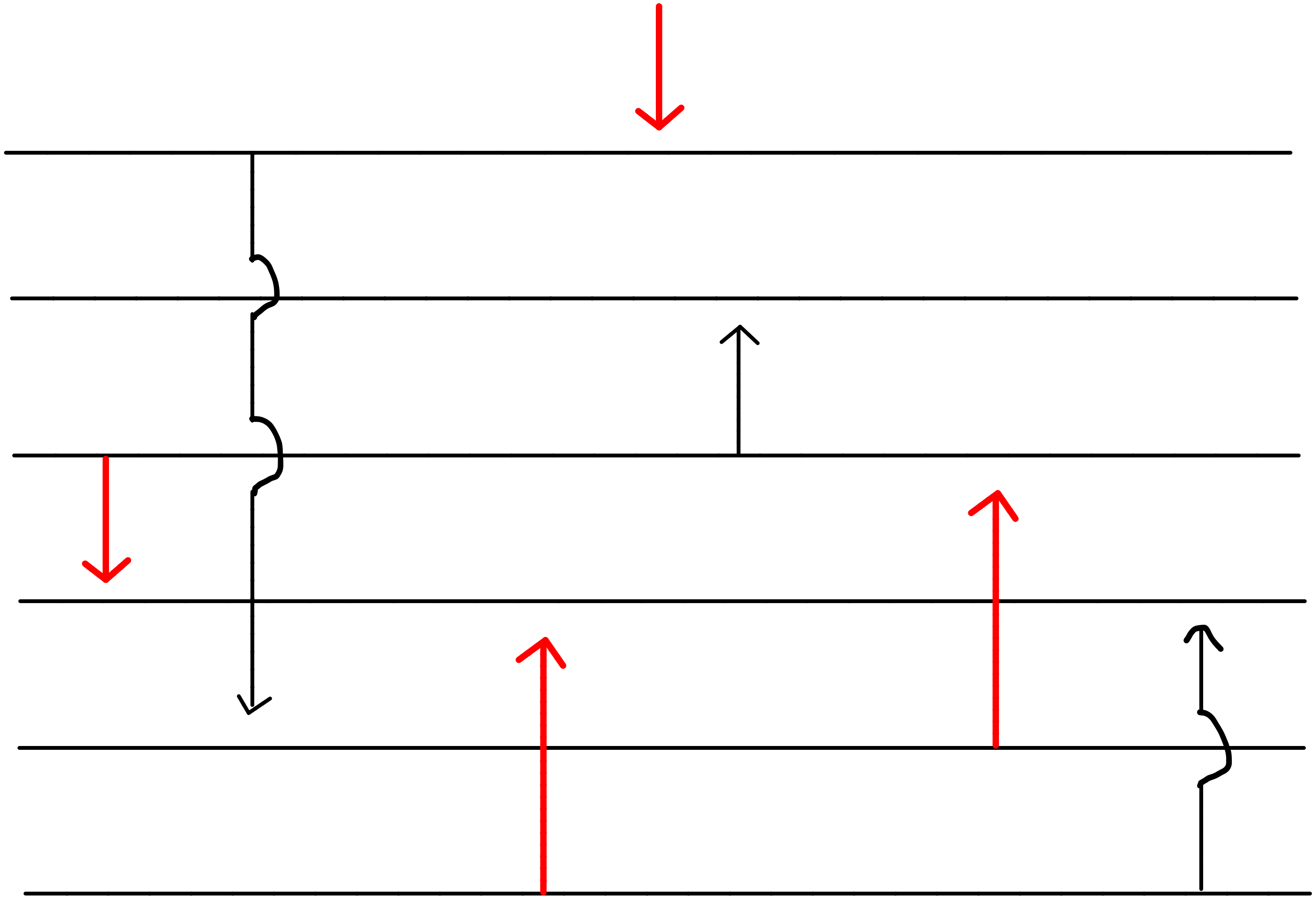
The Moran model

Each individual reproduces $\uparrow$ at Rate $\rho/2 > 0$

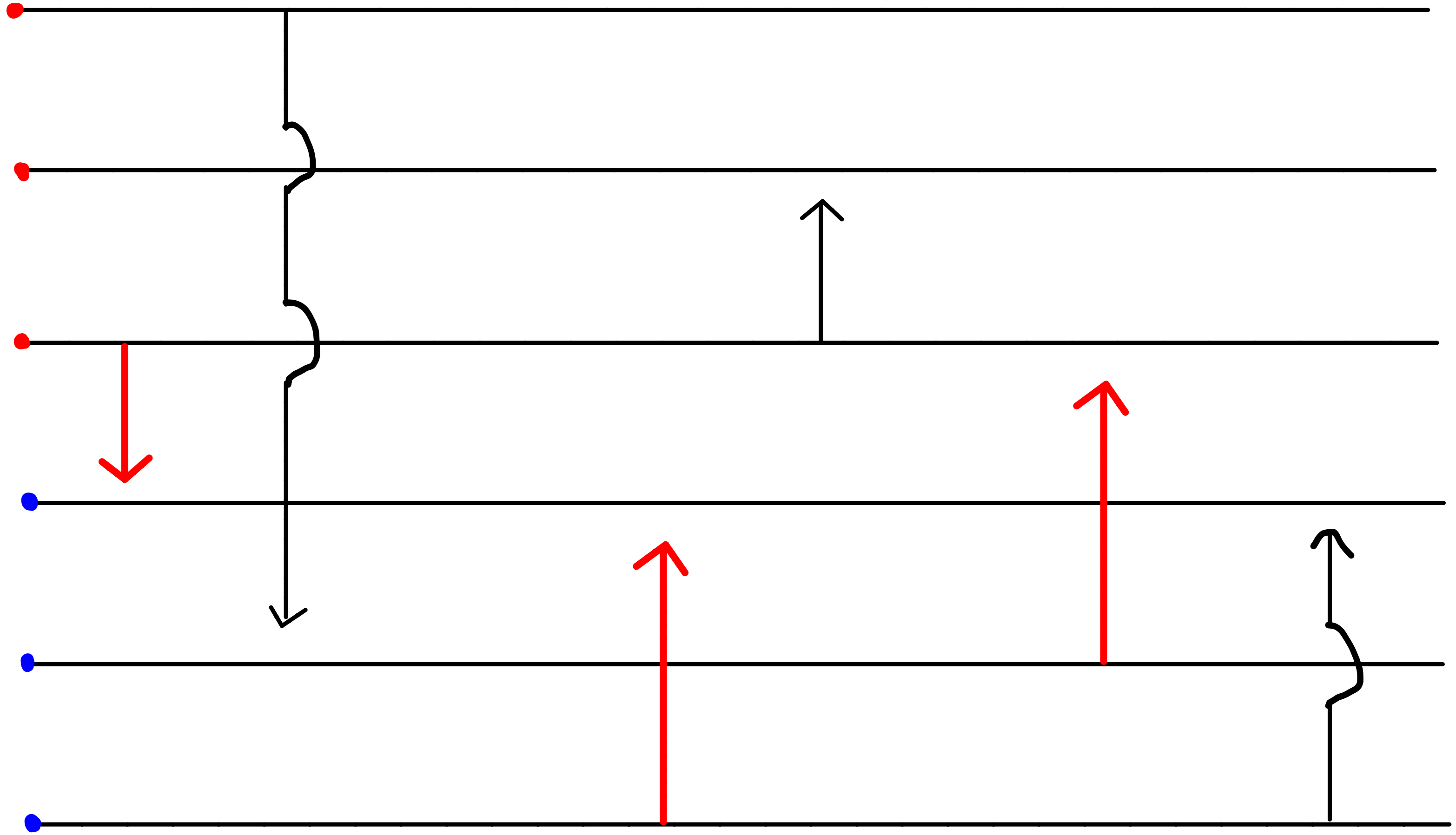


The Moran model

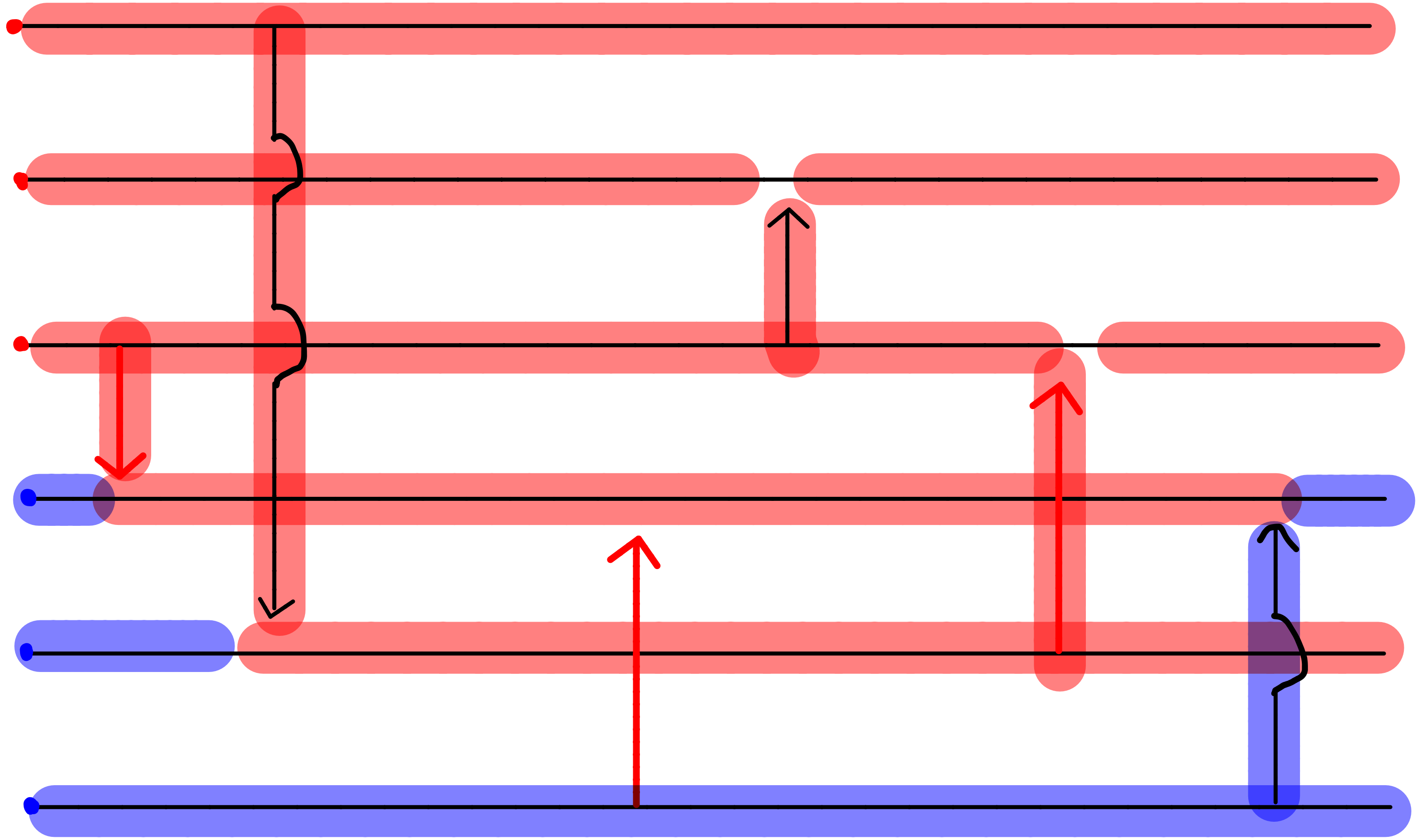
Individuals produce a selective arrow at rate s_N



The Moran model

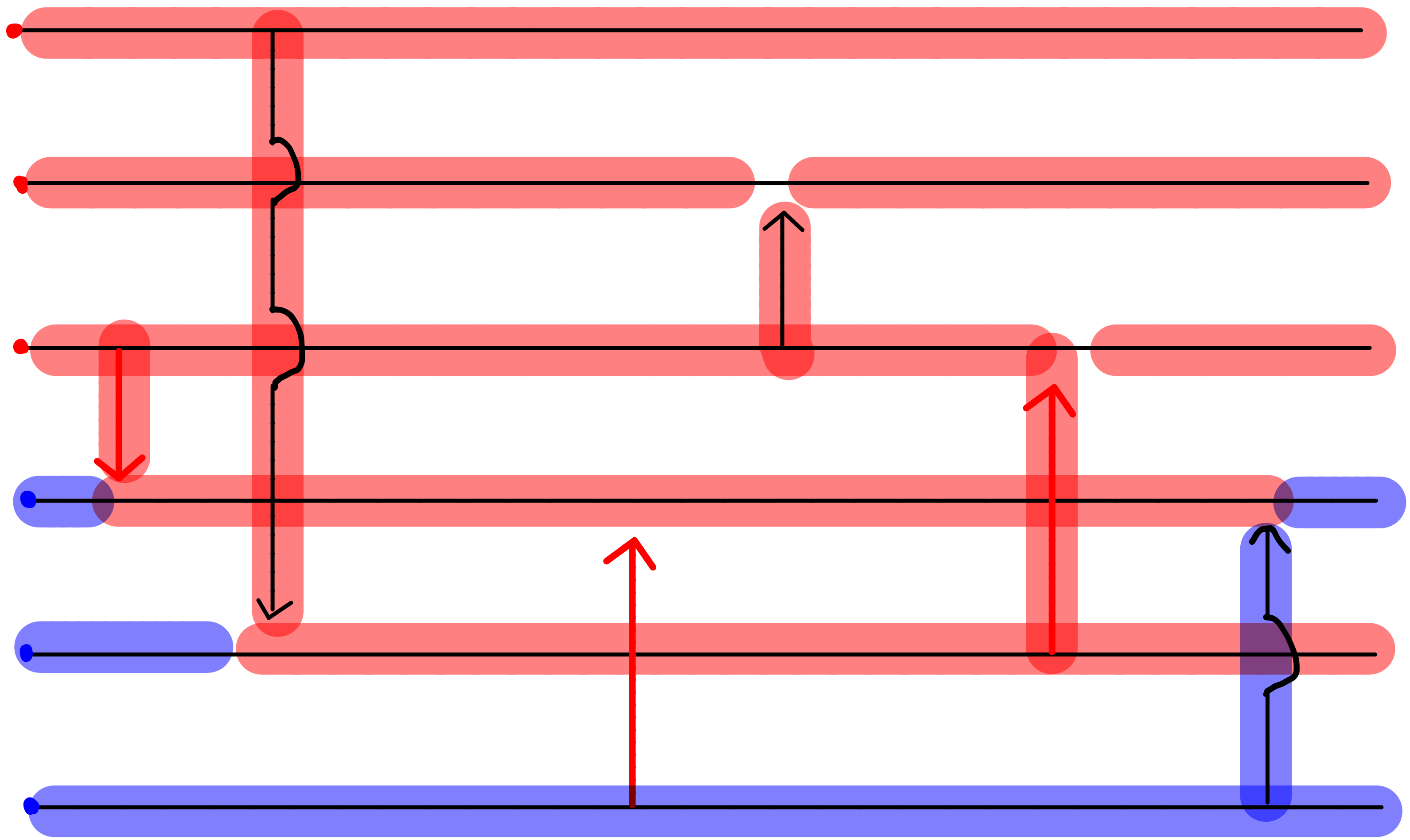


The Moran model

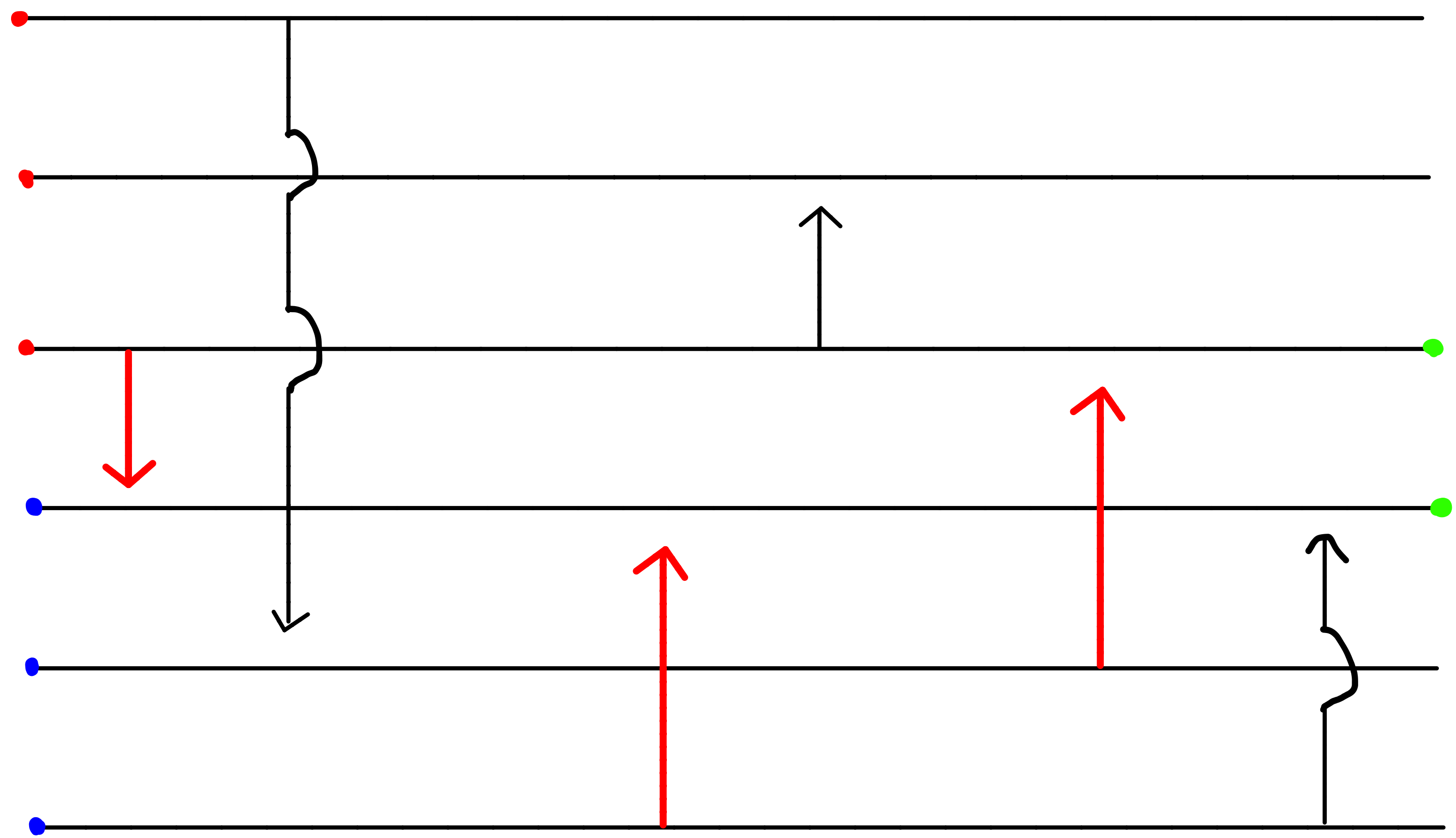


The Moran model

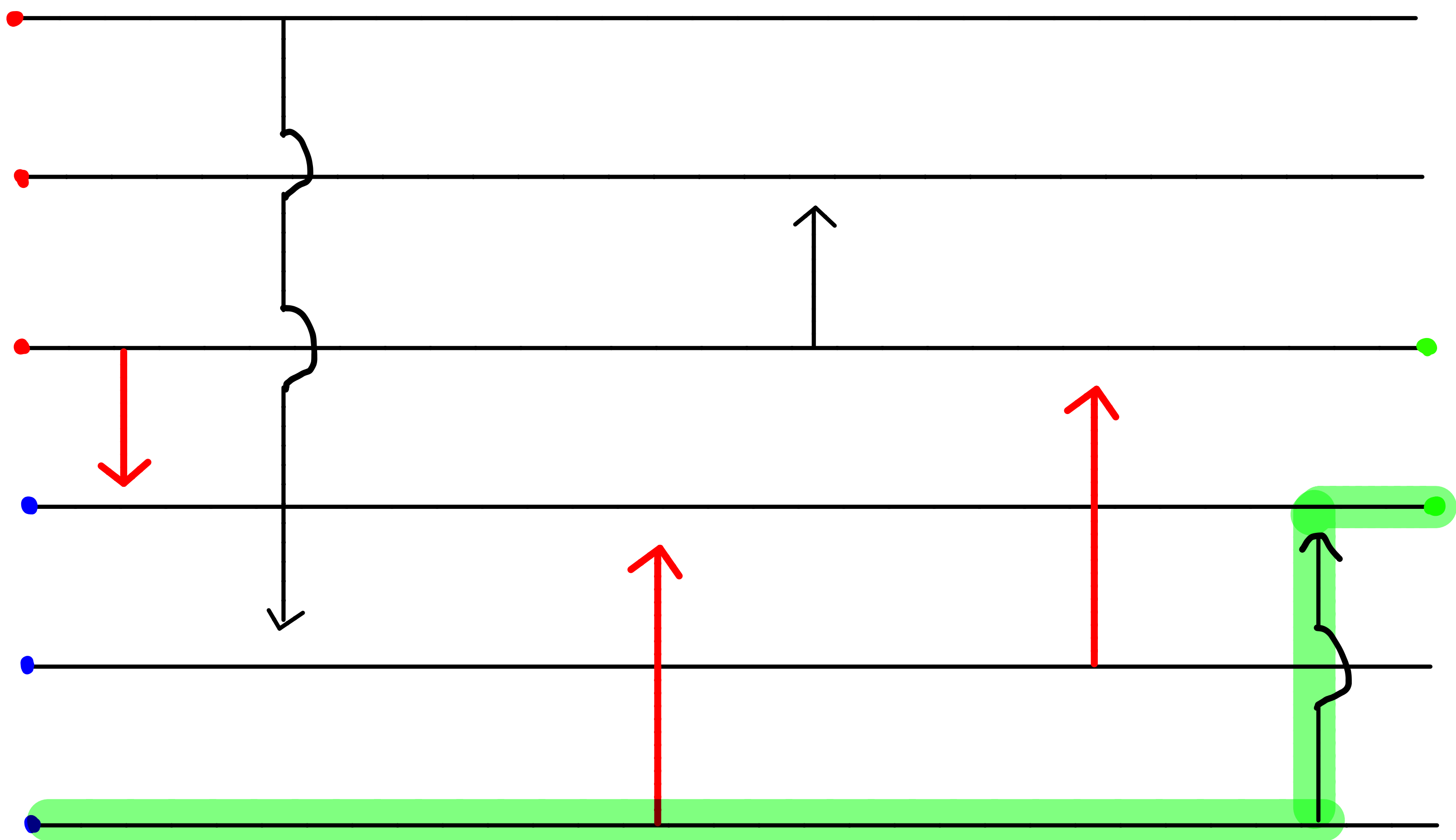
$Y_t = \#$ ● at time t



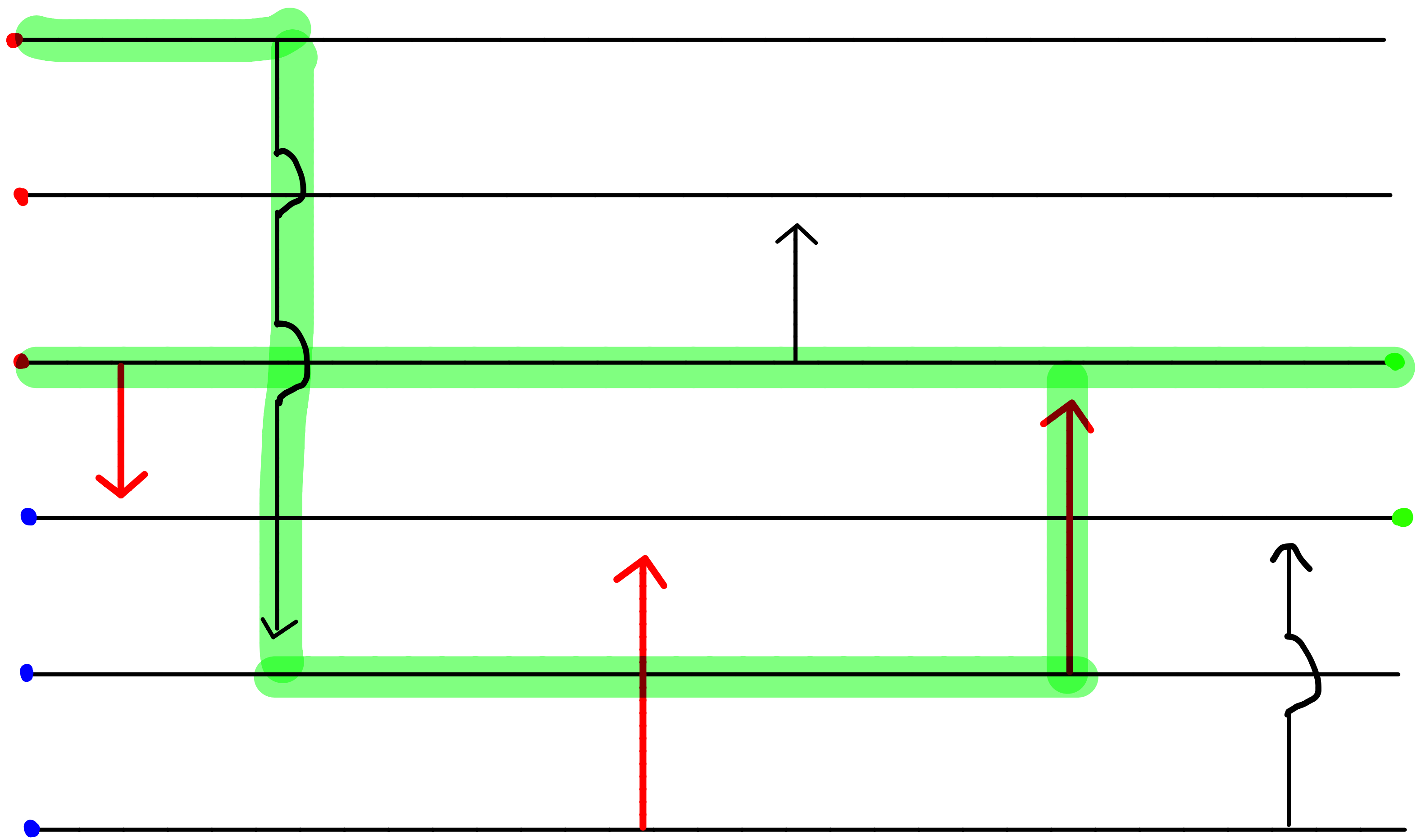
The Ancestral selection graph
(Krone, neuhauser)



The Ancestral selection graph

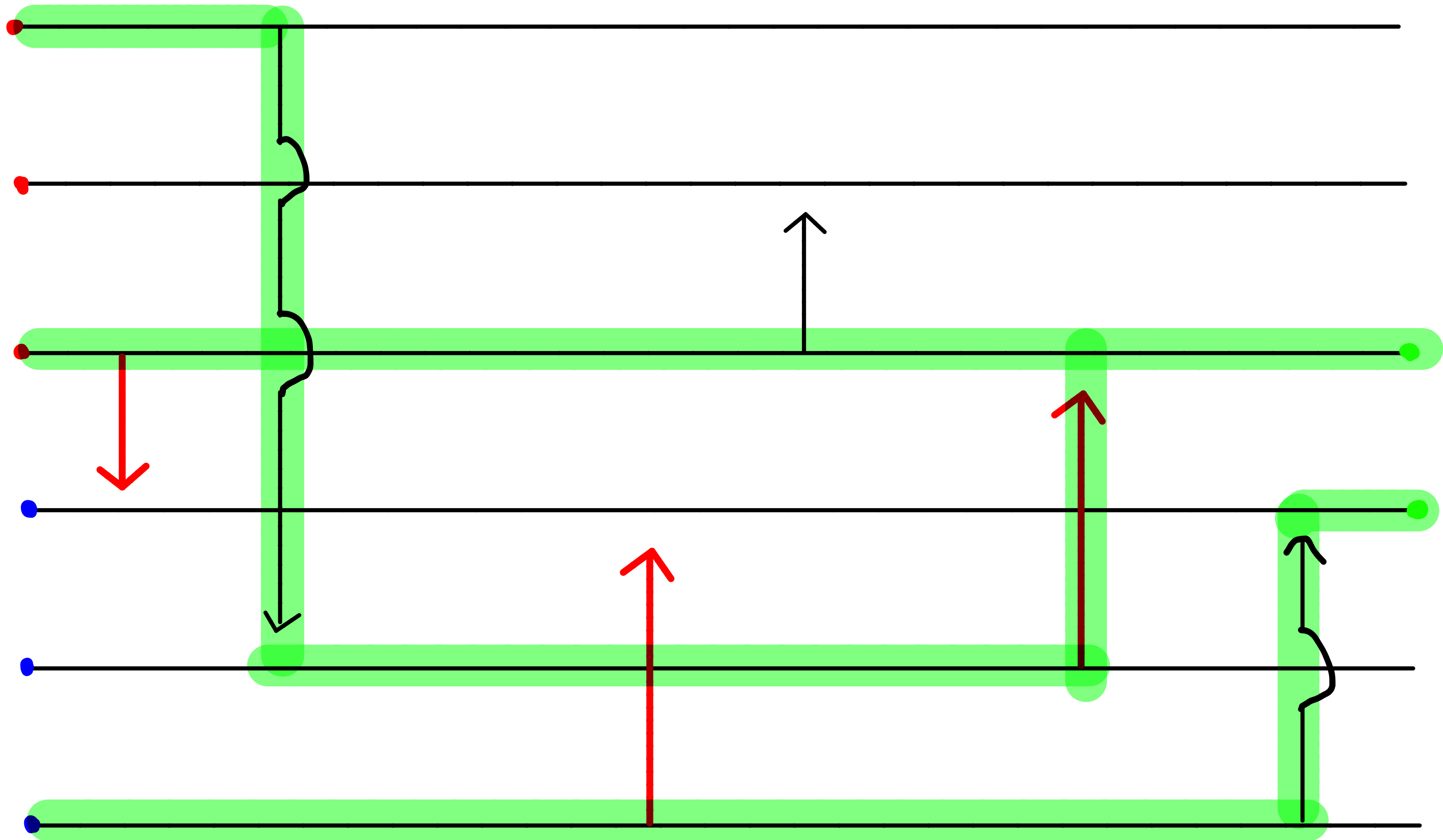


The Ancestral selection graph



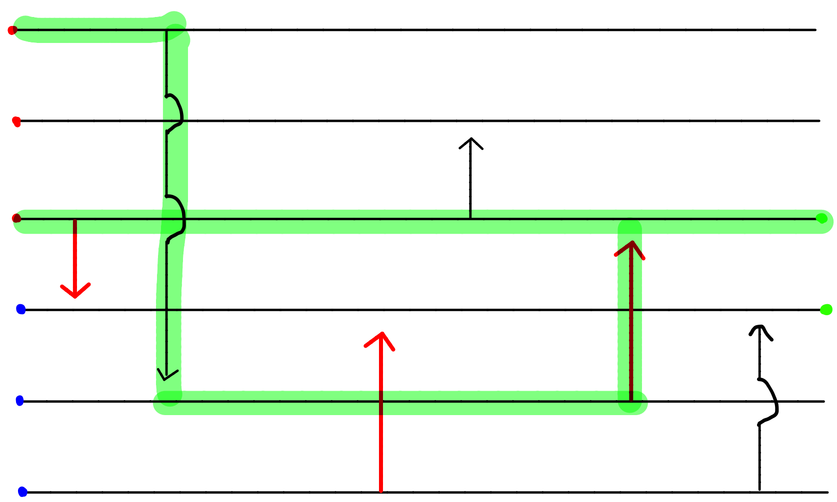
The Ancestral selection graph

$B_t = \#$ ● at time t



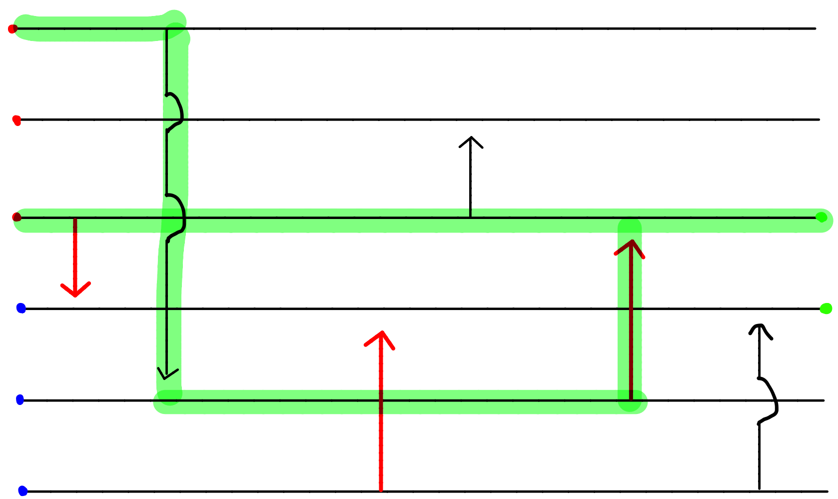
The Ancestral selection graph

$$B_t \text{ goes from } n \text{ to } \begin{cases} n+1 & \text{at Rate} \quad n s_N \left(\frac{N-n}{N}\right) \\ n-1 & \quad \frac{\rho}{N}\binom{n}{2} \end{cases}$$



Stationary distribution
(Möhle, Cordero)

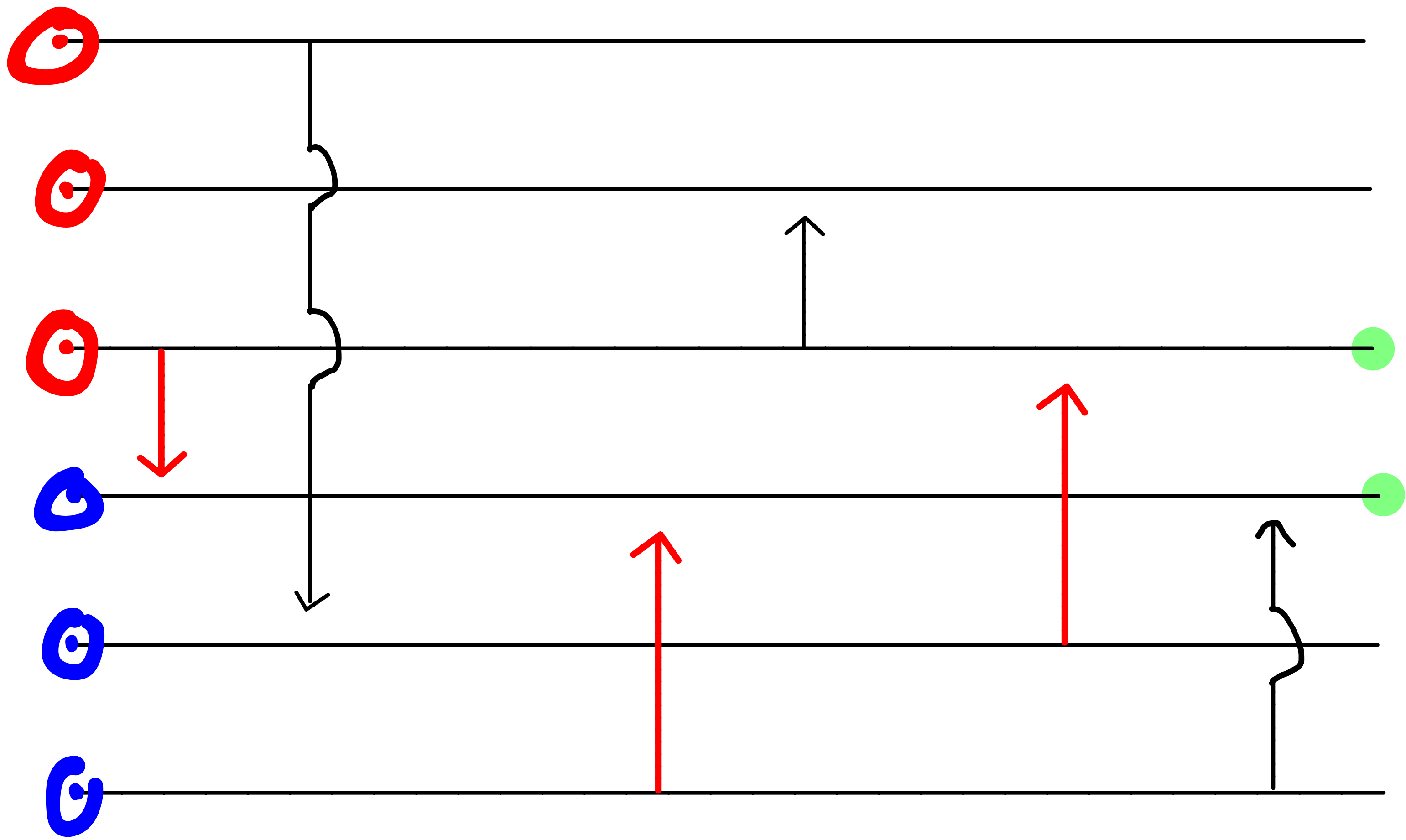
$B_t \xrightarrow{t \to \infty} B$ where B is a Binomial Random variable conditioned to be positive



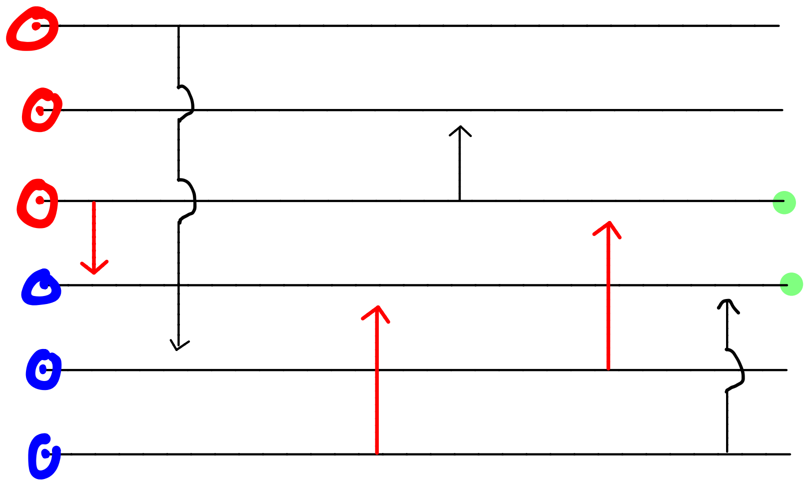
$$B \sim \mathrm{Bin}(N, p_N)\big|_{B>0}$$

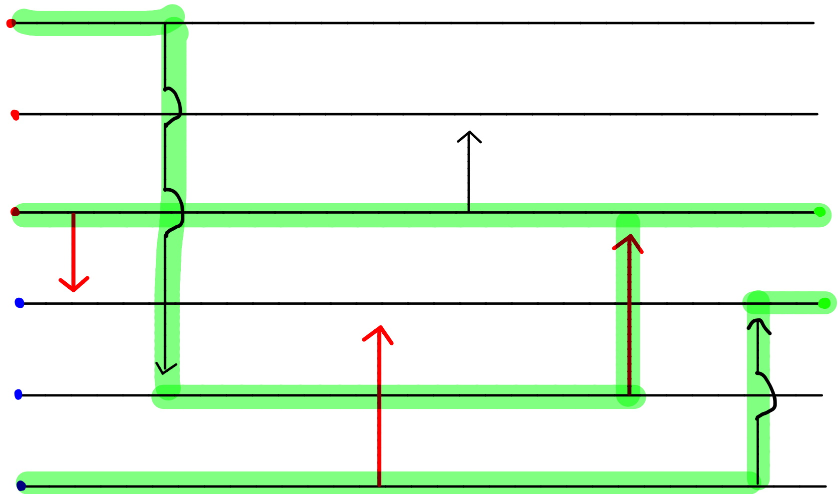
$$p_N = \frac{2s_N}{2s_N + \gamma}$$

Hypergeometric duality
(Hummel)

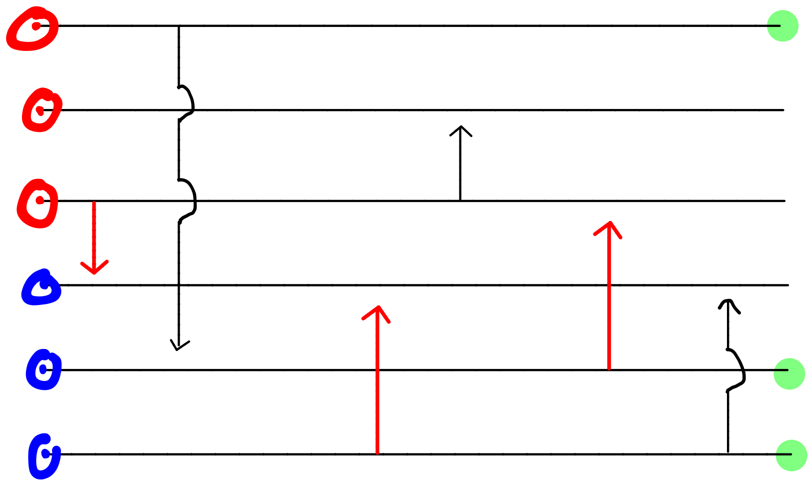


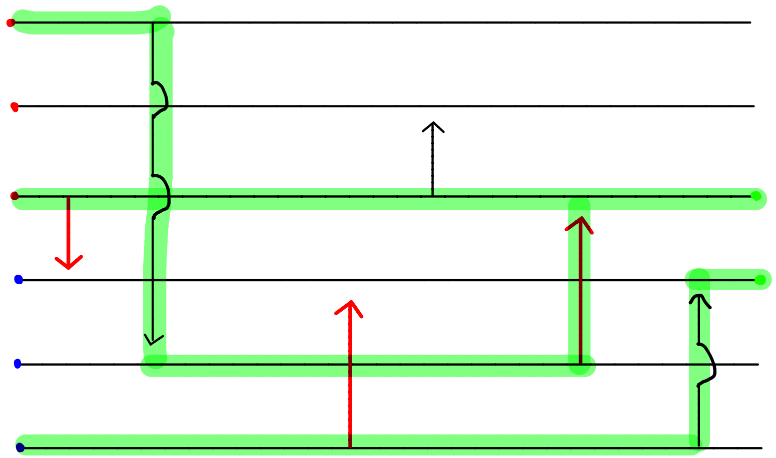
Hypergeometric dualits

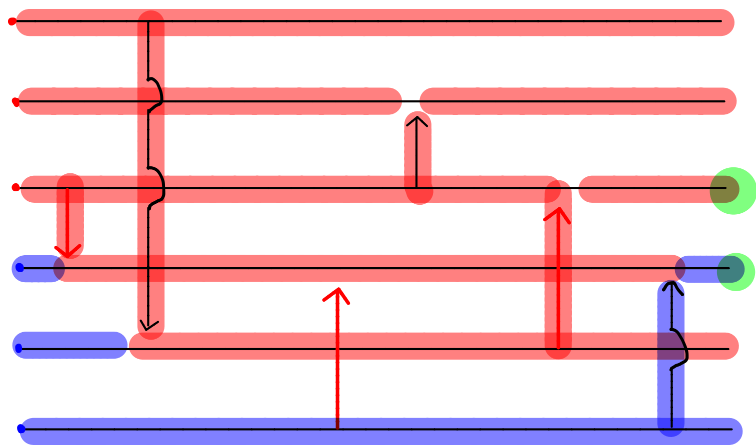




Hypergeometric dualities

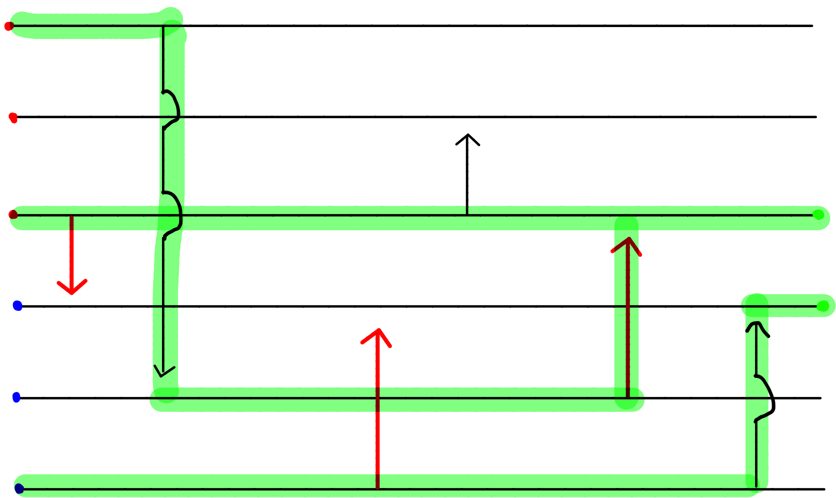


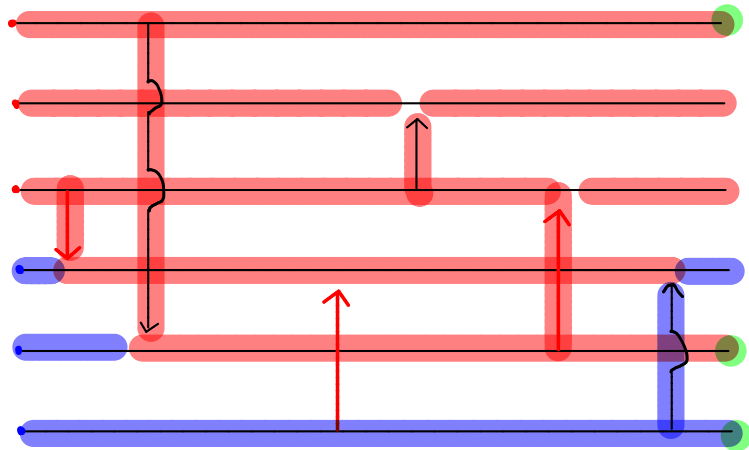




Hypergeometric dualities

$$\mathbb{E}_y\left[\frac{Y_t^N}{N}\cdot\frac{Y_t^N-1}{N}\cdots\frac{Y_t^N-(n-1)}{N}\right]=\mathbb{E}_n\left[\frac{K(K-1)}{N}\cdots\frac{K-(B_t^N-1)}{N-(B_t^N-1)}\right]$$



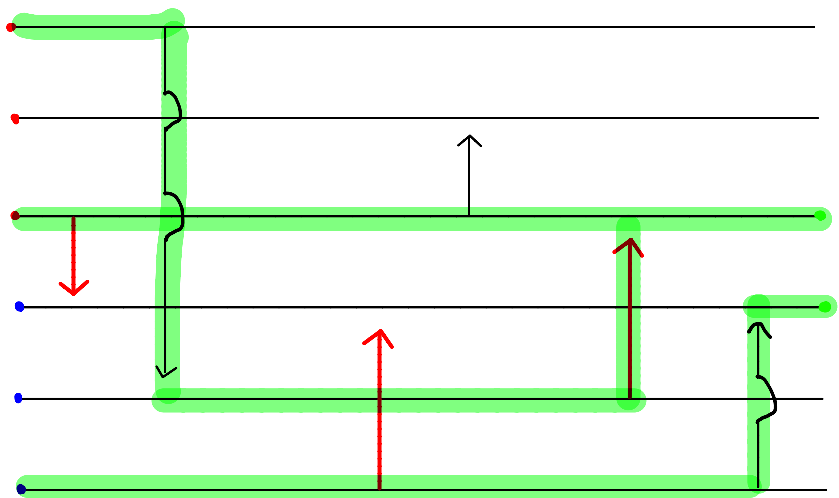


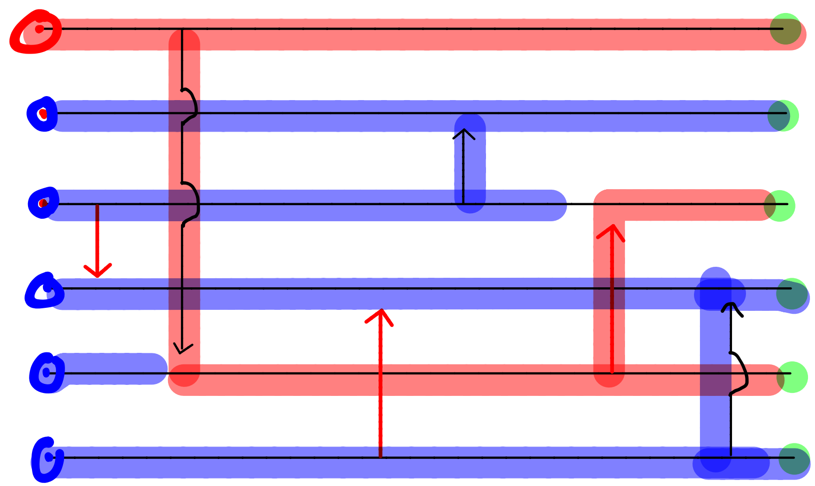
Haldane-Kimura Formula

$$\mathbb{E}_y\left[\frac{y_t^N}{N}\cdot\frac{y_t^N-1}{N}\cdots\frac{y_t^N-(n-1)}{N}\right]=\mathbb{E}_n\left[\frac{K(K-1)}{N}\cdots\frac{K-(B_t^N-1)}{N-(B_t^N-1)}\right]$$

$K=N-1,\ n=N$

$$\mathbb{E}_{N-1}\left[\frac{y_t}{N}\right]=\mathbb{E}_N\left[\frac{N-B_t}{N}\right]$$



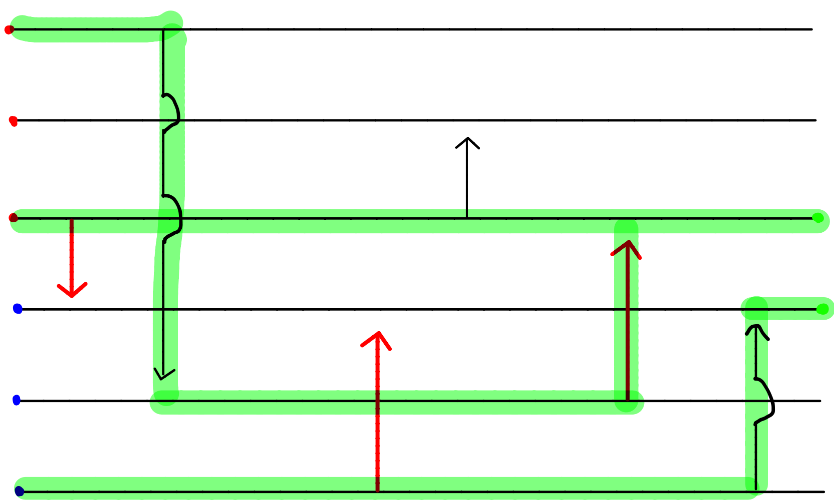


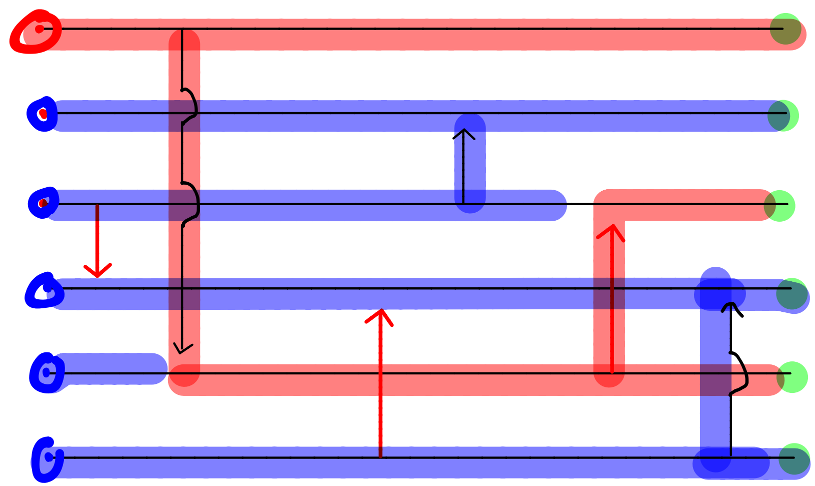
Haldane-Kimura Formula

$$\mathbb{E}_{N-1}\left[\frac{Y_t}{N}\right] = \mathbb{E}_{N}\left[\frac{N-B_t}{N}\right]$$

$K = N-1$, $n = N$ and $t \to \infty$

$$\mathbb{P}_{1/N}(\mathrm{Fix}) = \frac{\mathbb{E}[B]}{N}$$



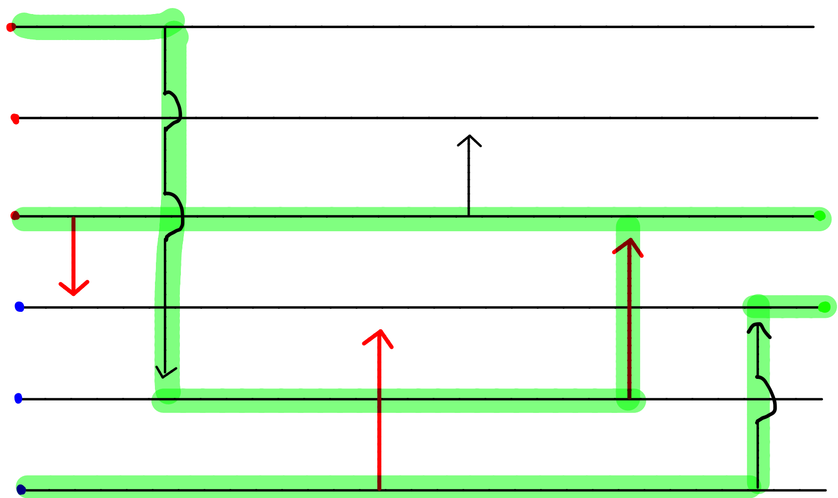


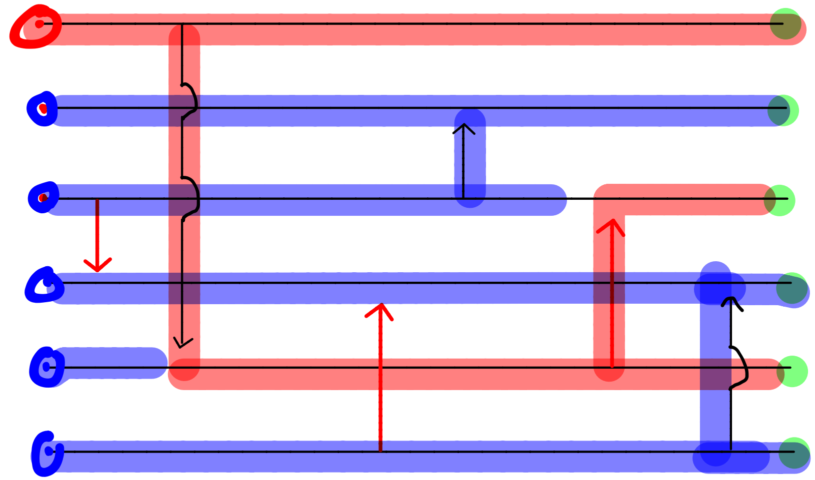
Haldane-Kimura Formula

Boenkost, GC, Pokalyuk, Wakolbinger

$$\mathbb{P}_{1/N}(\text{Fix}) = \frac{\mathbb{E}[B]}{N} = \frac{p_N}{1-(1-p_N)^N}$$

$$p_N = \frac{2s_N}{2s_N + \rho^2}$$





$$\mathbb{P}_{1/N}(\text{Fix}) = \frac{\mathbb{E}[B]}{N} = \frac{p_N}{1-(1-p_N)^N}$$

$$p_N = \frac{2s_N}{2s_N + \rho^2}$$

Haldane ~ 27

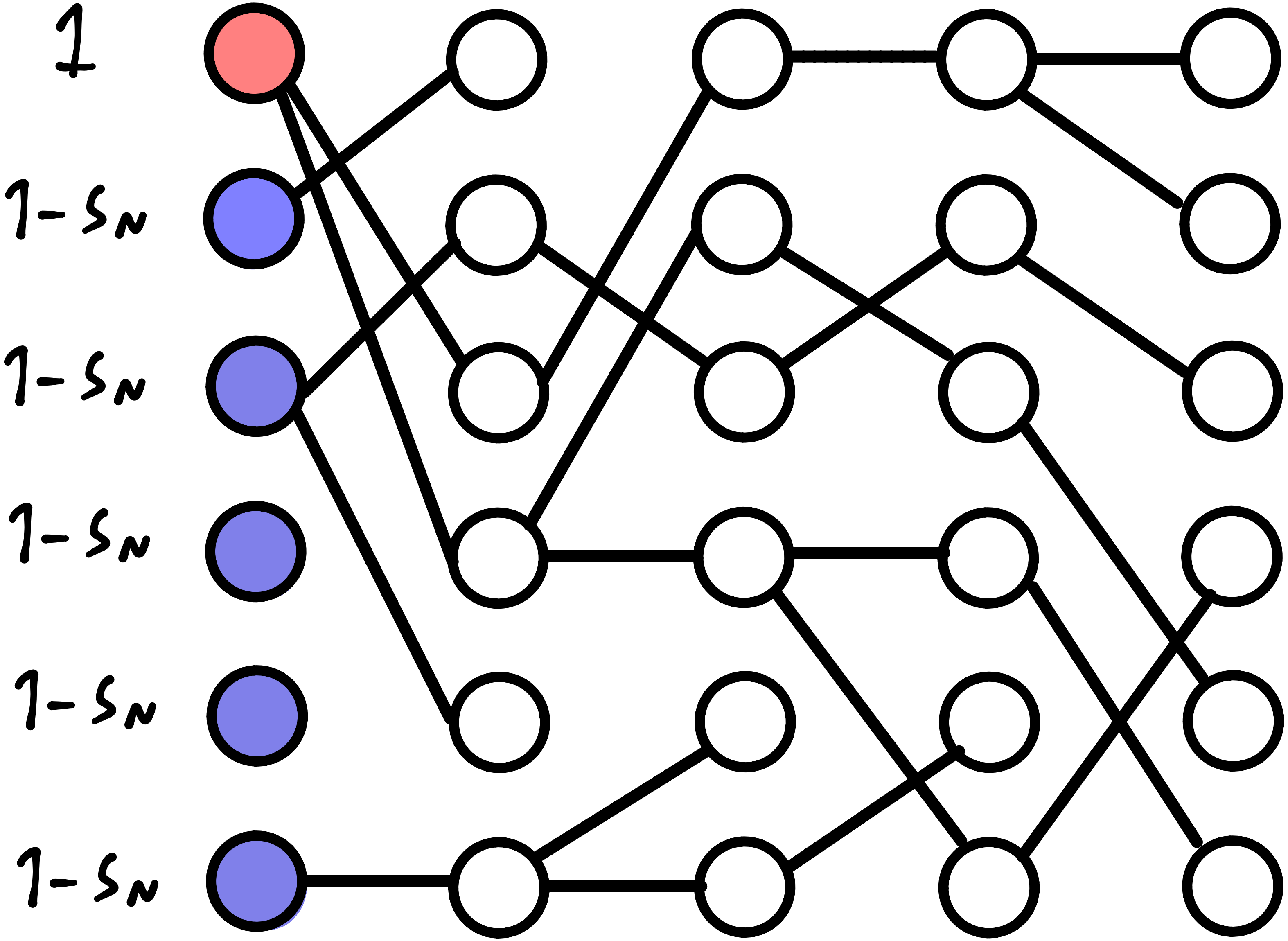
$s_N = s$

$$\mathbb{P}^N_{1/N}(\text{Fix.}) = 2s + o(N^{-1})$$

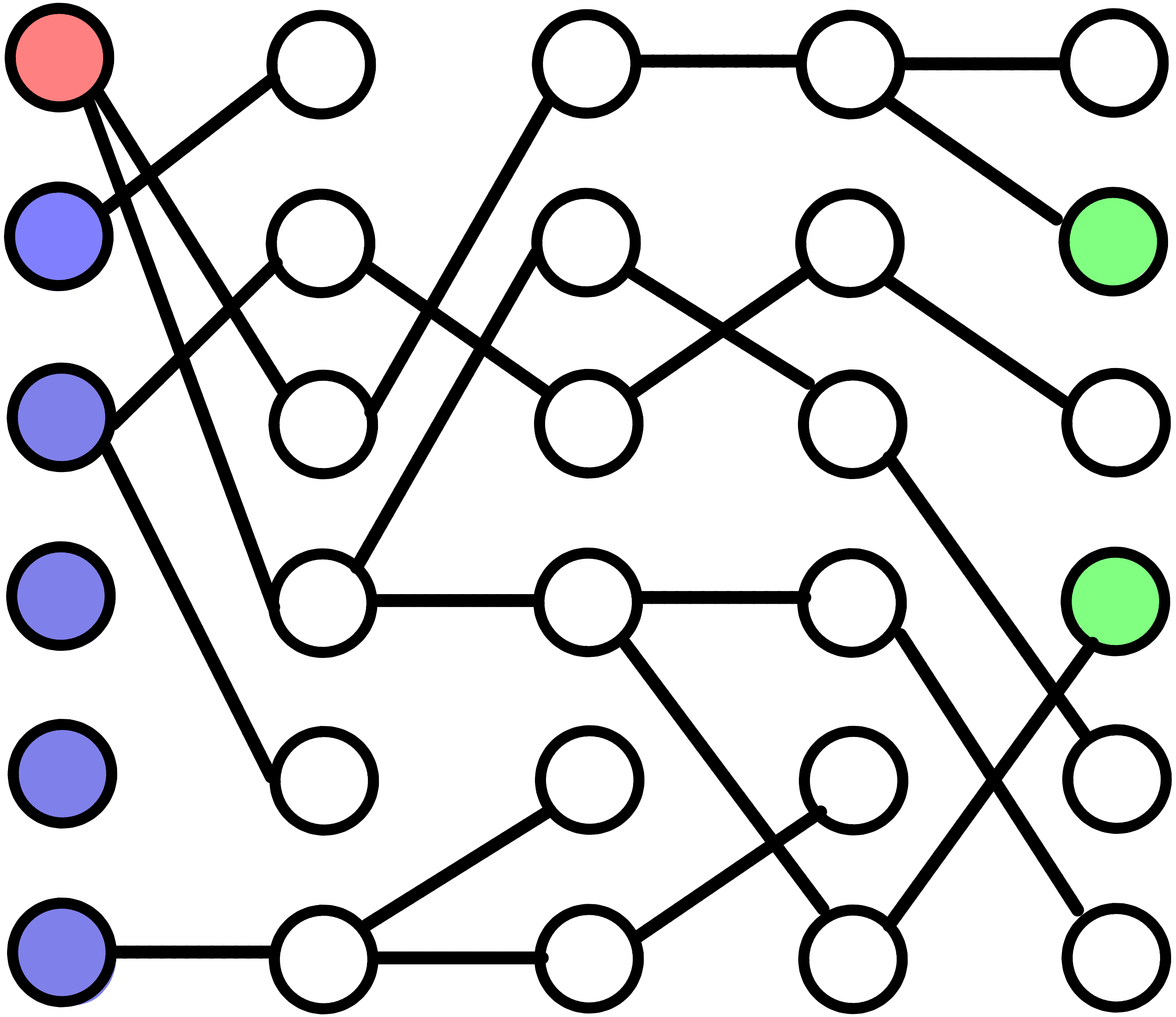
Kimura ~ 62

$s_N \simeq s/N$

$$\mathbb{P}^N_{1/N}(\text{Fix.}) = \frac{2s}{1-e^{-2s}} + o(1)$$

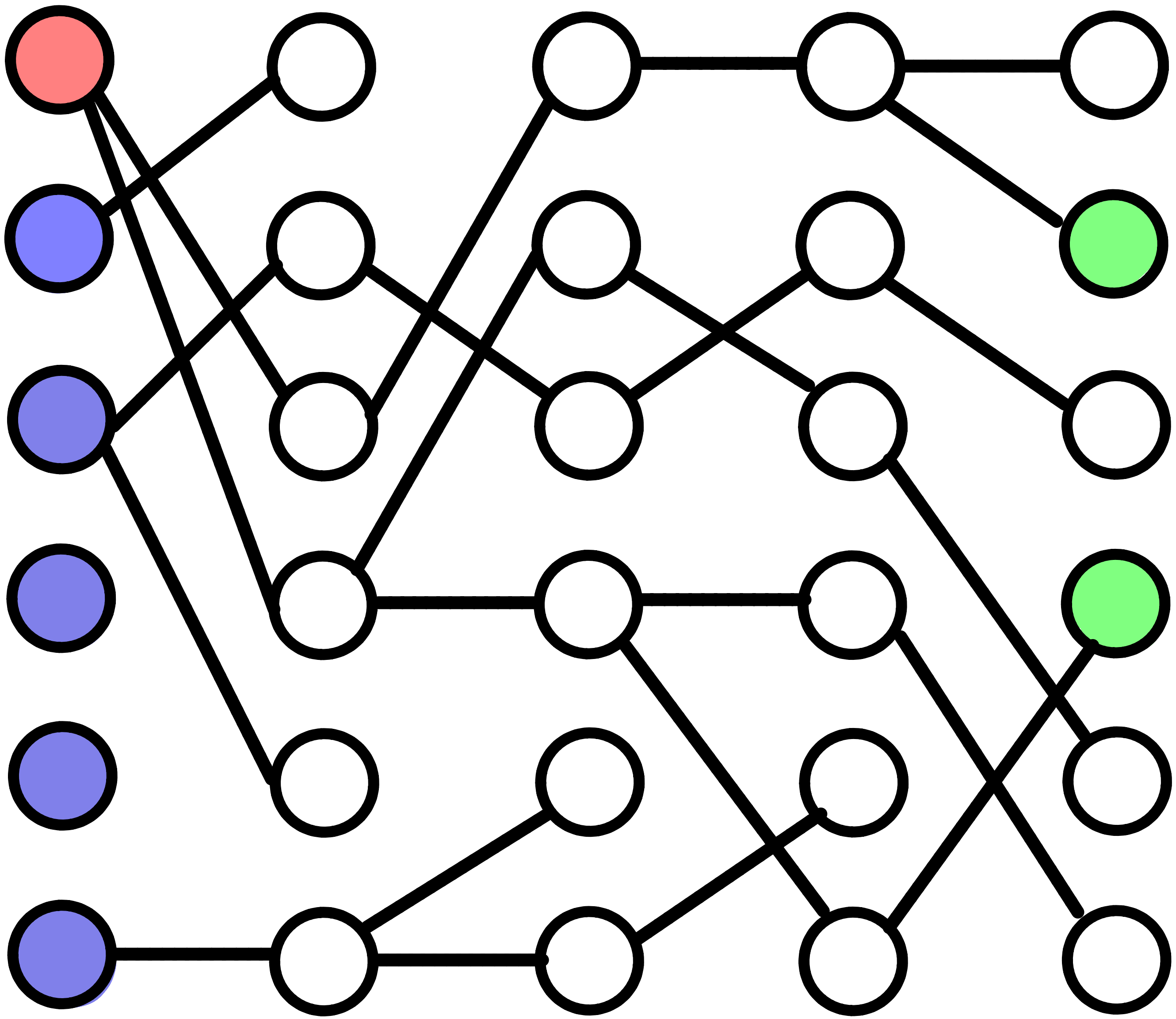


$$X_g = \frac{\#\ \text{in generation } g}{N}$$



$A_g = ?$

The Discrete ancestral selection graph
(G.C. Spanò)

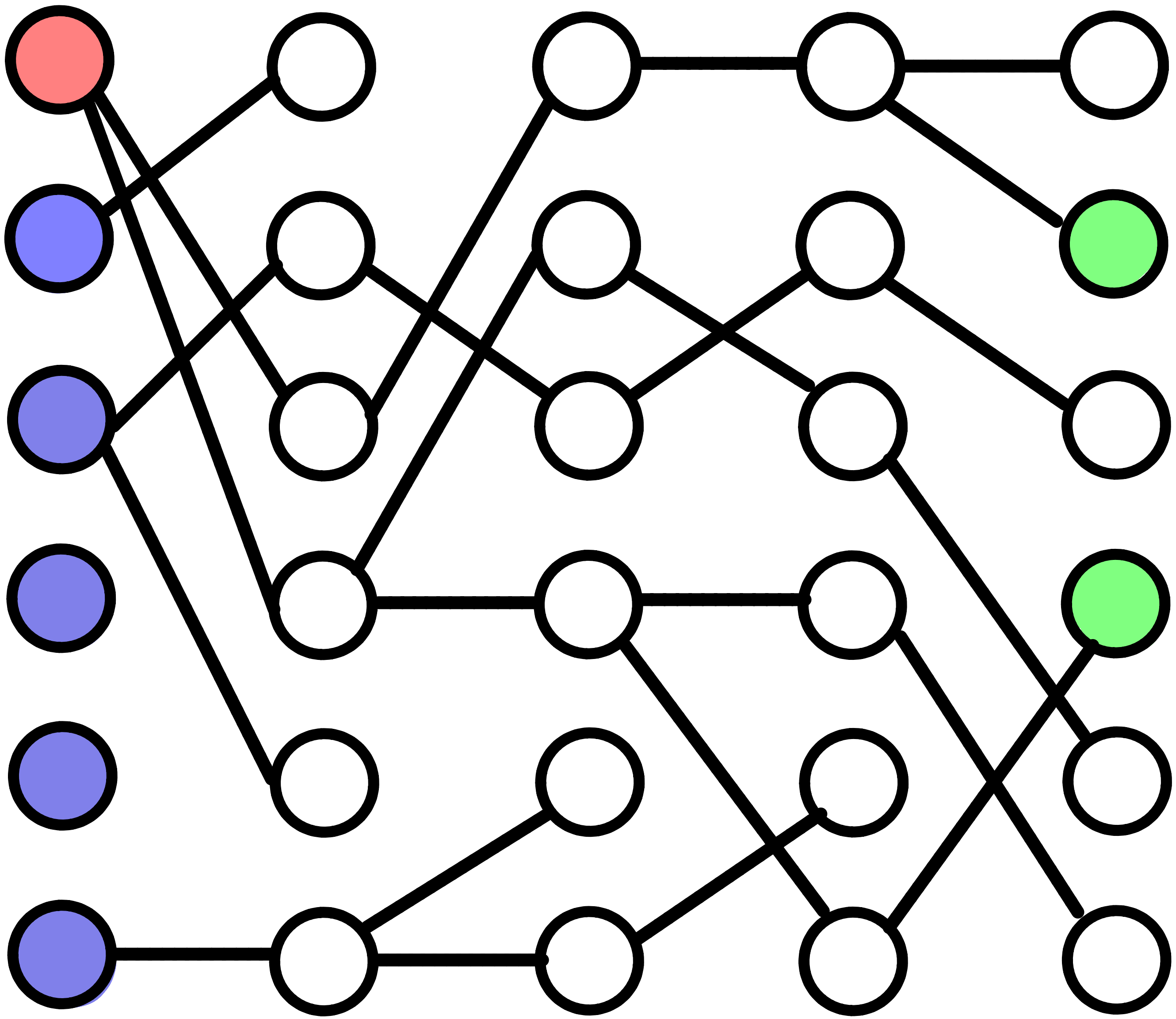


1) Vertex v has

$K_v \geq 1$

Directed edges
(each of them as
in the WF model)

The Discrete ancestral selection graph (G.C. Spanò)



1) Vertex v has

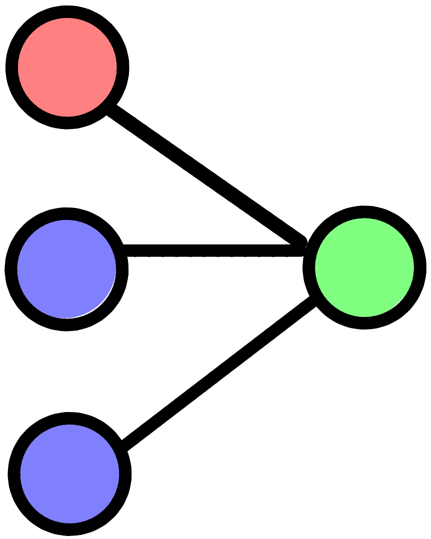
$$K_v \geq 1$$

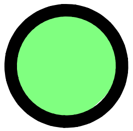
Directed edges
(each of them as
in the WF model)

See also G.C. Spanò and Wilke Berenguer

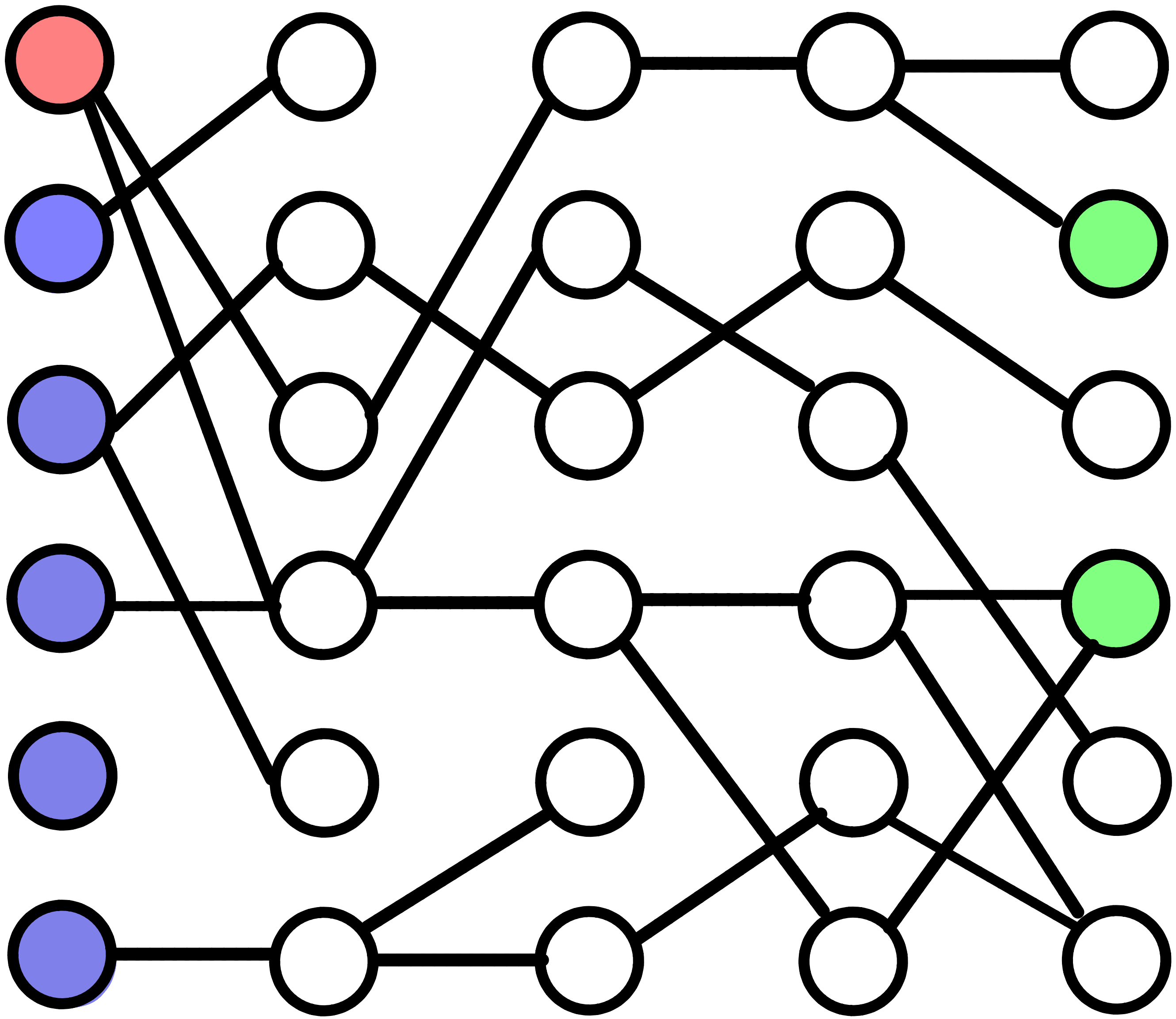
The Discrete ancestral selection graph

Coloring Rule:

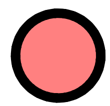


 is Red if and only if it has a Red Parent

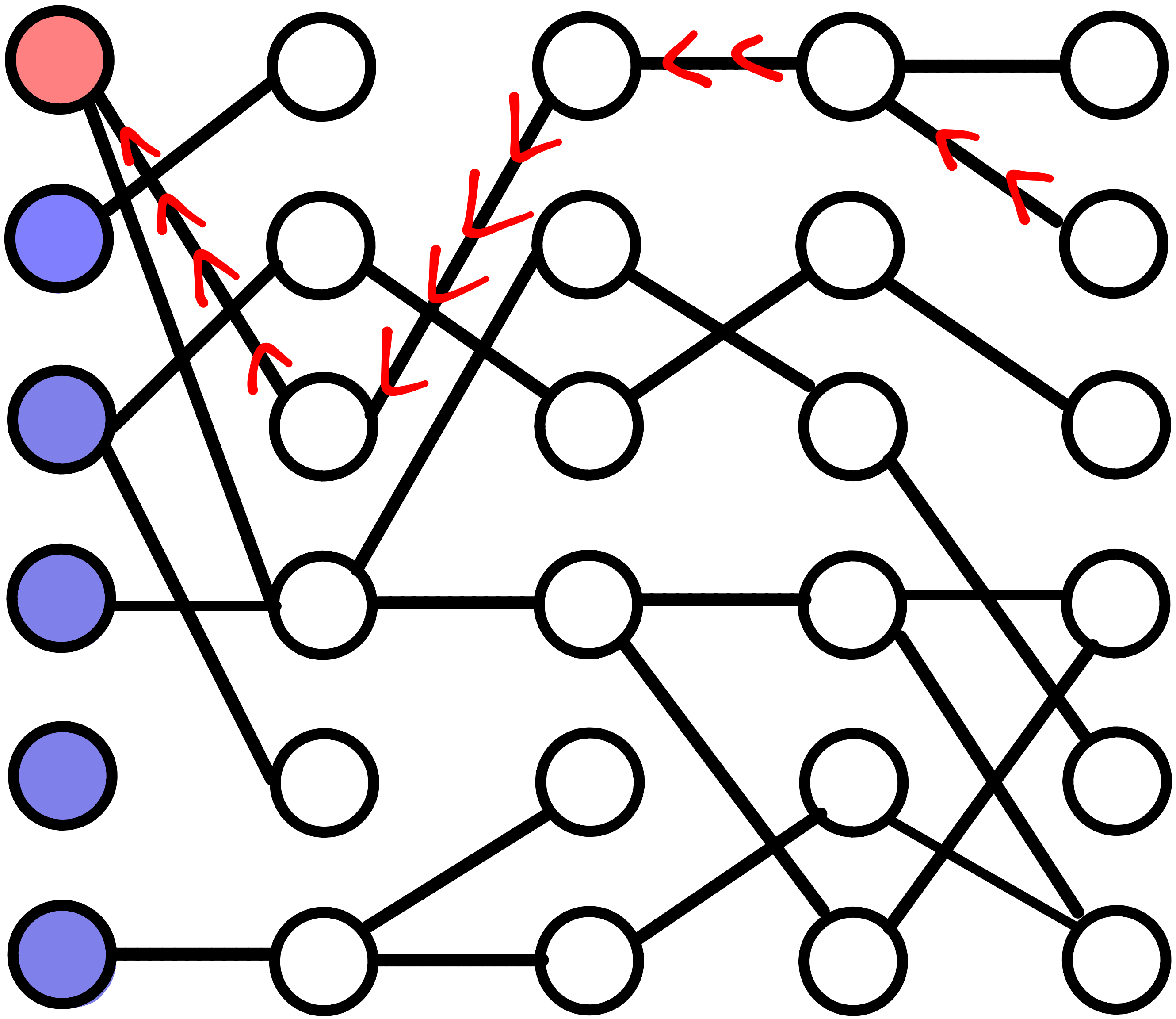
The Discrete ancestral selection graph



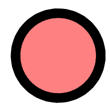
1) Vertex v has K_v Directed edges (each of them as in the WF model)

2) $\tau(v) =$  if and only if v is connected to a red vertex

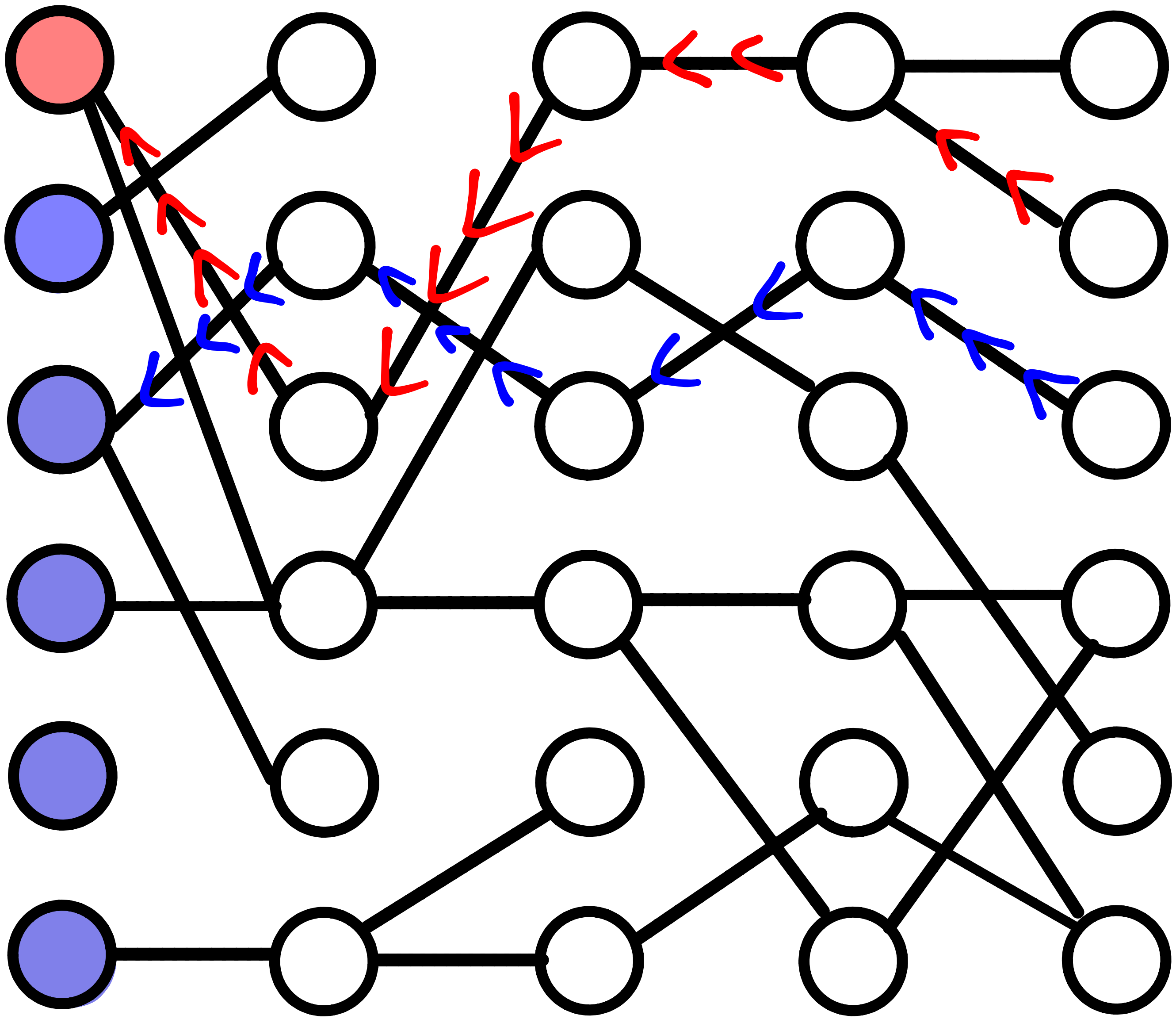
The Discrete ancestral selection graph



1) Vertex v has K_v Directed edges (each of them as in the WF model)

2) $\tau(v) =$  if and only if v is connected to a Red vertex

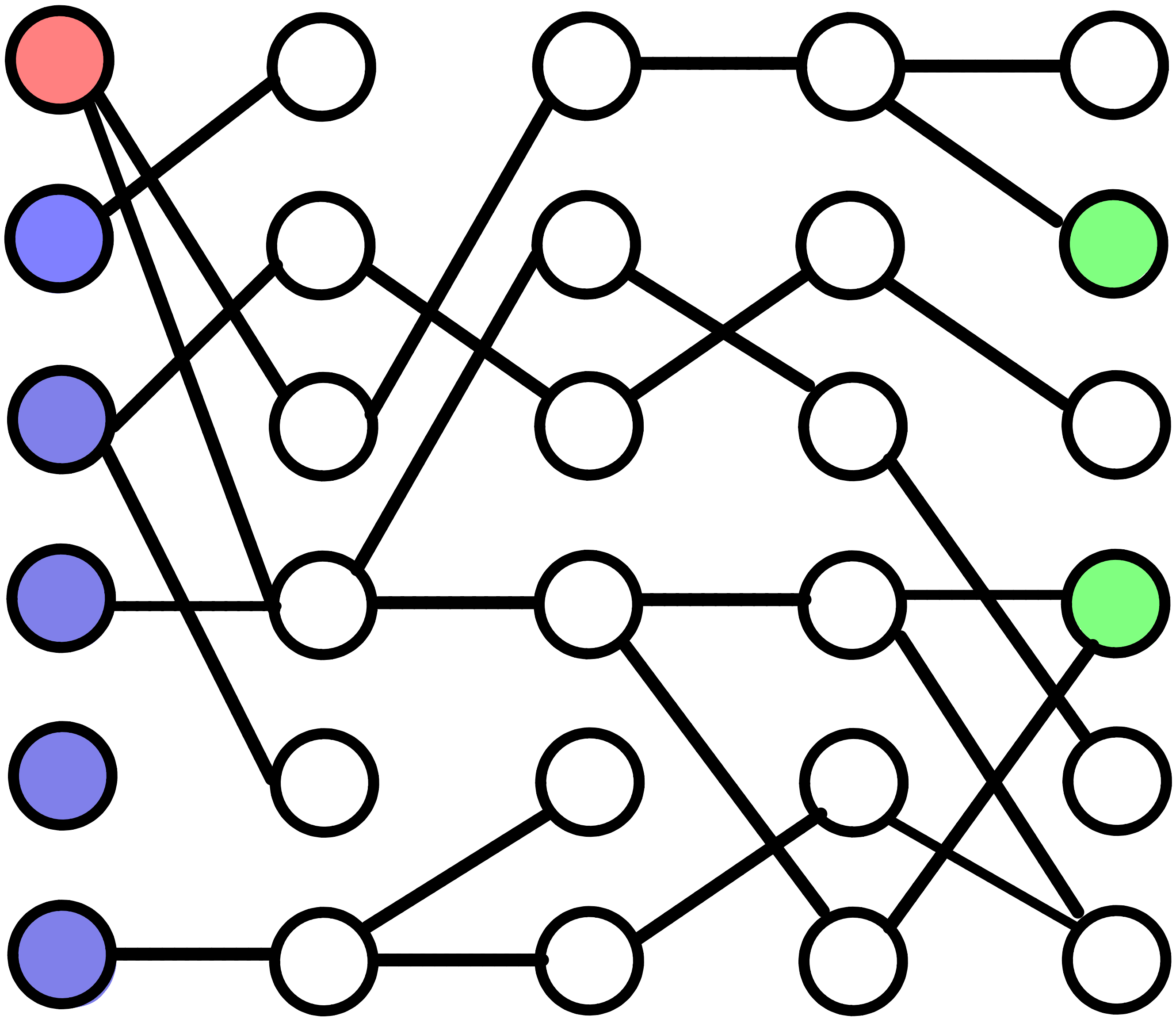
The Discrete ancestral selection graph



1) Vertex v has K_v Directed edges (each of them as in the WF model)

2) $\tau(v) =$ (red vertex) if and only if v is connected to a red vertex

The Discrete ancestral selection graph

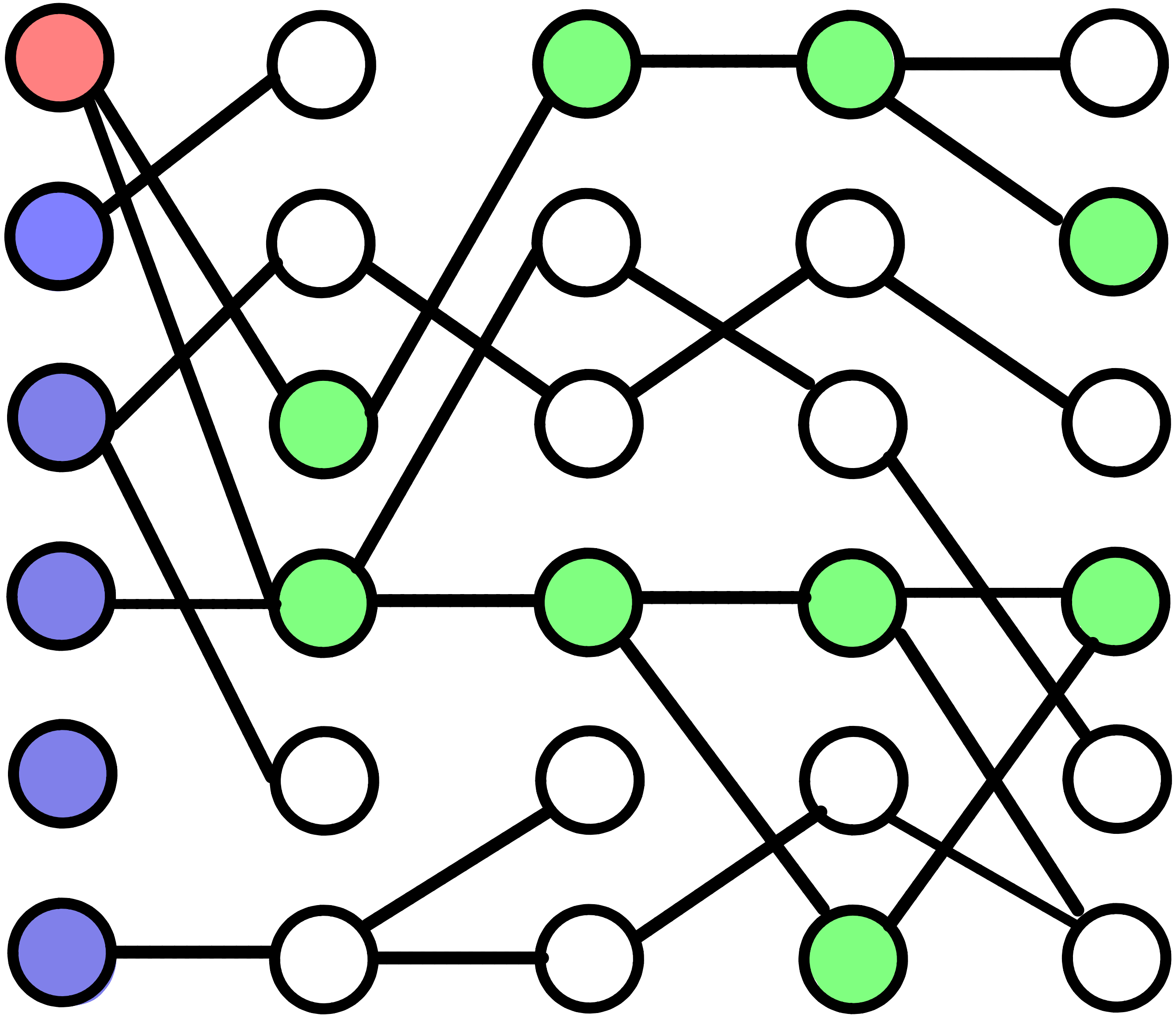


1) Vertex v has K_v Directed edges (each of them as in the WF model)

2) $\tau(v) =$ ● if and only if v is connected to a Red vertex

$A_g =$

The Discrete ancestral selection graph

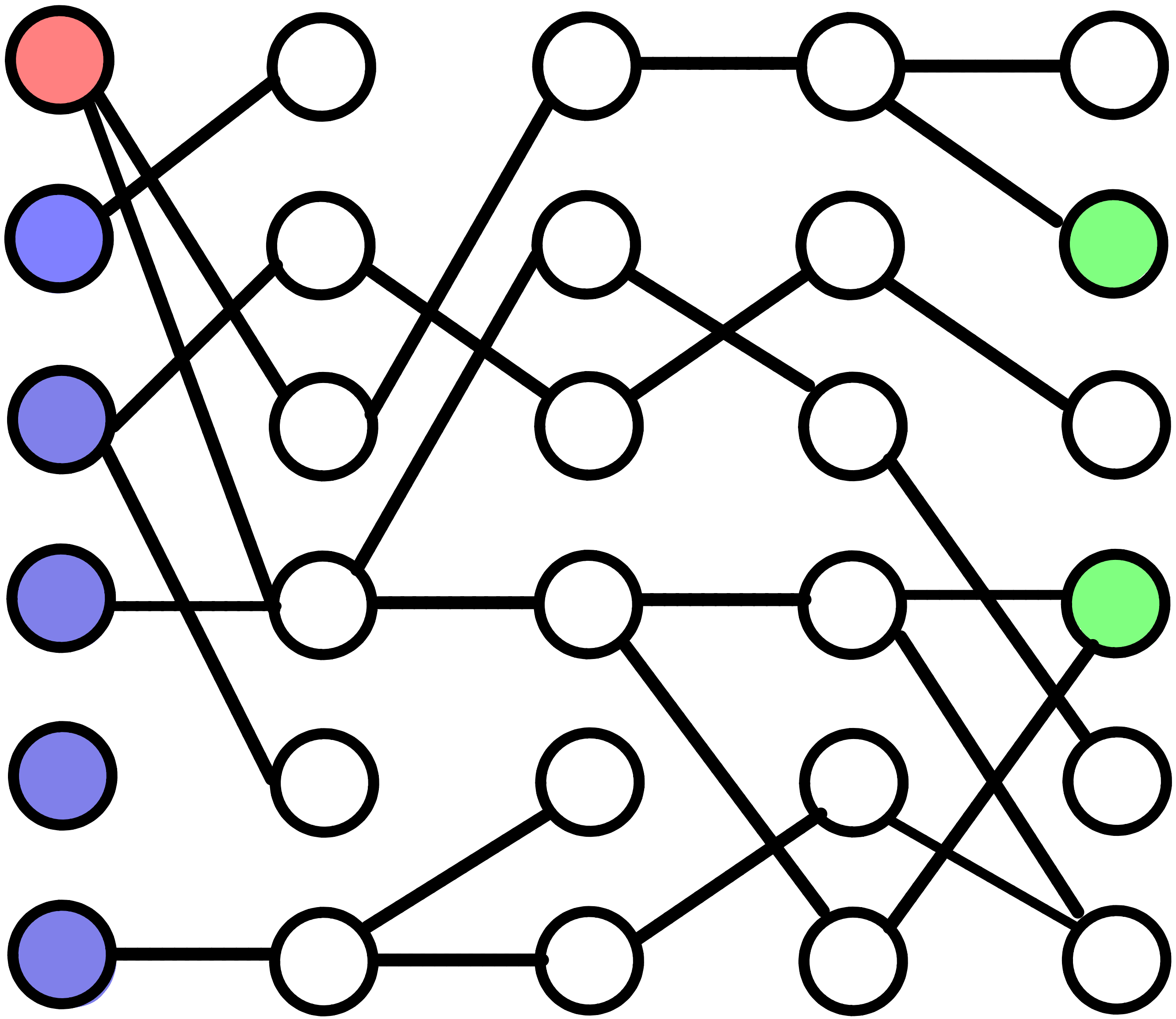


1) Vertex v has K_v Directed edges (each of them as in the WF model)

2) $\tau(v) =$ if and only if v is connected to a red vertex

$A_g = \#$ of in generation g

The Discrete ancestral selection graph



1) Vertex v has

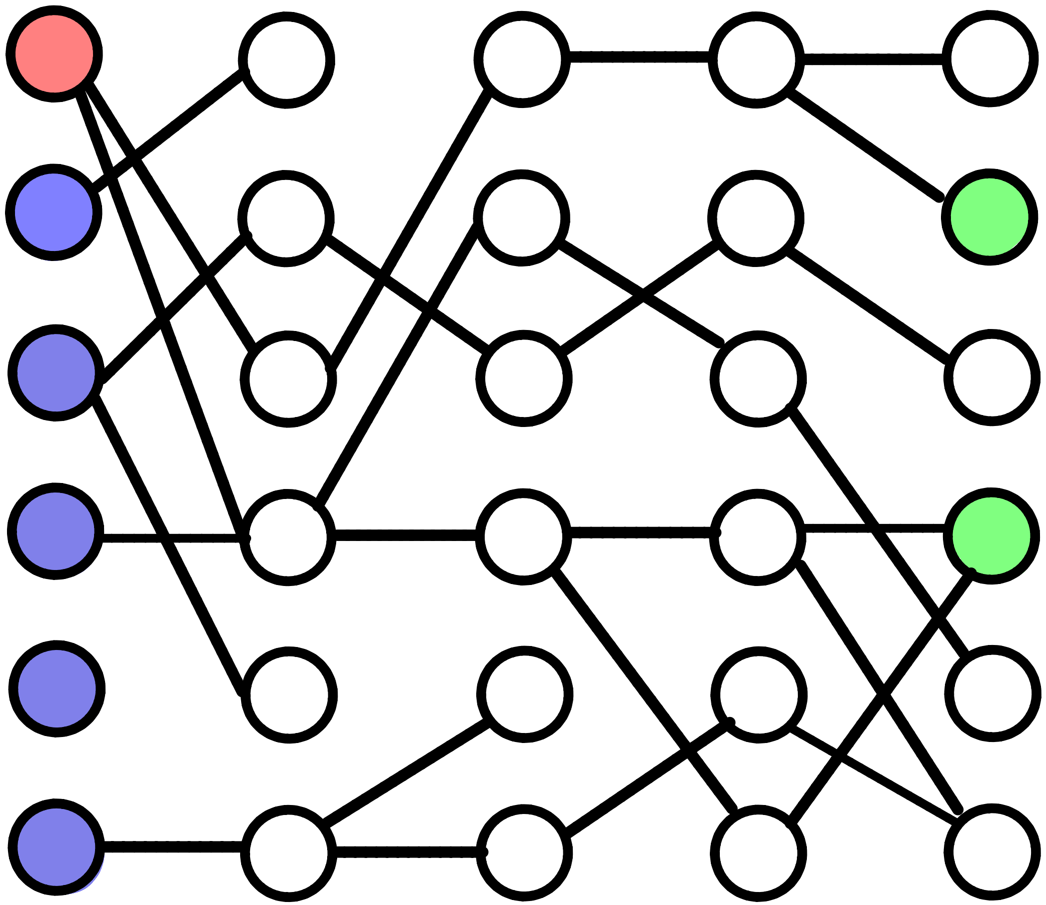
K_v

Directed edges (each of them as in the WF model)

2) $\tau(v) =$ ● if and only if v is connected to a Red vertex

If $\{K_v\}_{v\in V}$ are IID and

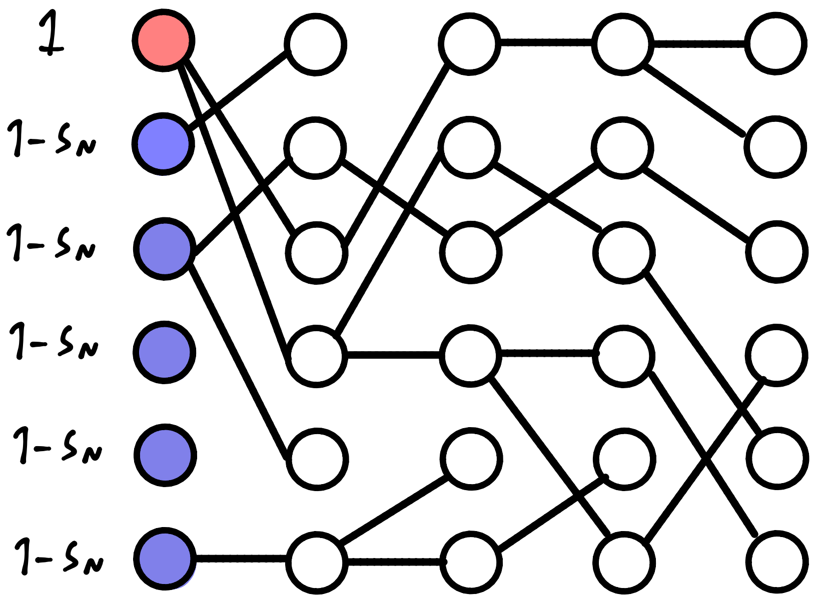
$$\mathbb{P}(K_v = k) = (1-s_N)\, s_N^{k-1}$$



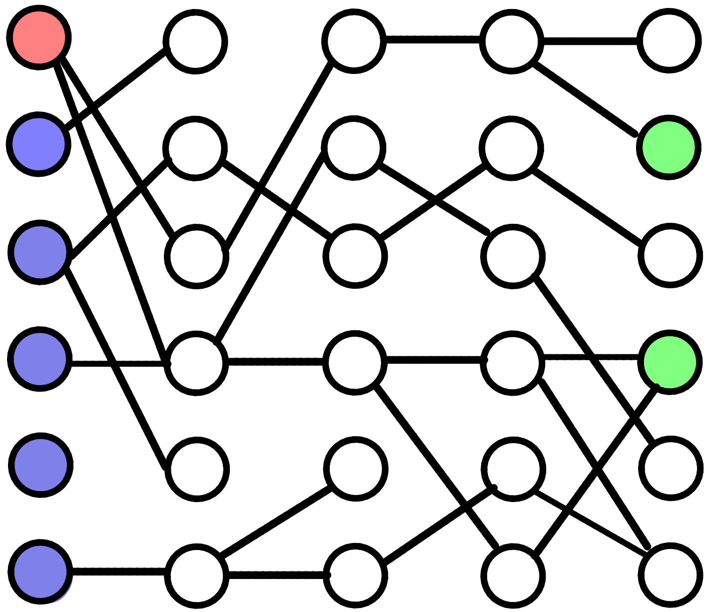
$$\bar{X}_g = \frac{\#\ \text{in generation } g}{N}$$

IF $\{K_v\}_{v \in V}$ are IID and

$$\mathbb{P}(K_v = k) = (1 - s_N)\, s_N^{k-1}$$



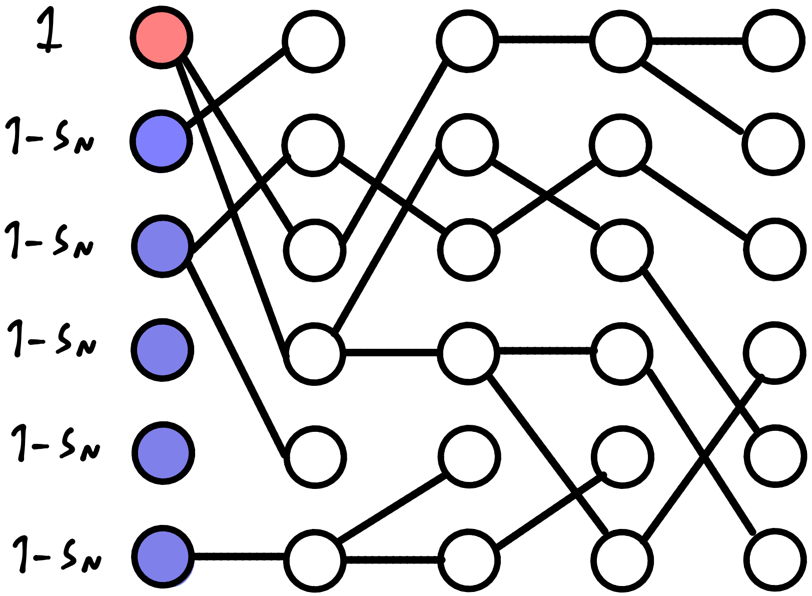
$X_g = \dfrac{\#\ \text{in generation } g}{N}$



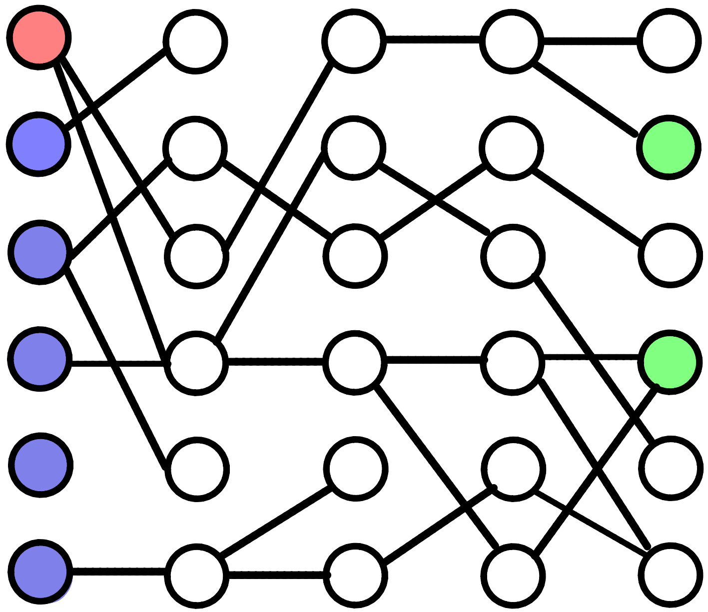
$\bar{X}_g = \dfrac{\#\ \text{in generation } g}{N}$

If $\{K_v\}_{v \in V}$ are IID and

$$\mathbb{P}(K_v = k) = (1 - s_N)\, s_N^{\,k-1}$$



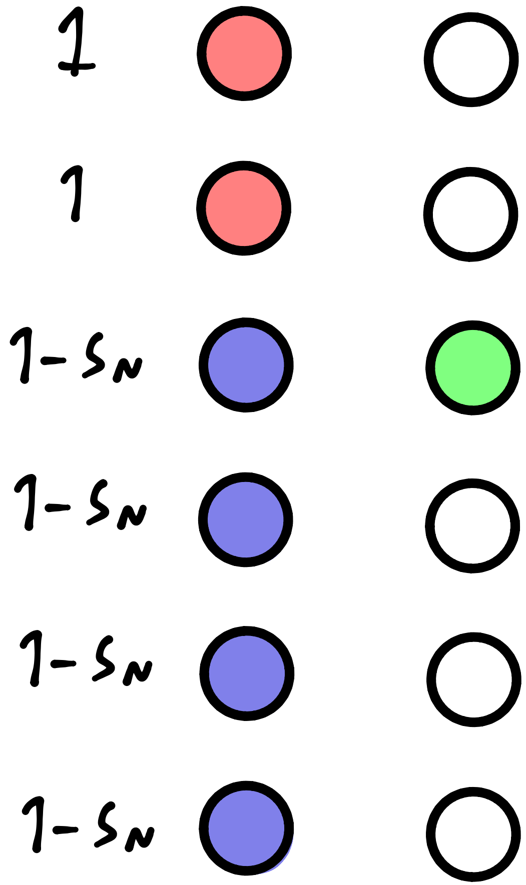
$$X_g = \frac{\#\ \text{in generation } g}{N}$$



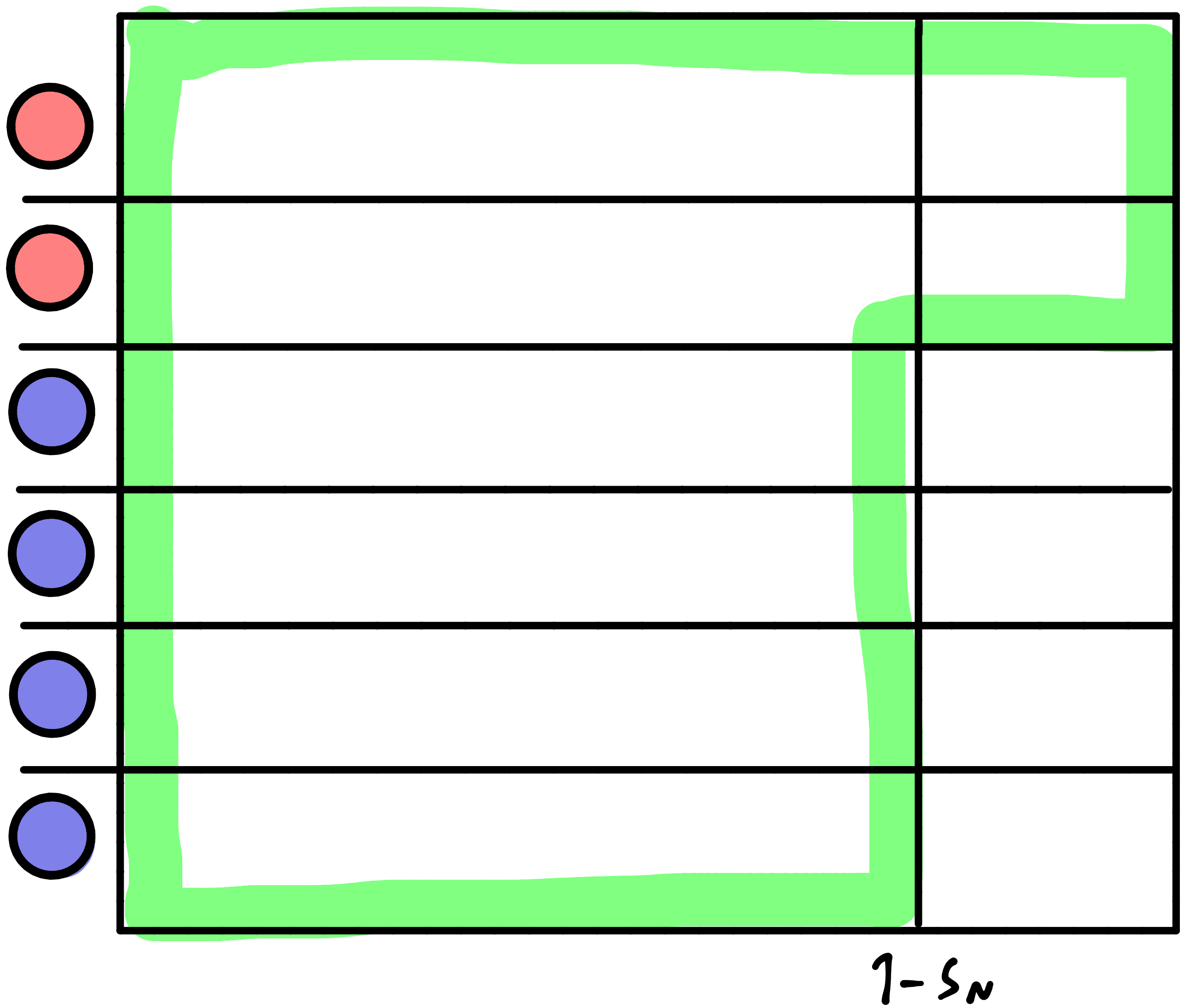
$$\bar{X}_g = \frac{\#\ \text{in generation } g}{N}$$

$$(X_g) \overset{d}{=} (\bar{X}_g)$$

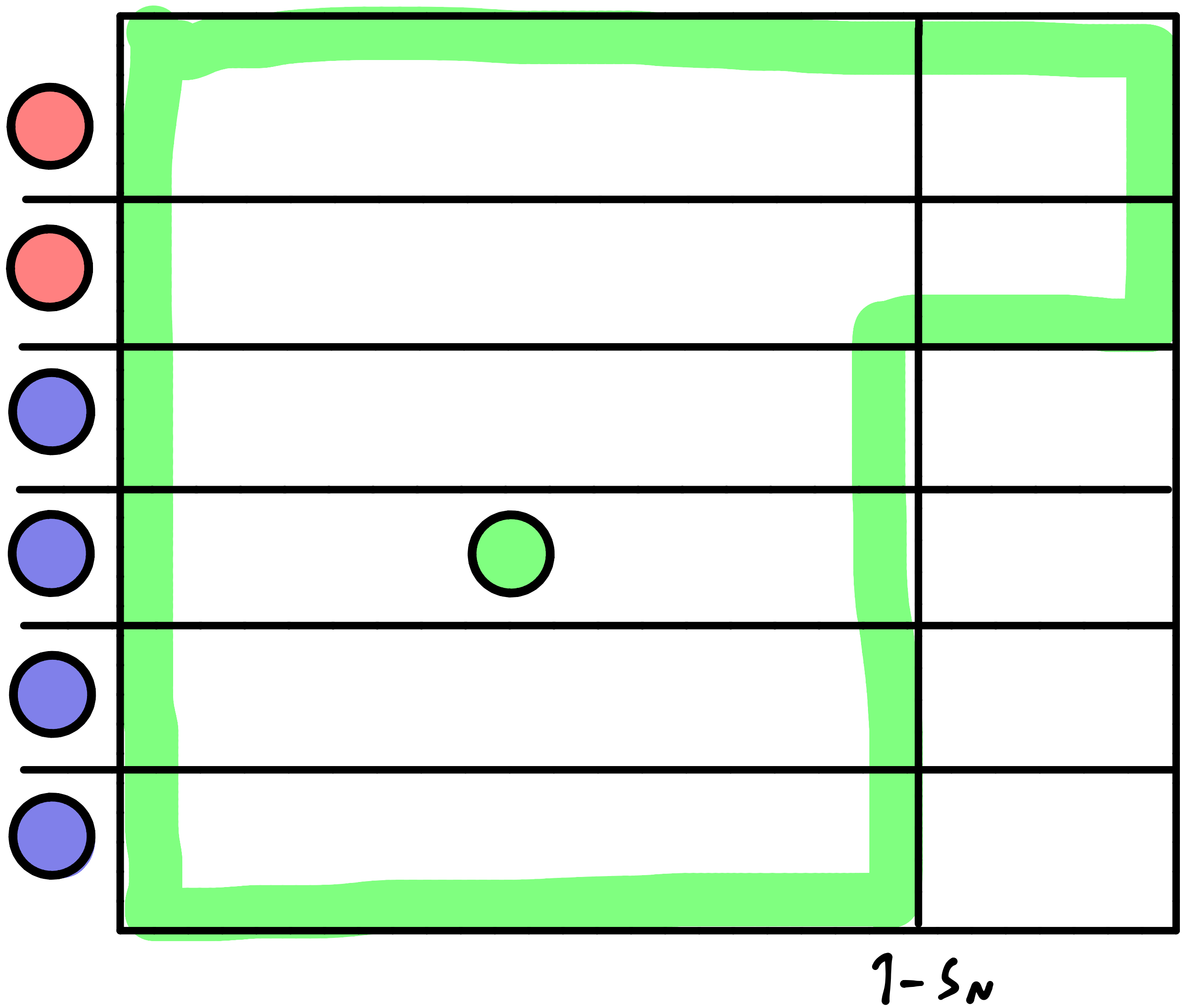
Why geometric?



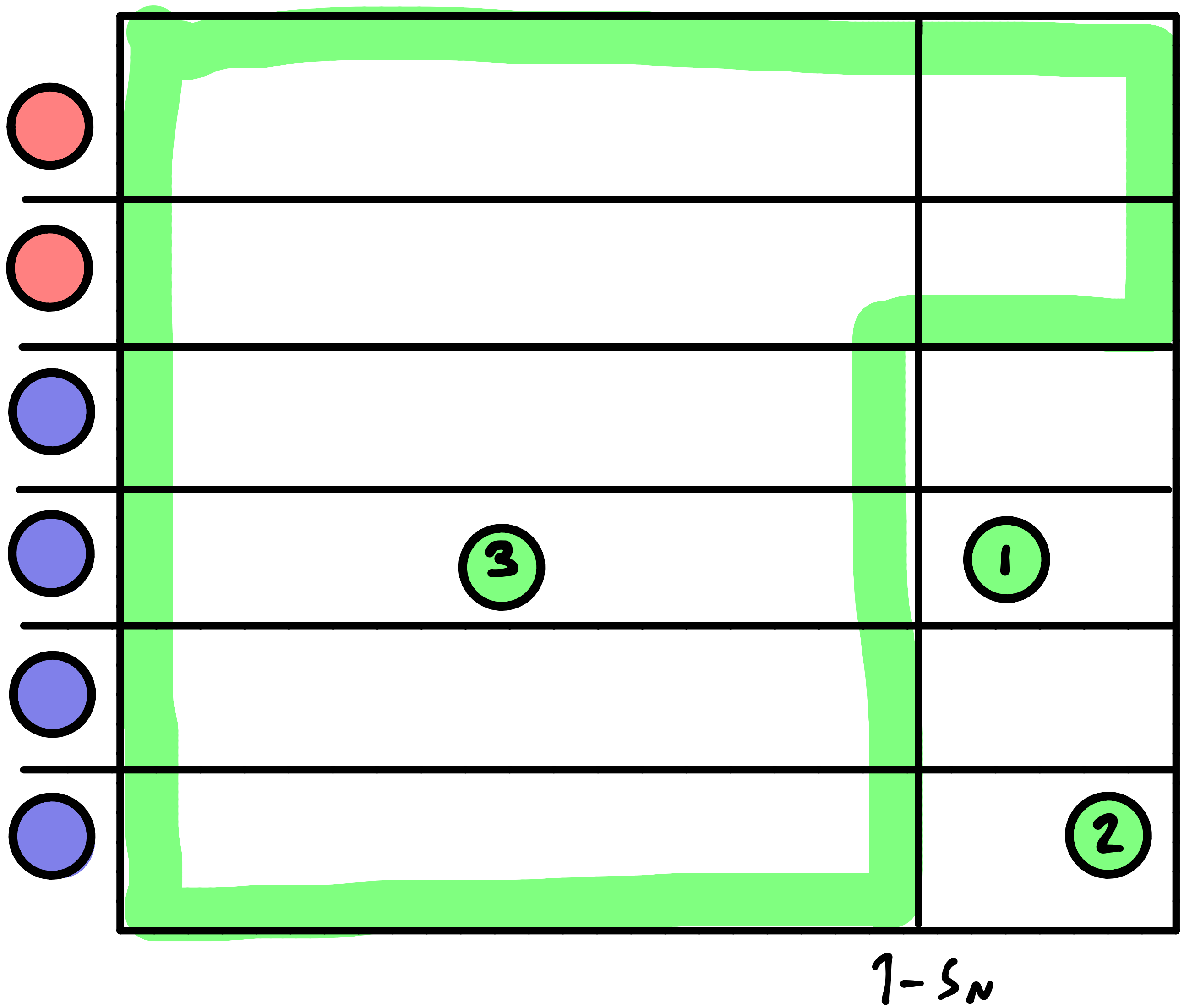
Why geometric?



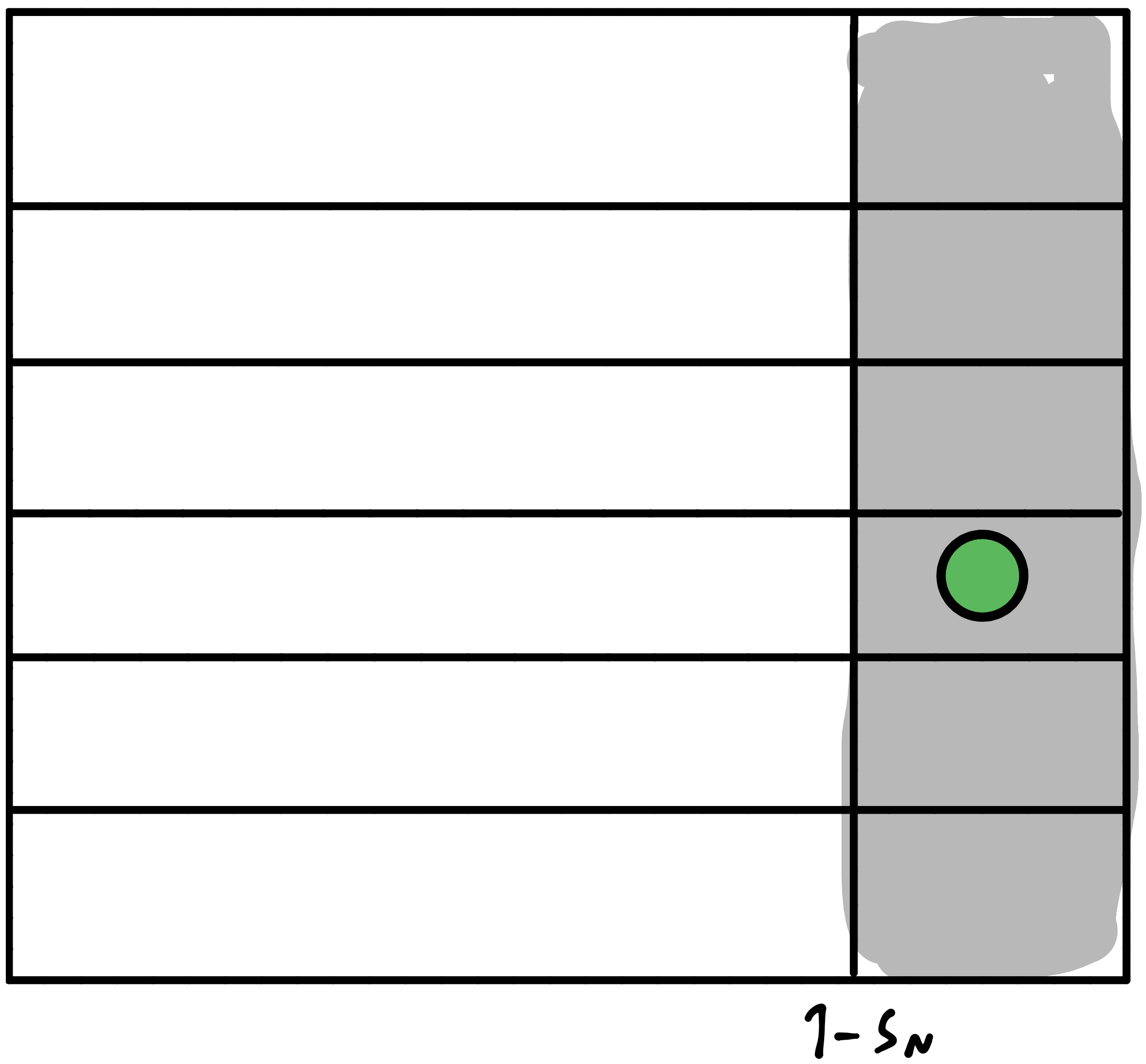
Why geometric?



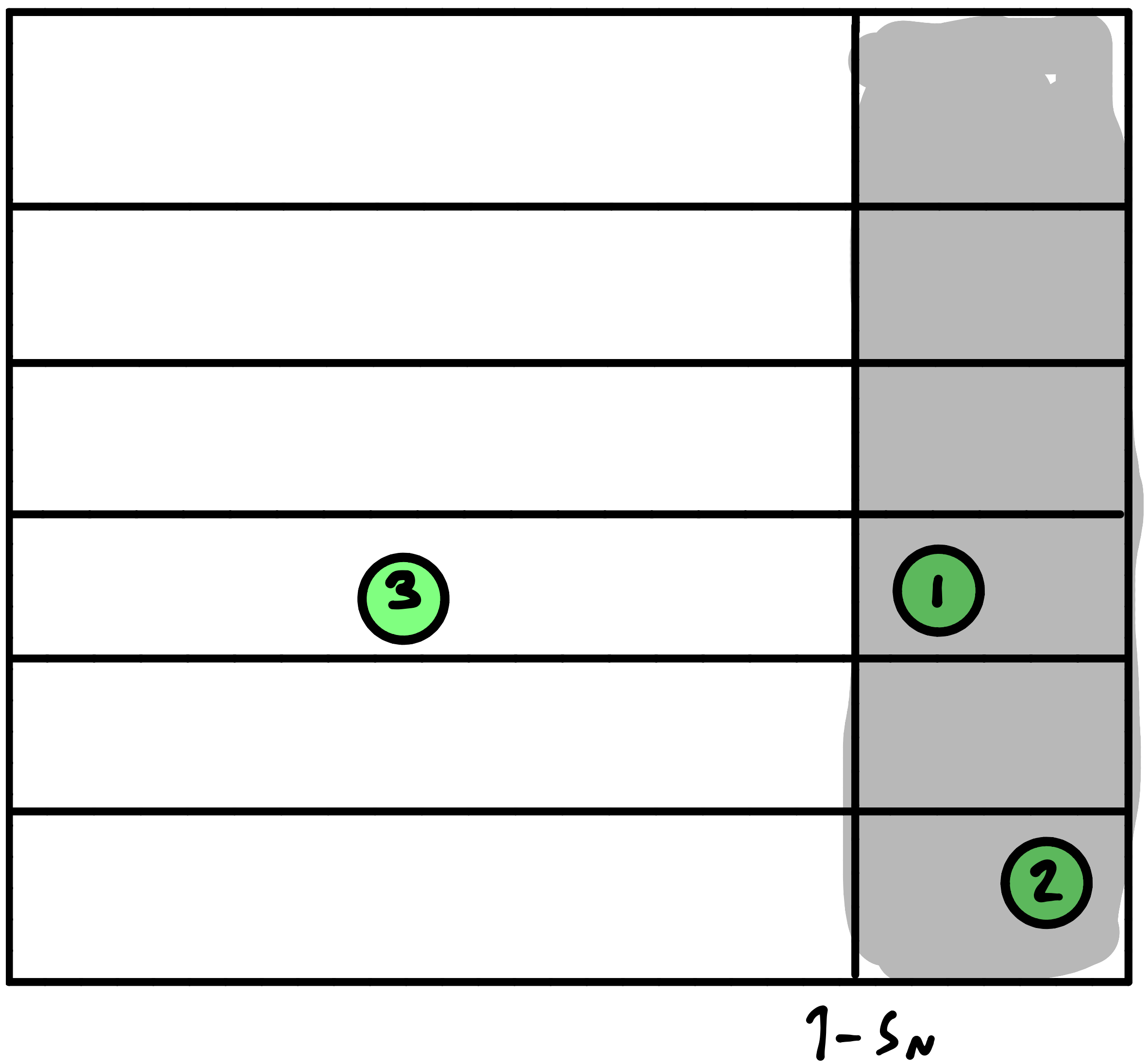
Why geometric?



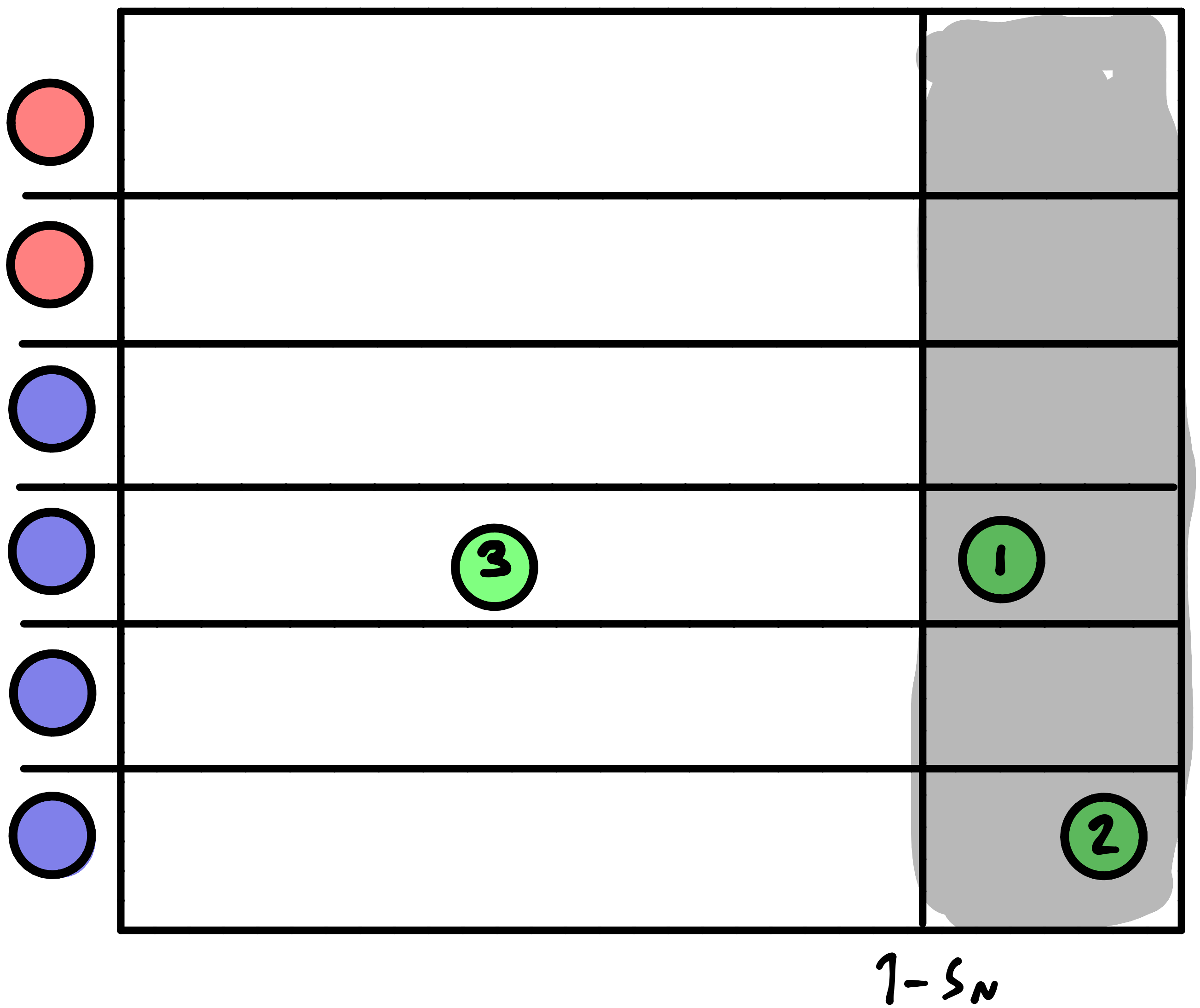
Why geometric?



Why geometric?

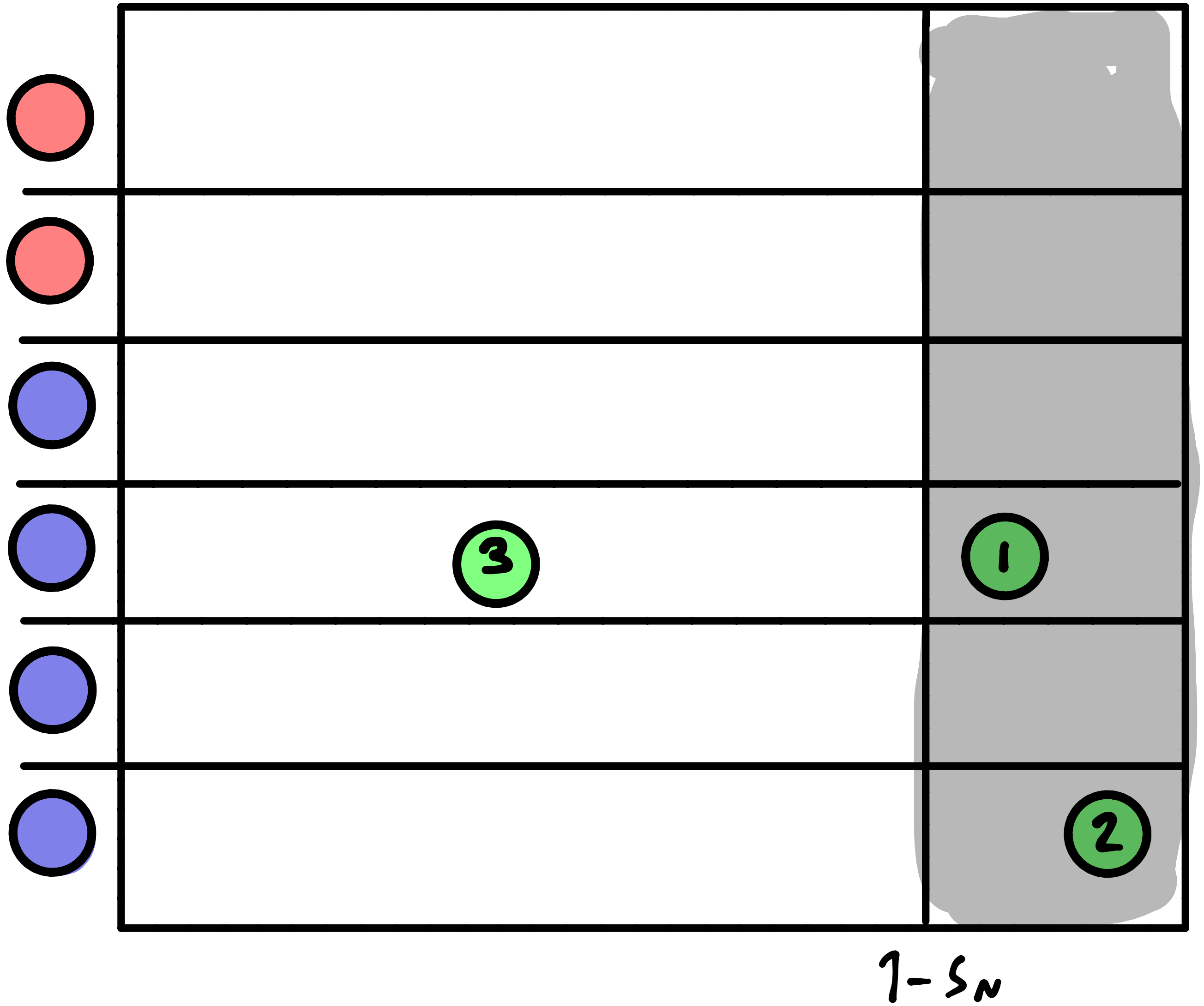


Why geometric?



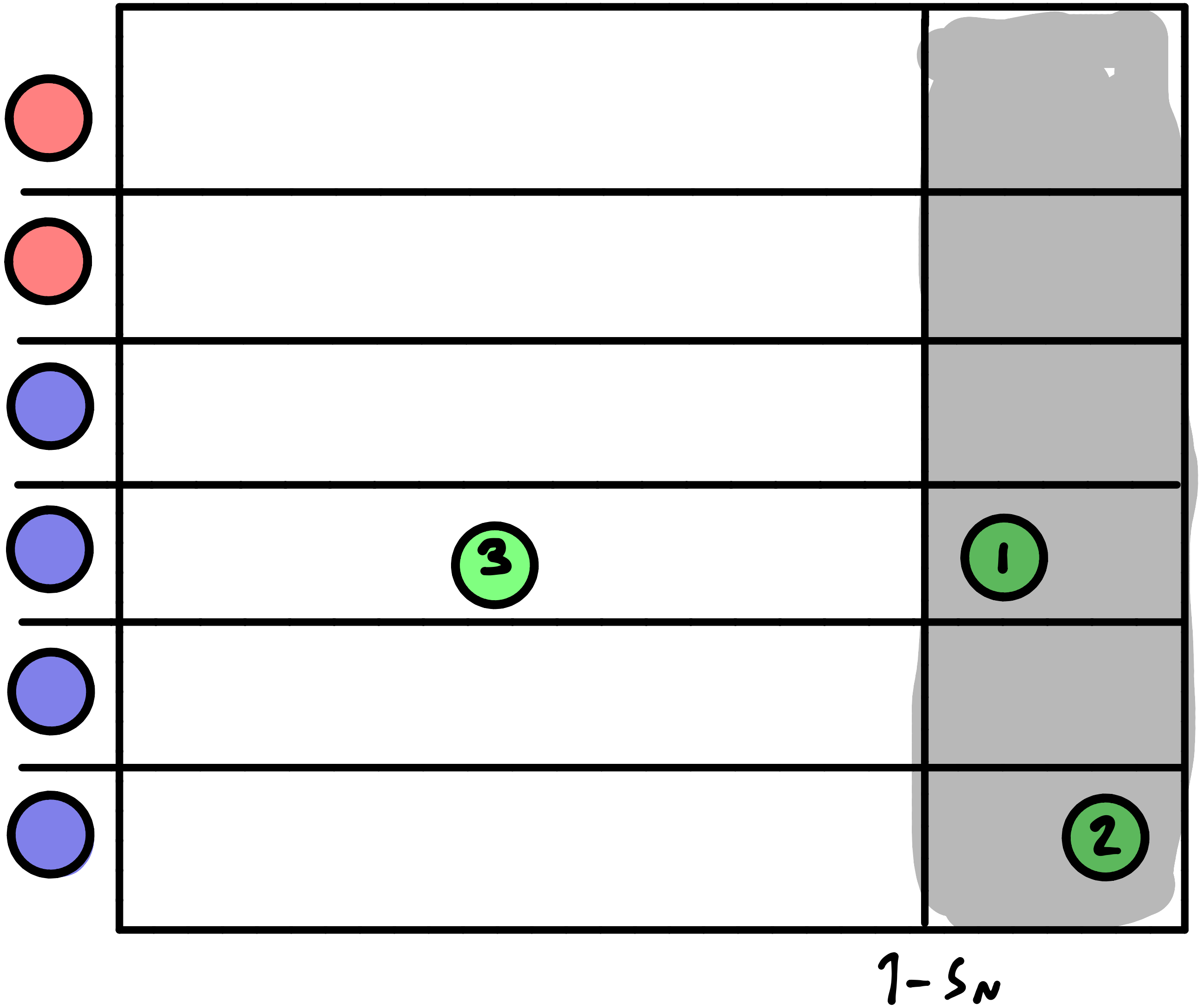
Why geometric?

$\# \text{Picks} \sim \text{geo}(1 - S_N)$



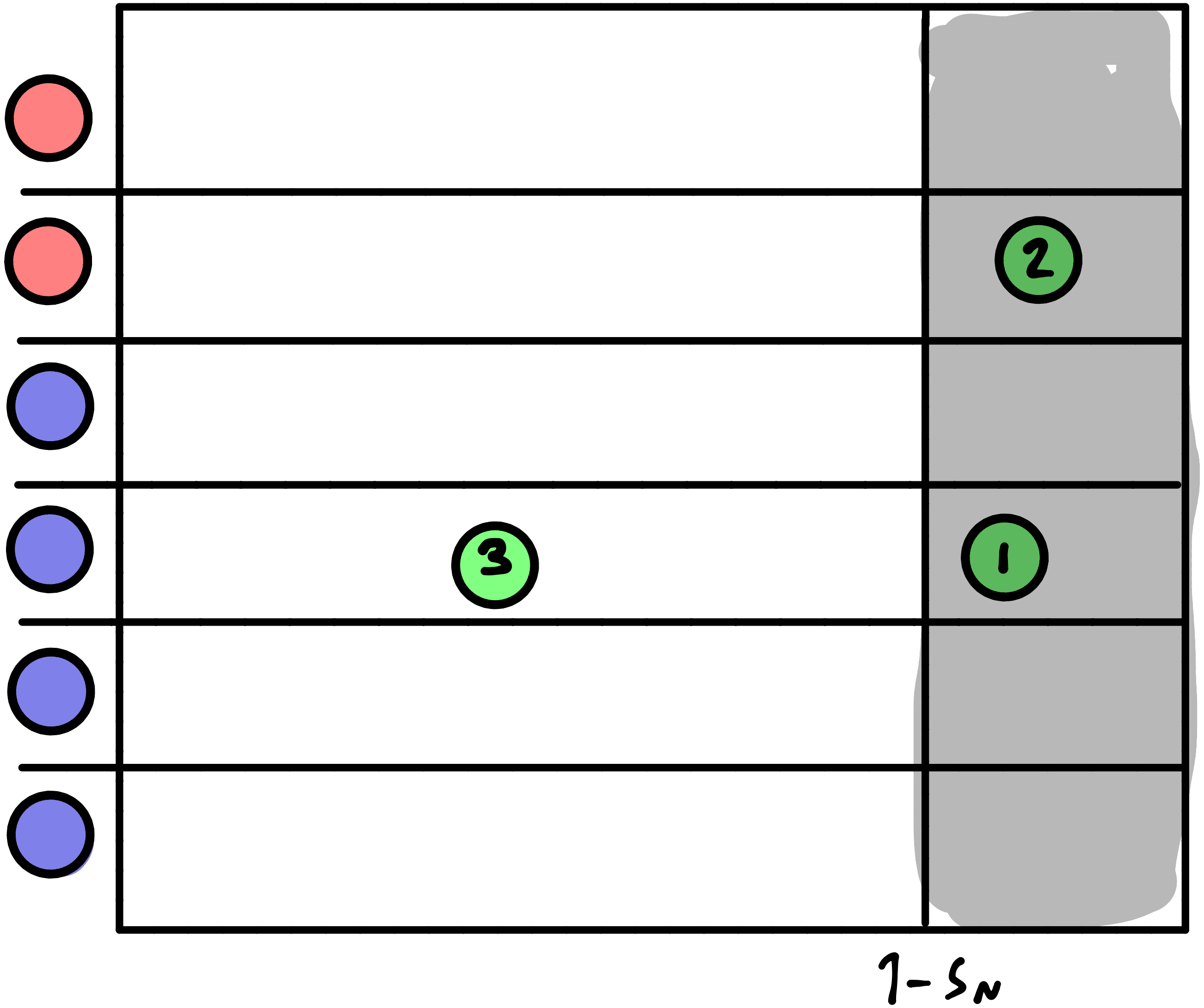
Why geometric?

(green circle) = (blue circle) if and only if all picks are (blue circle)

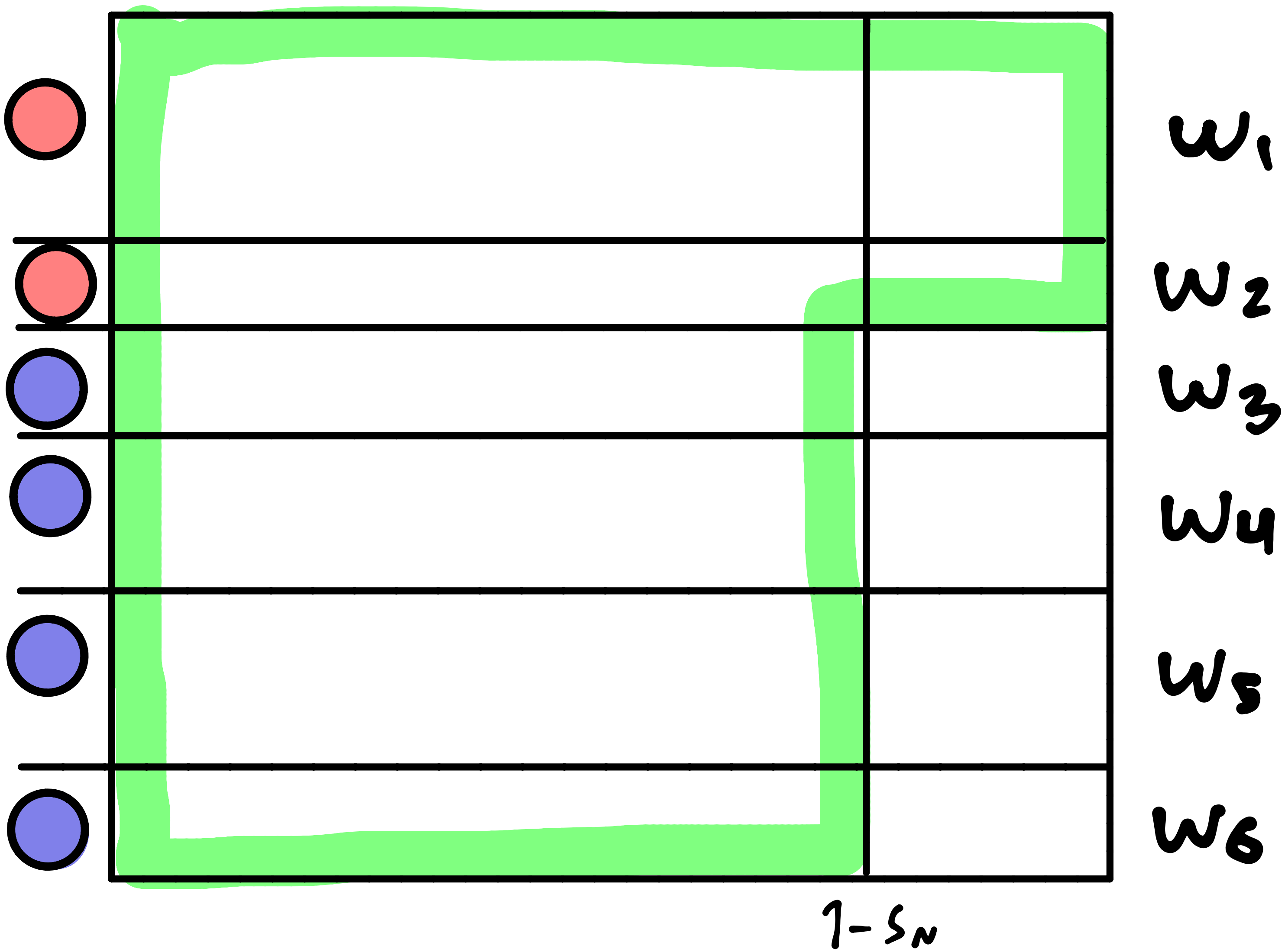


Why geometric?

= if and only if all picks are



Paint box



Theorem: Haldane's formula for Cannings-Paintbox models

(Boenkost, GC, Pokalyuk, Wakolbinger)

Theorem: Haldane's formula for Canning-Paintbox models

Let $(W_{1,g}, \dots, W_{N,g})_{g \in \mathbb{Z}}$ be iid RV such that $\sum_{i=1}^{N} W_{i,g} = 1$ for all $g \in \mathbb{Z}$, $E[W_v^2] = \frac{\rho^2}{N^2}$ and $\mathbb{E}[W_v^3] = O(N^{-3})$,

Theorem: Haldane's formula for Cannings-Paintbox models

Let $(W_{1,g}, \dots, W_{N,g})_{g \in \mathbb{Z}}$ be iid RV such that $\sum_{i=1}^{N} W_{i,g} = 1$ for all $g \in \mathbb{Z}$, $E[W_r^2] = \frac{\rho^2}{N^2}$ and $\mathbb{E}[W_r^3] = O(N^{-3})$

$$S_N = N^{-b} \text{ for } b \in (2/3, 1]$$

Theorem: Haldane's formula for Cannings-Paintbox models

Let $(W_{1,g},\dots,W_{N,g})_{g\in\mathbb{Z}}$ be iid RV such that $\sum_{i=1}^{N} W_{i,g}=1$ for all $g\in\mathbb{Z}$ $E[W_v^2]=\frac{\rho^2}{N^2}$ and $\mathbb{E}[W_v^3]=O(N^{-3})$

$S_N=N^{-b}$ for $b\in(2/3,1]$

Let (X_g) be the frequency process of the Paintbox model with parameters S_N and (W_v)

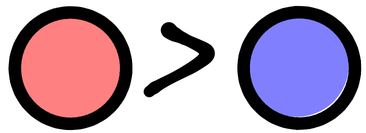
$$\mathbb{P}_{1/N}(\mathrm{Fix})=\frac{S_N}{(\rho/2)}+o(S_N)$$

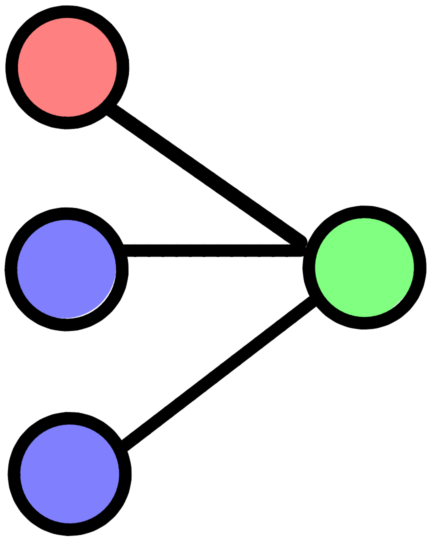
Proof:

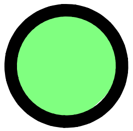
1) Hypergeometric Duality

2) Couple the DASG with the ASG

The Discrete ancestral selection Graph

Coloring Rule: 

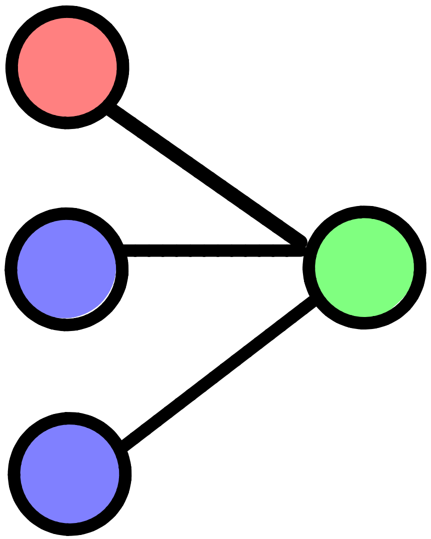


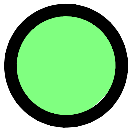
 is Red if and only if it has a Red Parent

Voting Schemes in Sarah's talk

The Discrete ancestral Selection Graph

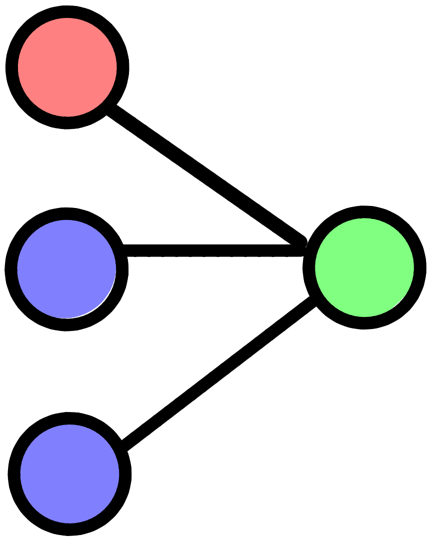
Coloring Rule: Majority Rule



 is Red if and only if it has more Red Parent than blue Parents

The Discrete ancestral selection graph

Coloring Rule:

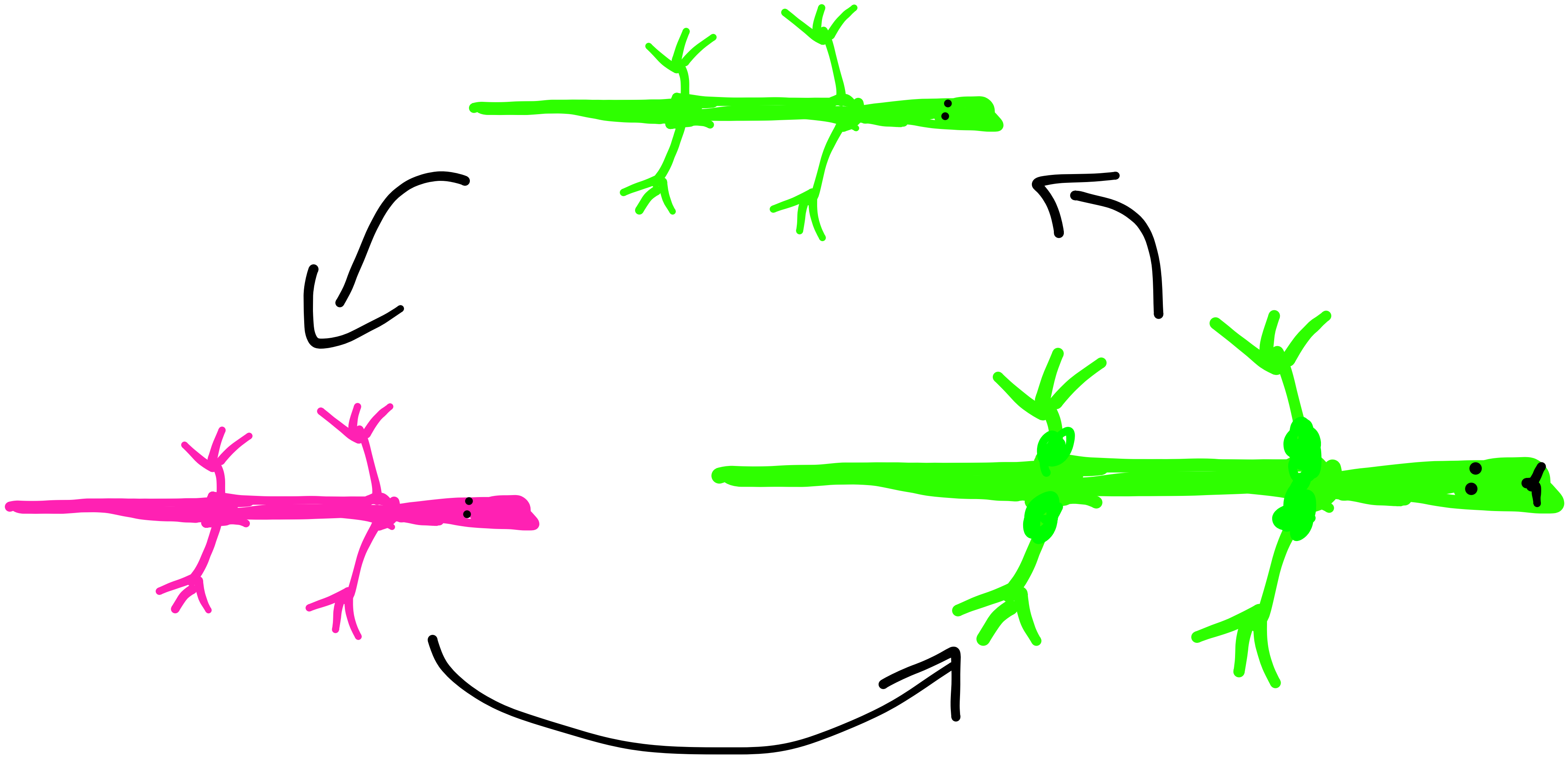


$$dX_t = (P(X_t) - X_t)\,dt + \sqrt{X_t(1-X_t)}\,dB_t$$

See Cordero, Hummel, Schertzer and G.C. Smadi

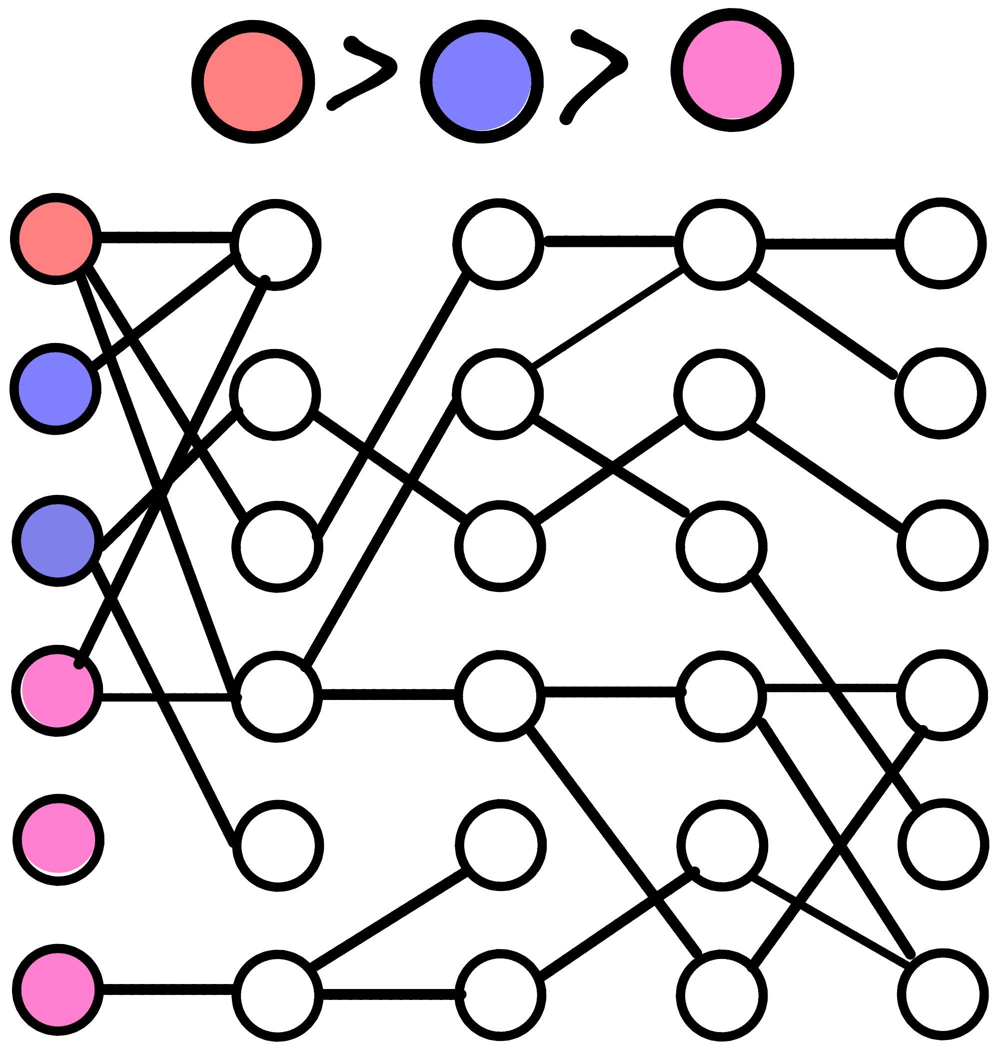
Multitype Complex interaction

G.C. Smadi

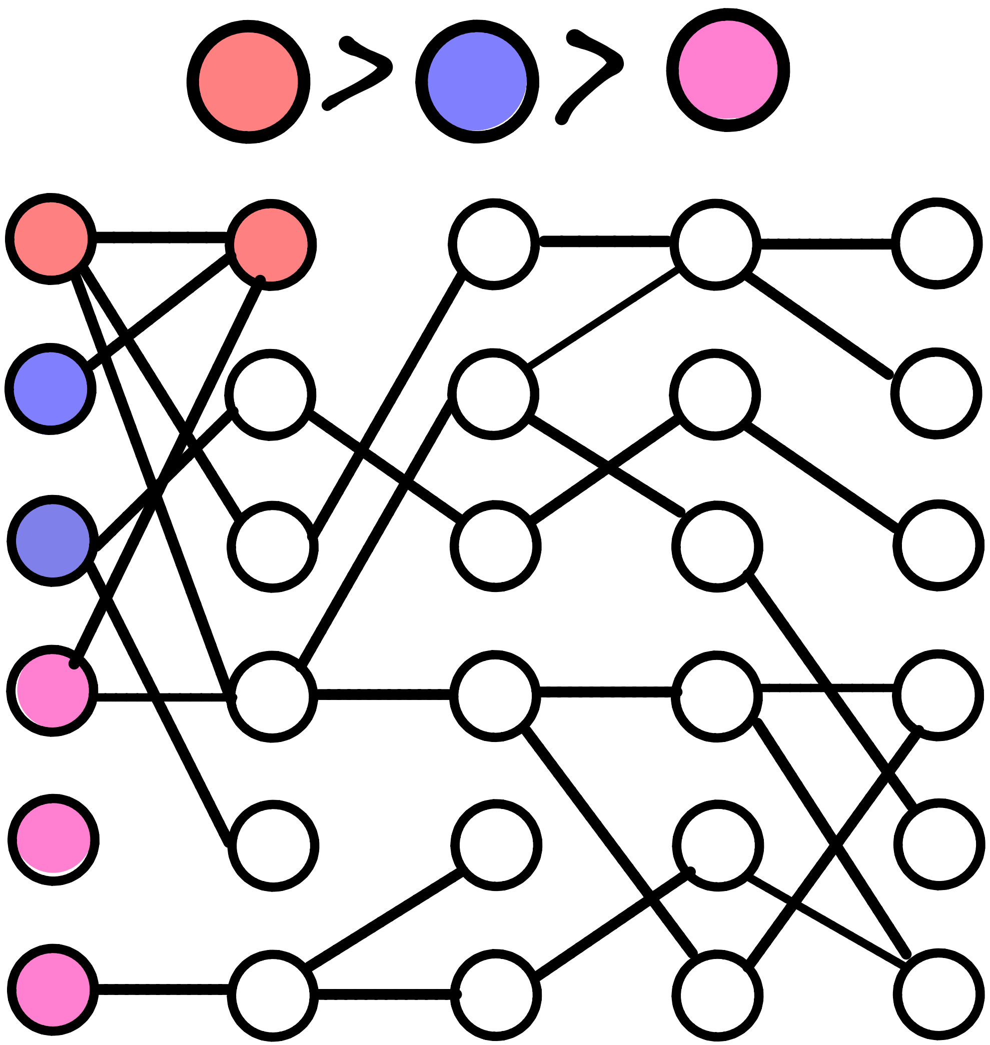


Multitype Complex interaction

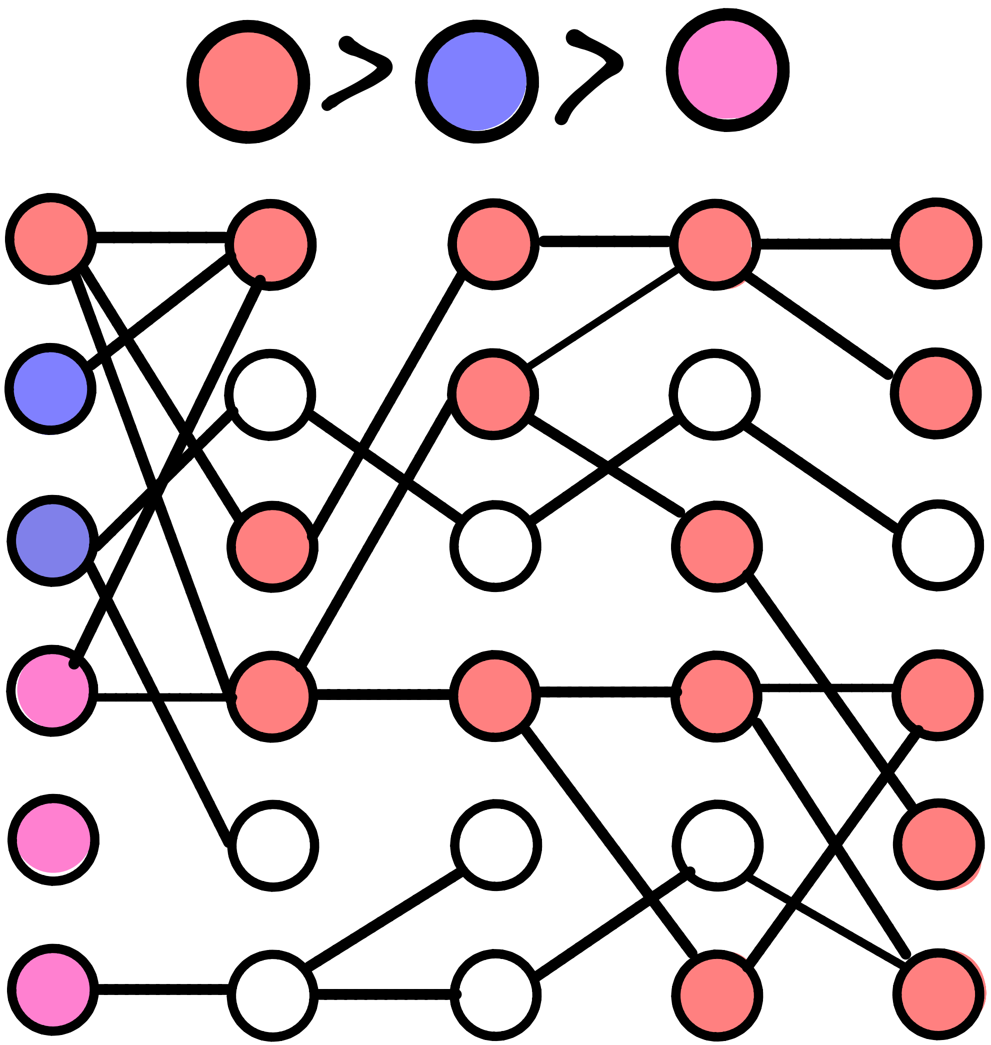
(G.C. Smadi)



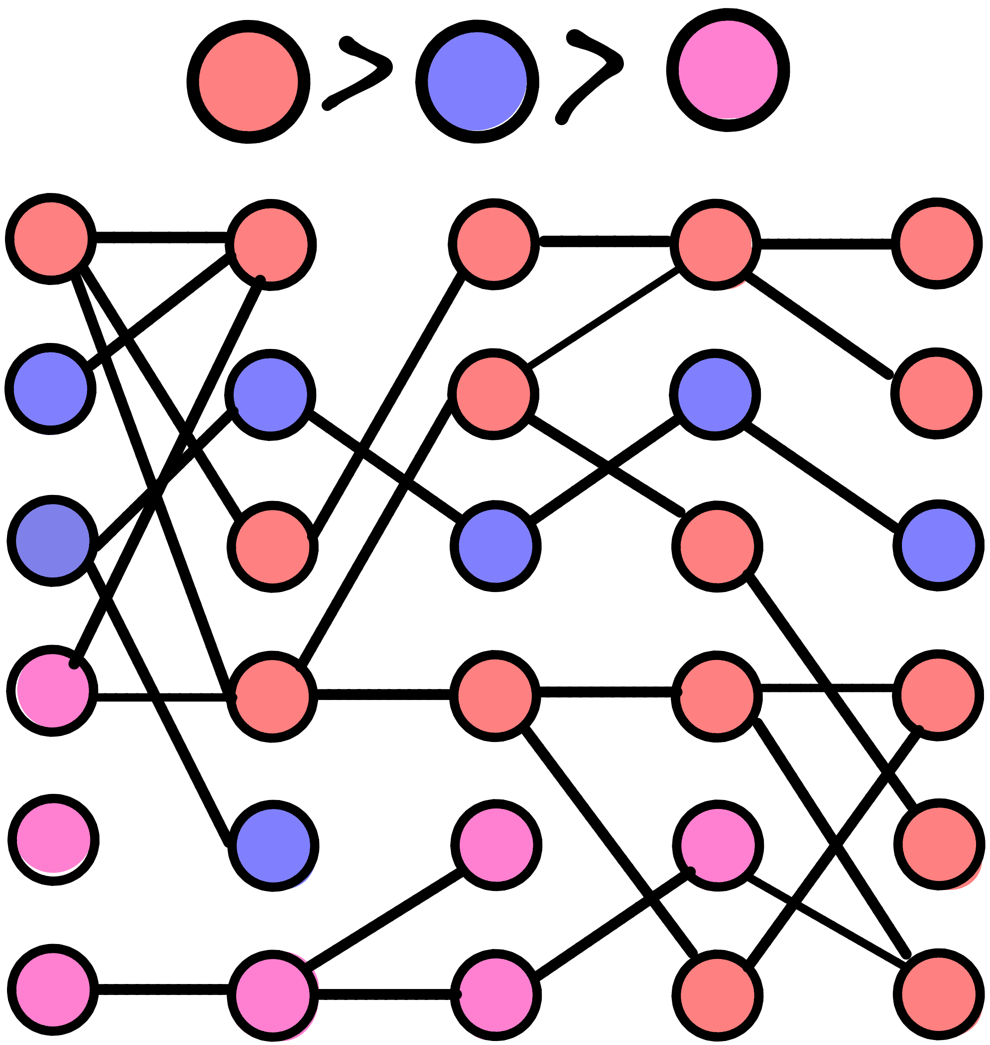
Multitype Complex interaction



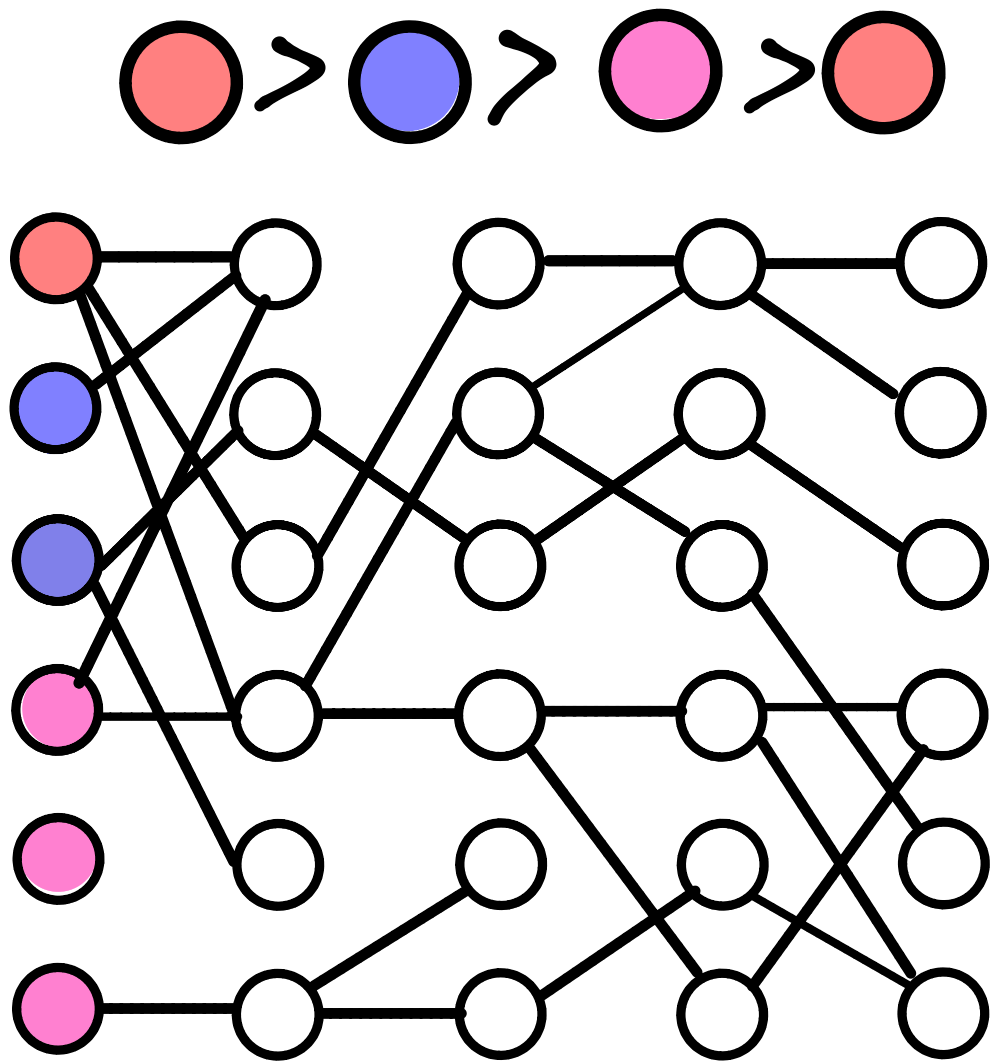
Multitype Complex interaction



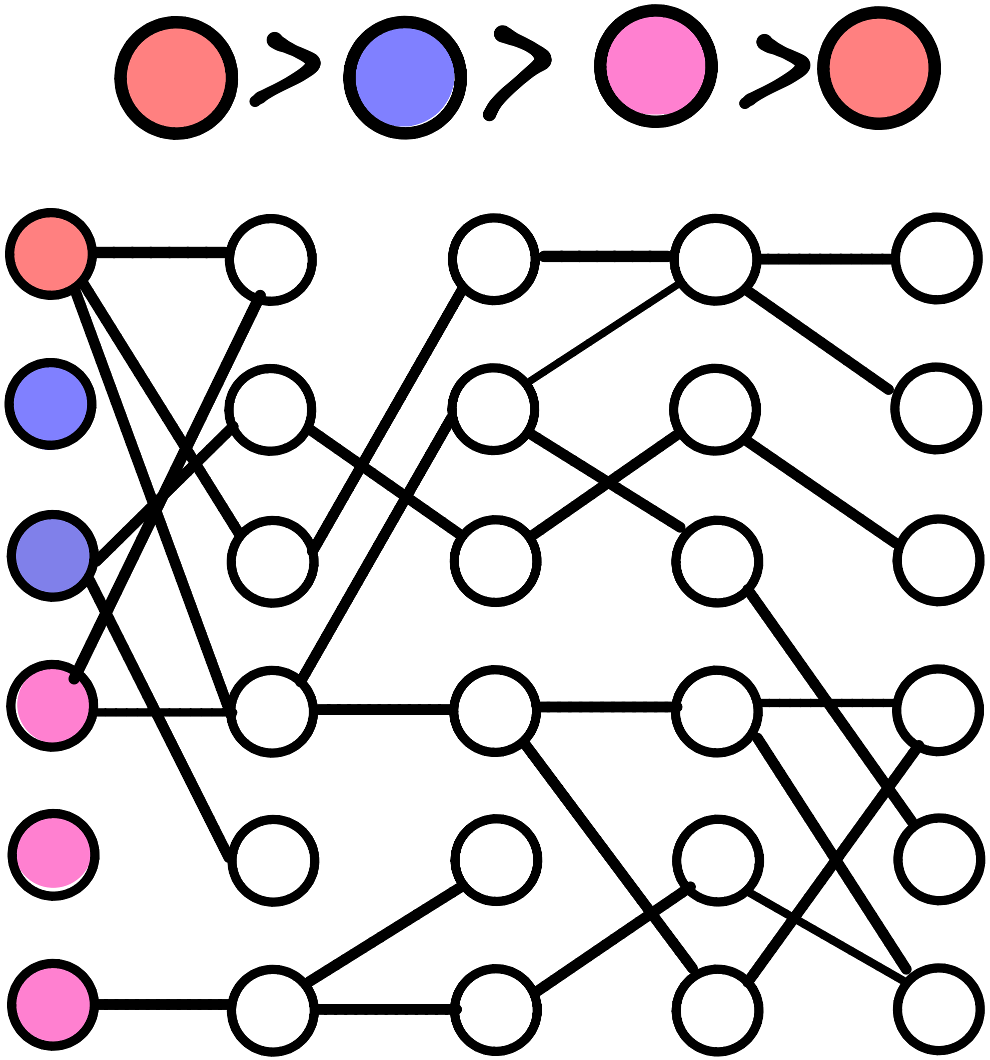
Multitype Complex interaction



Multitype Complex interaction

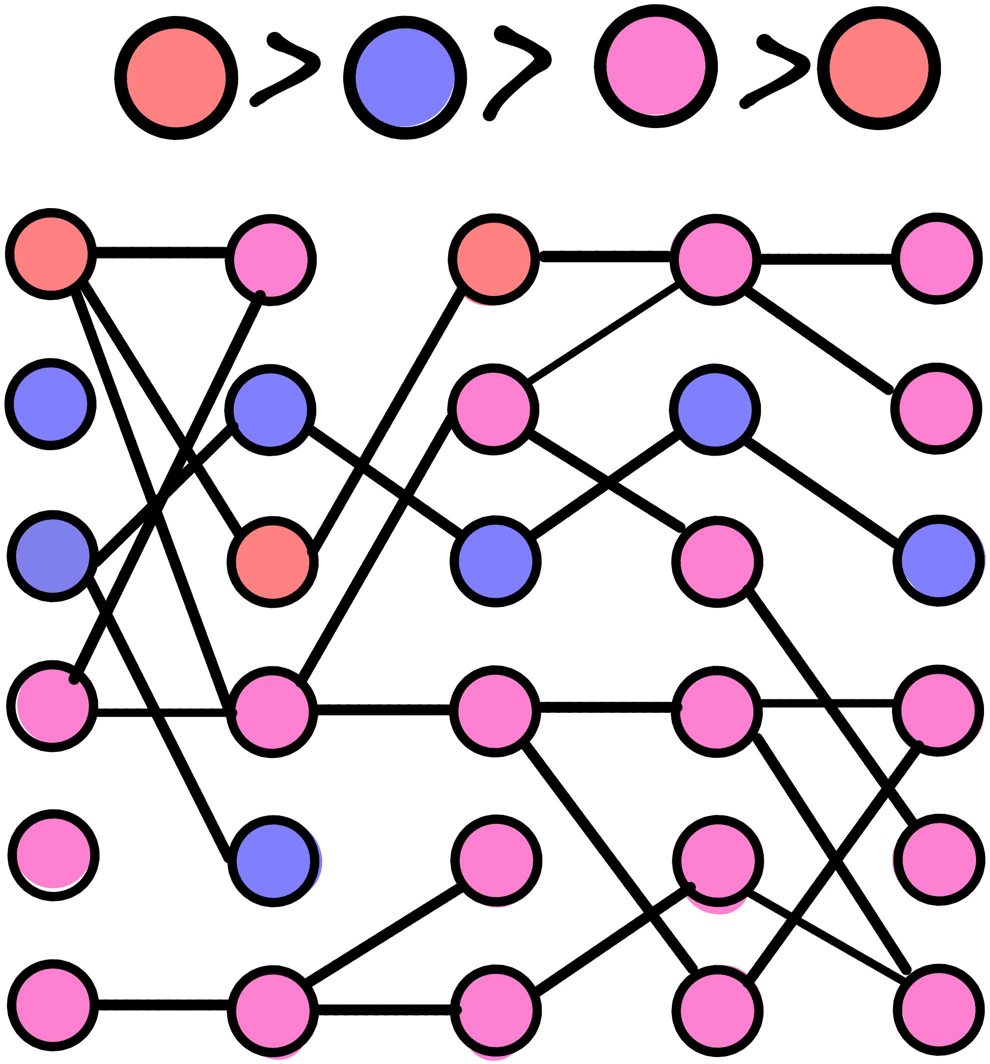


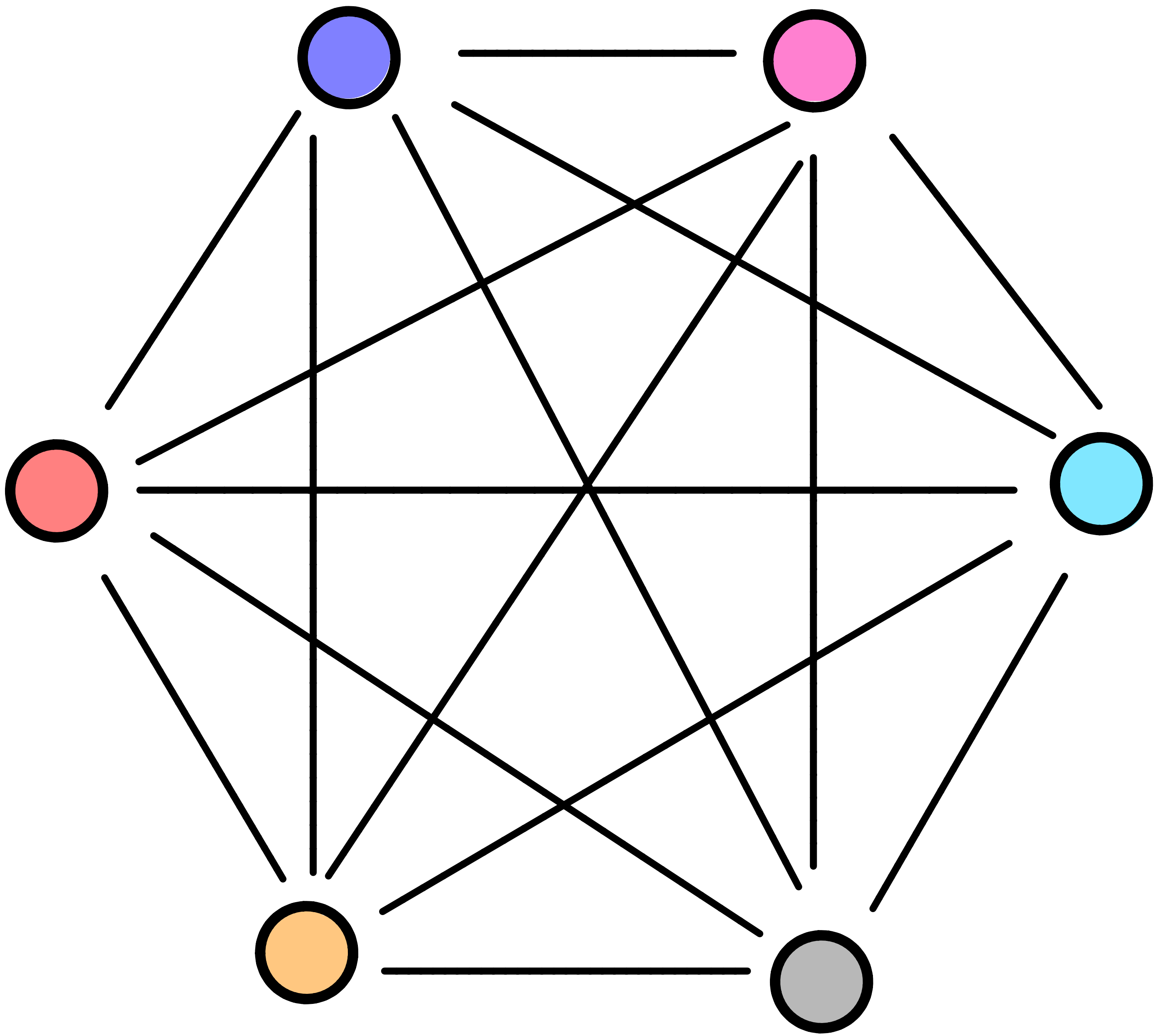
Multitype Complex interaction

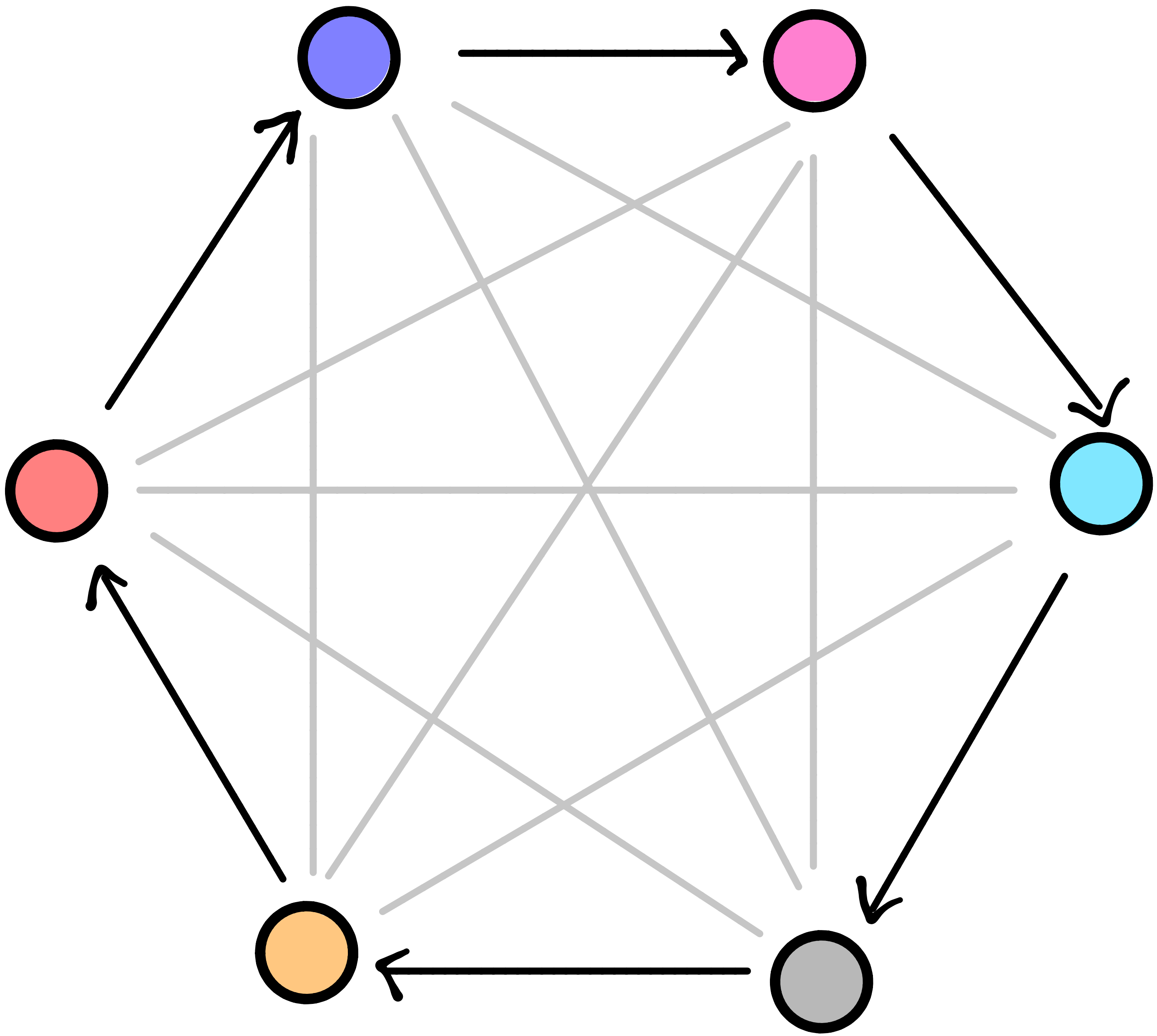


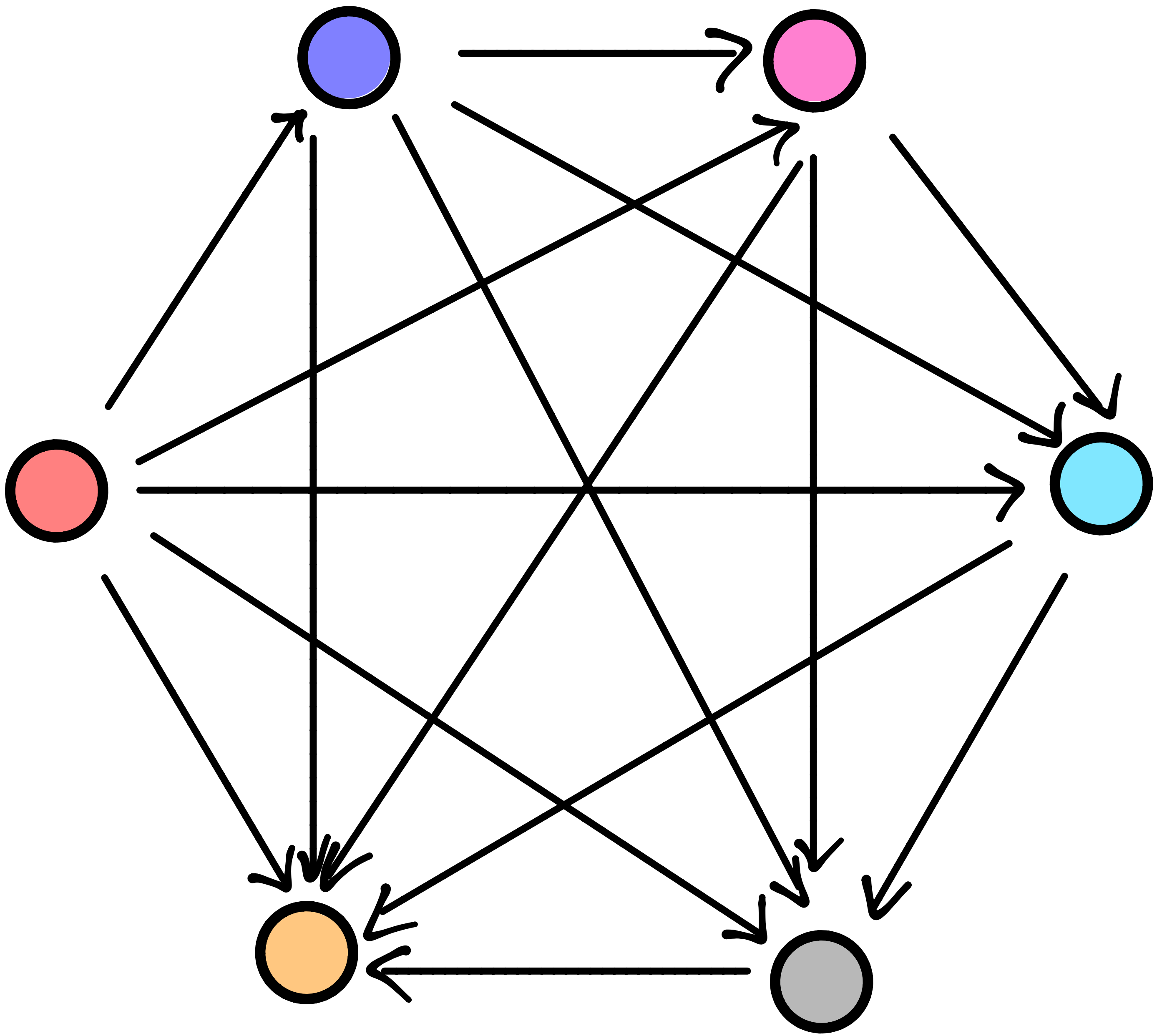
$K_v \in \{1, 2\}$

Multitype Complex interaction

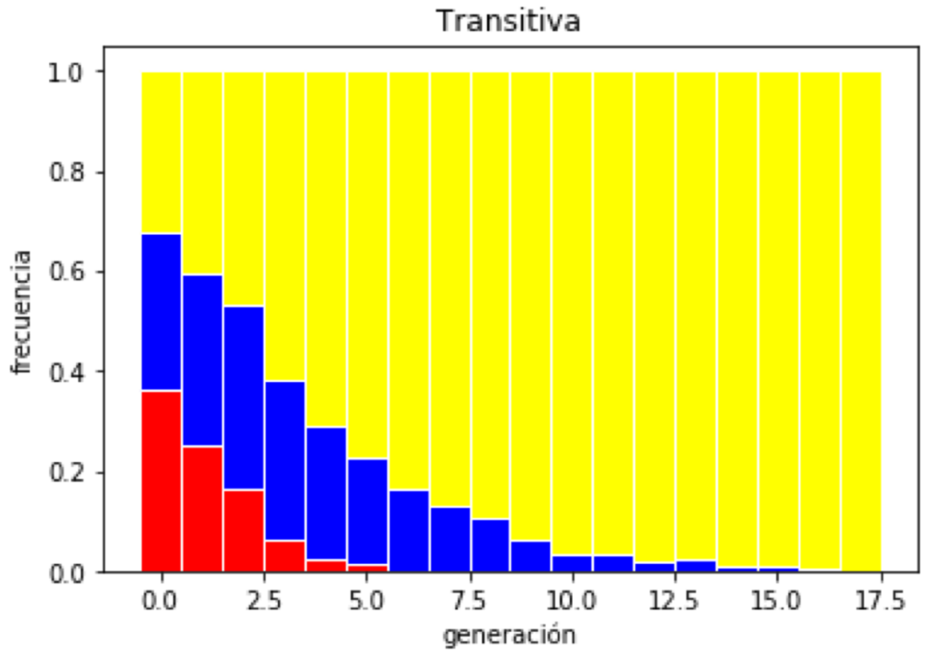


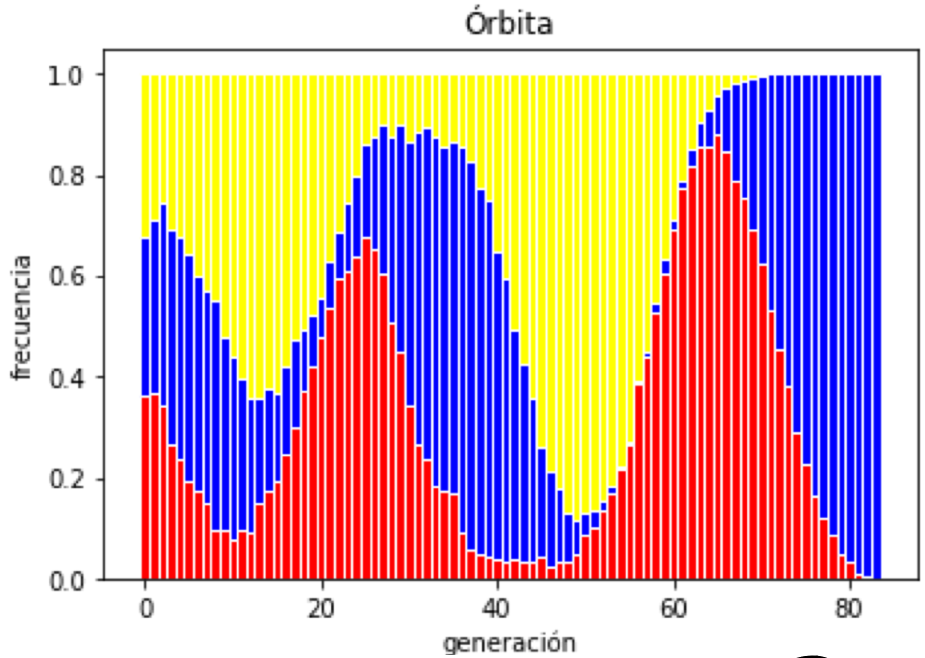






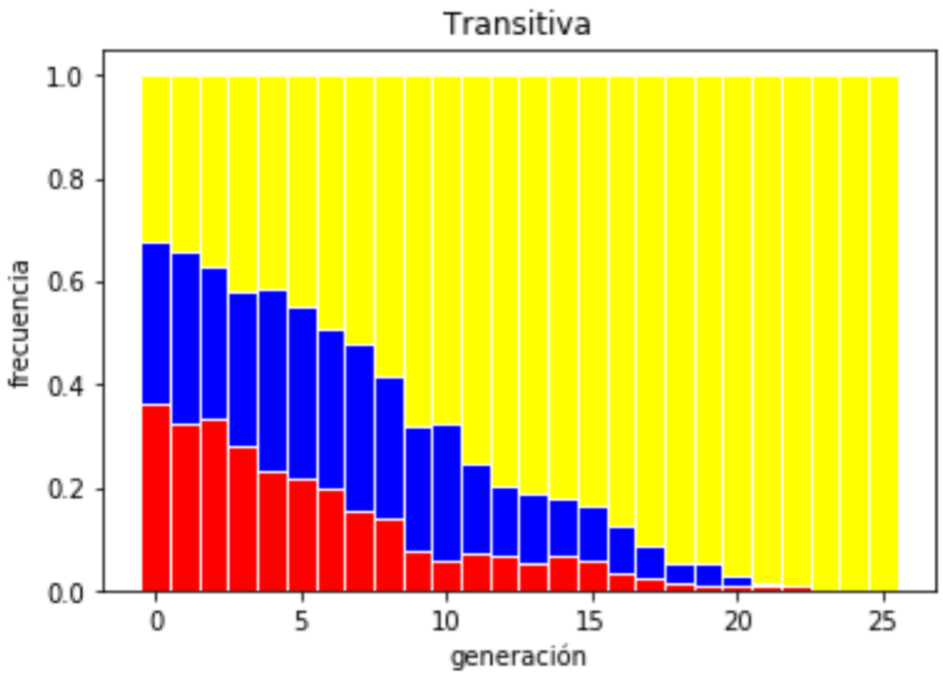
N= 300

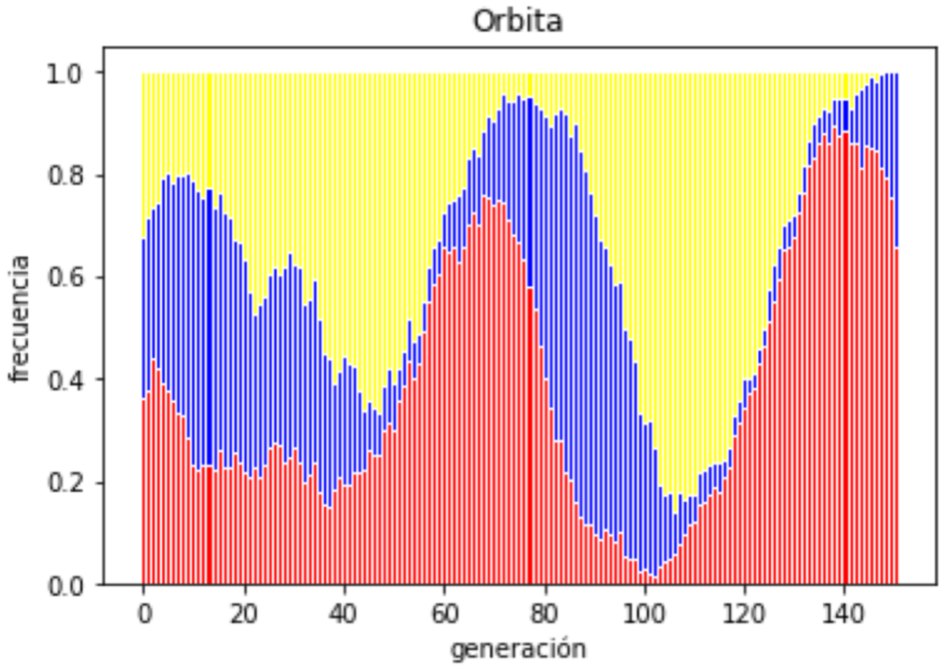




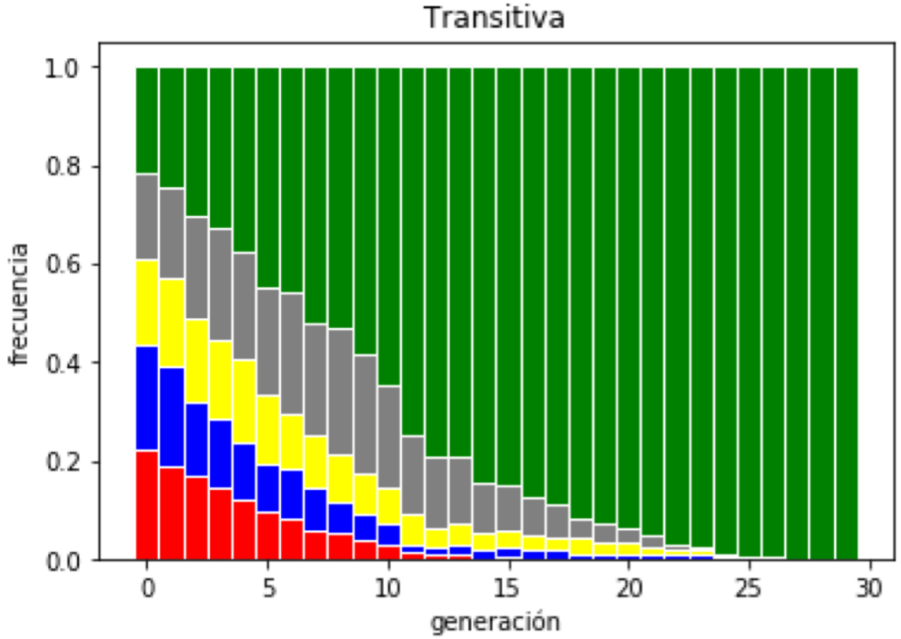
4 Sofia Rozanes

N= 300

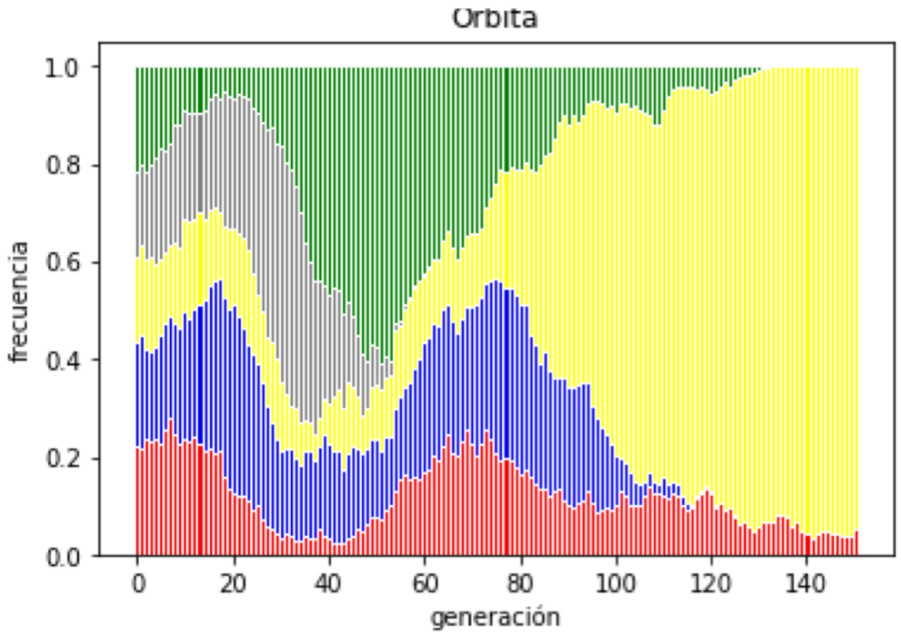

Transitiva
1.0
0.8
0.6
0.4
0.2
0.0
frecuencia
0
5
10
15
20
25
generación


Orbita
1.0
0.8
0.6
0.4
0.2
0.0
frecuencia
0
20
40
60
80
100
120
140
generación

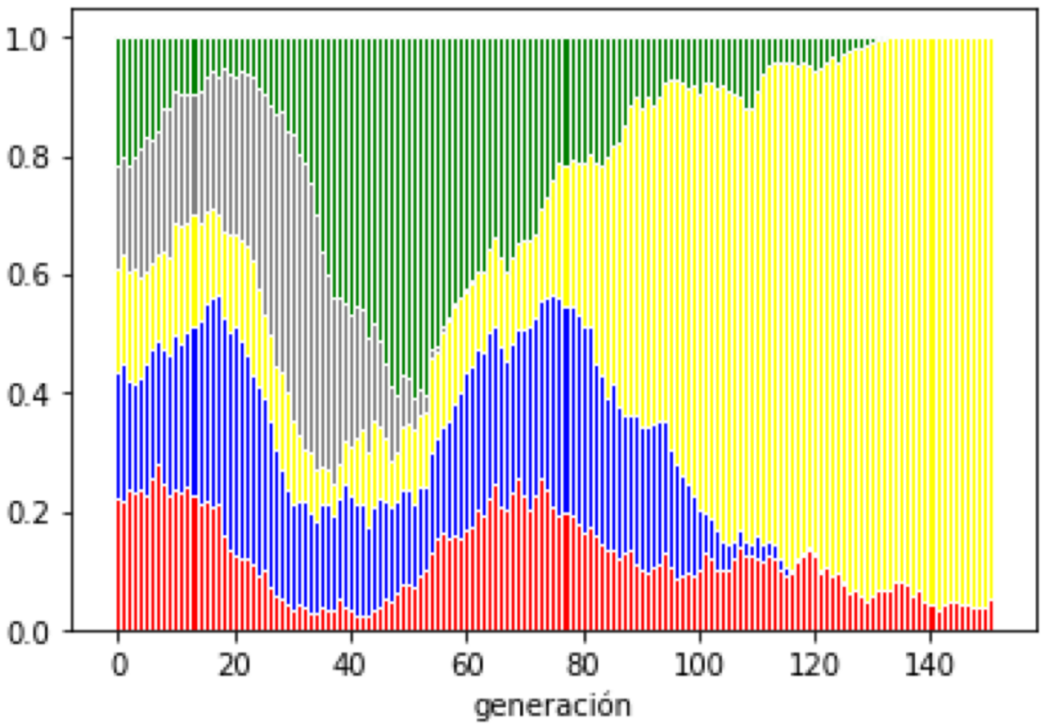
N= 600, k=5

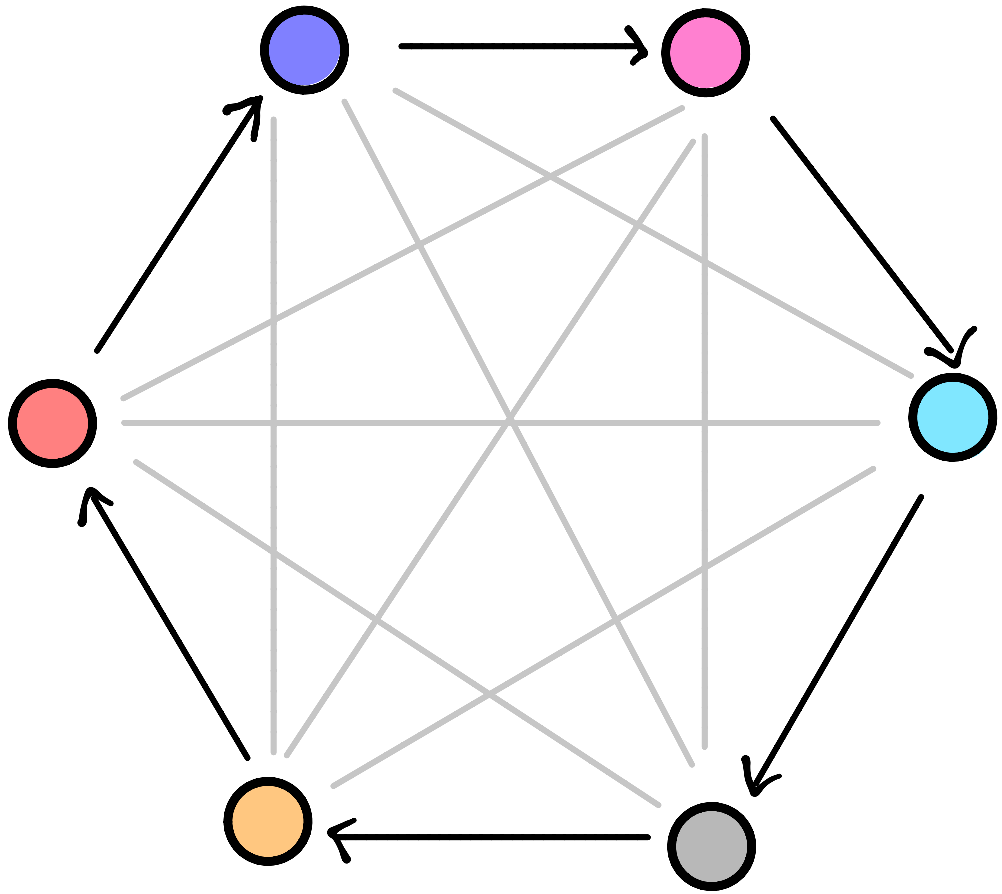


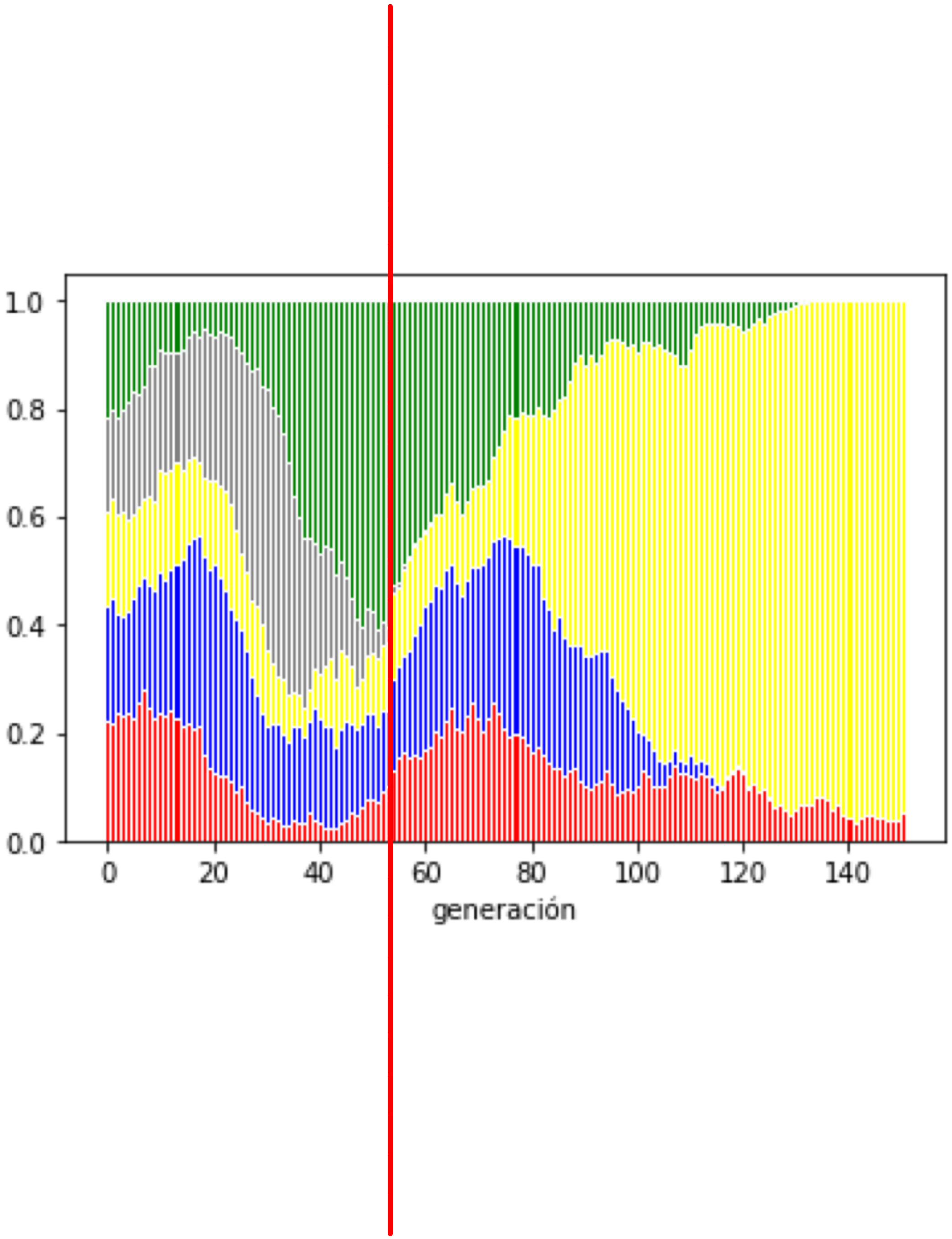
Transitiva
frecuencia
generación
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30



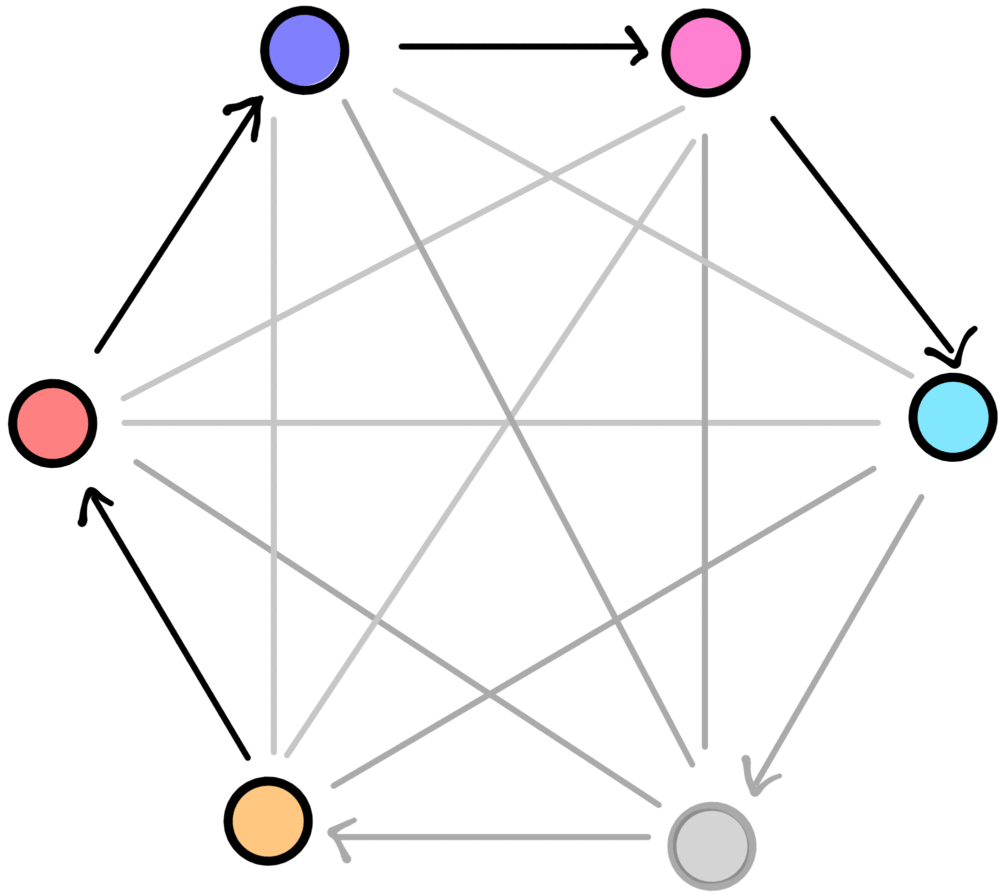
Orbita
frecuencia
generación
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100
120
140

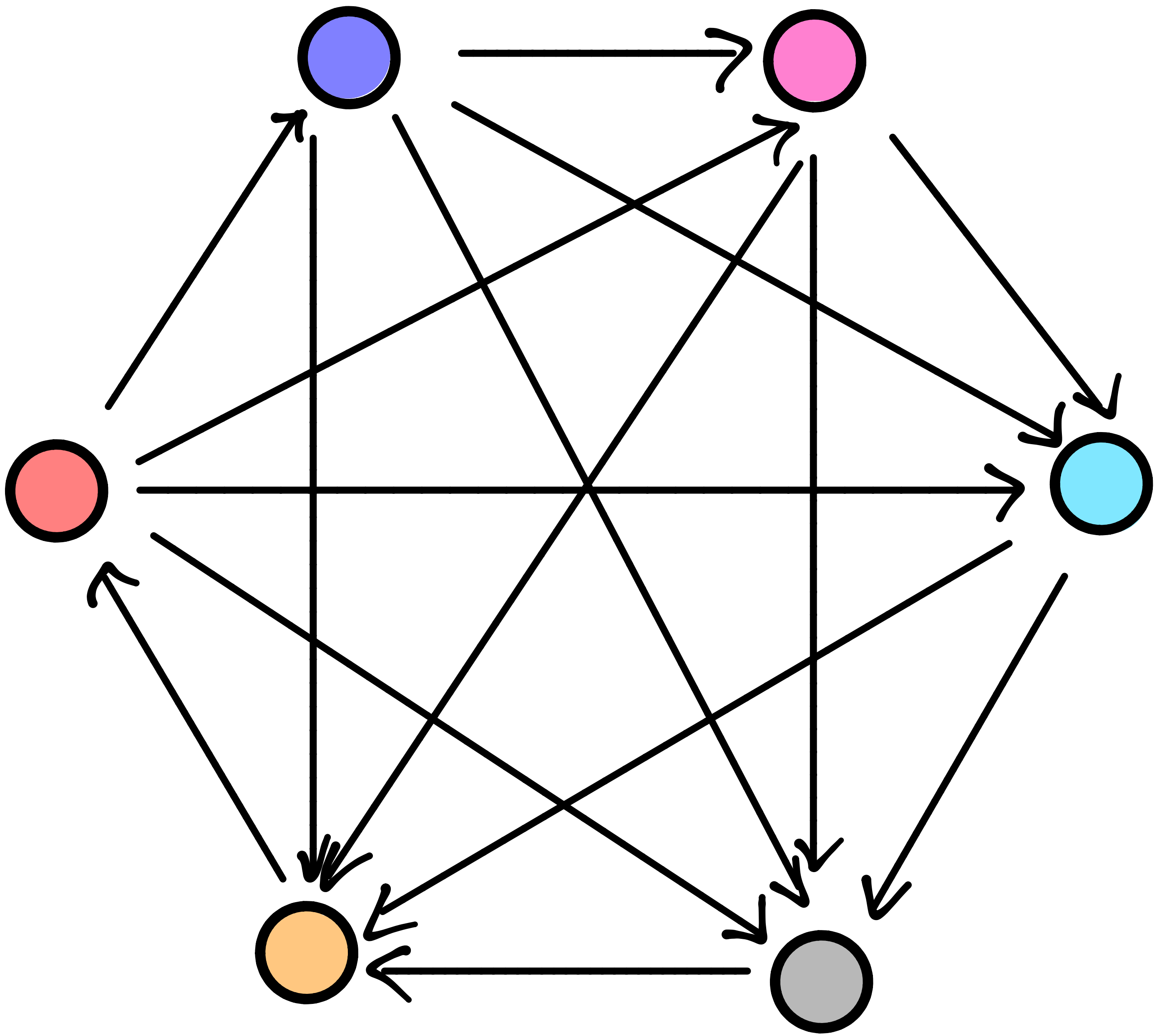

1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140
generación

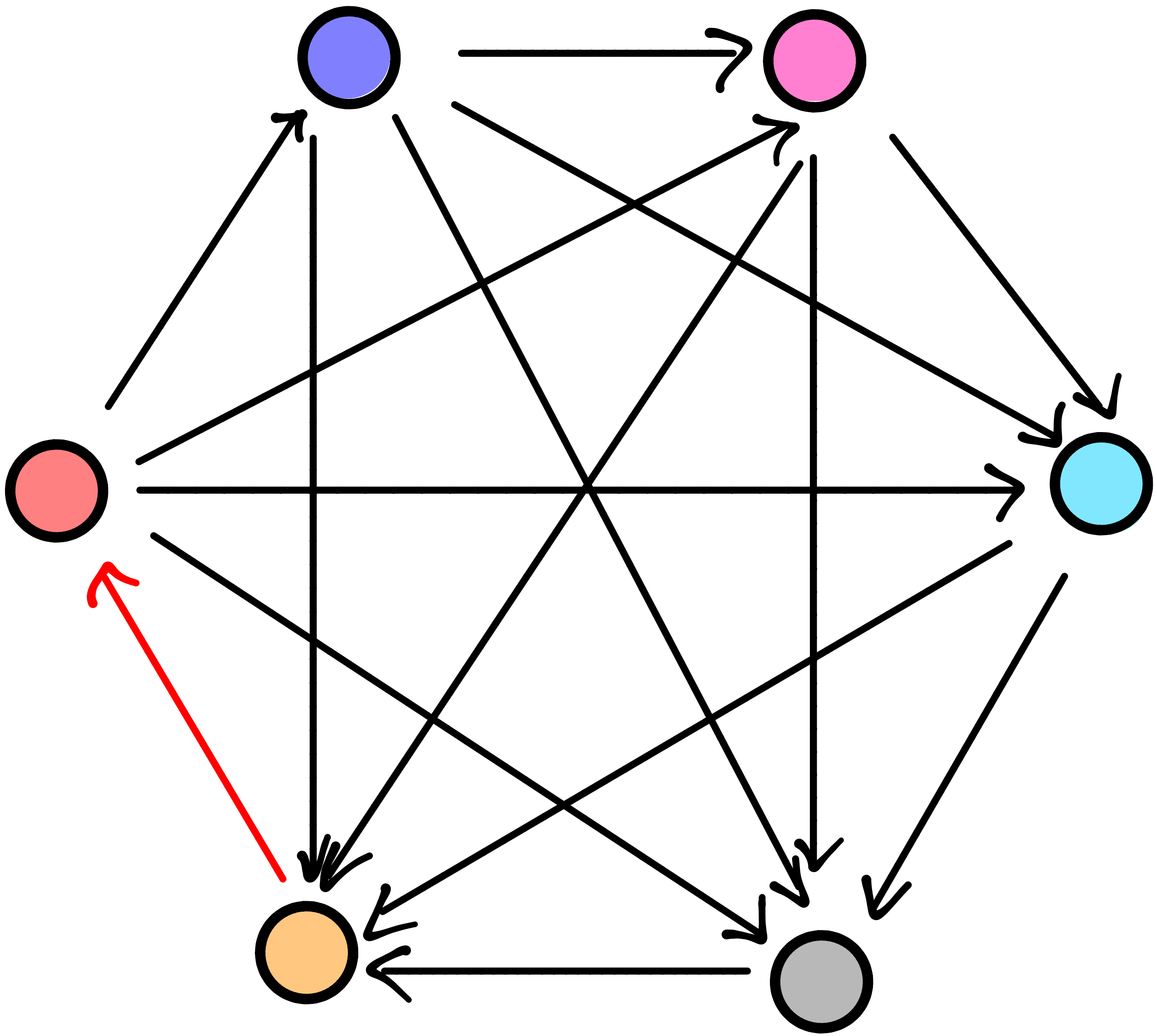


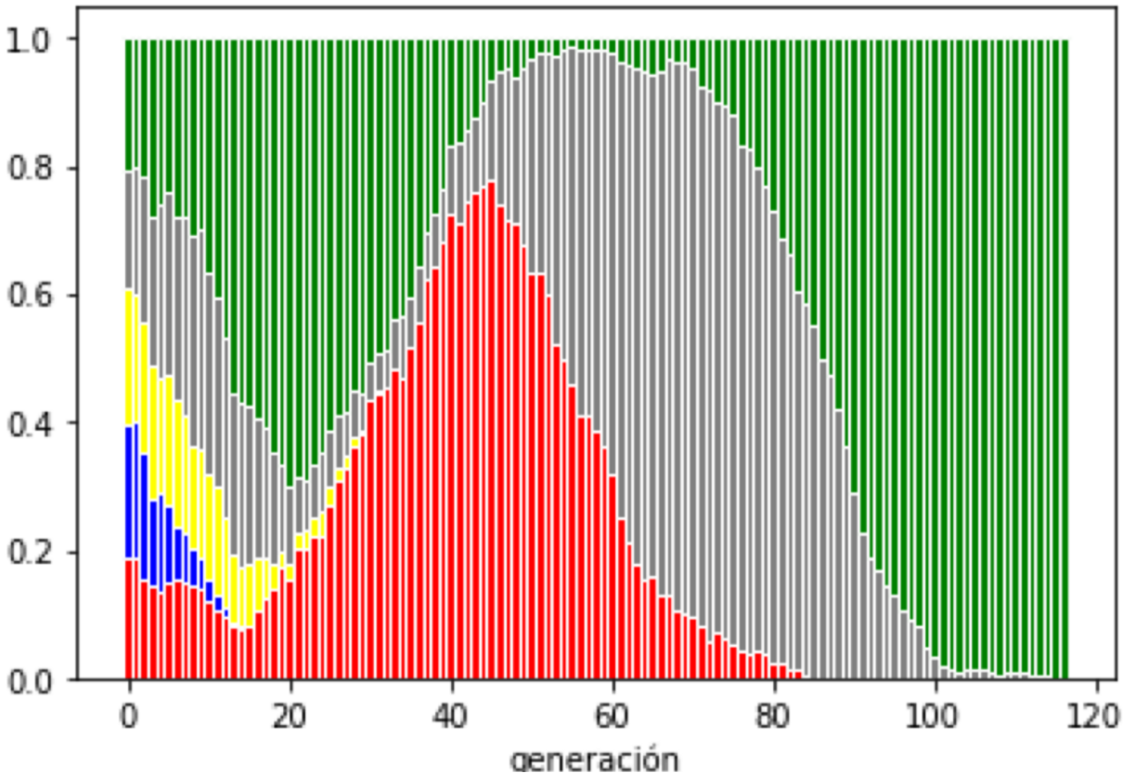


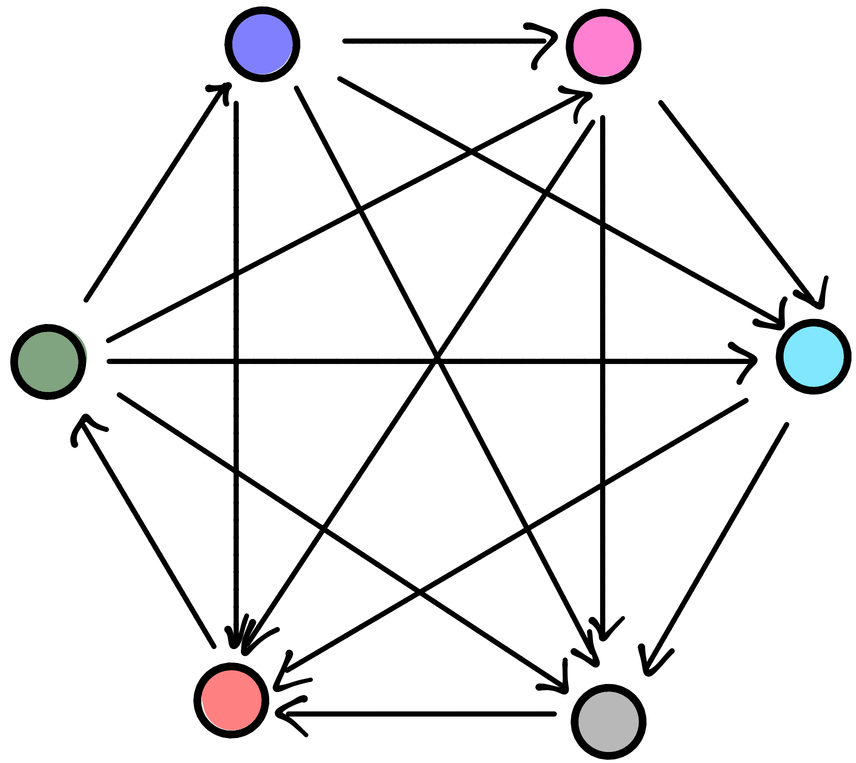
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140
generación

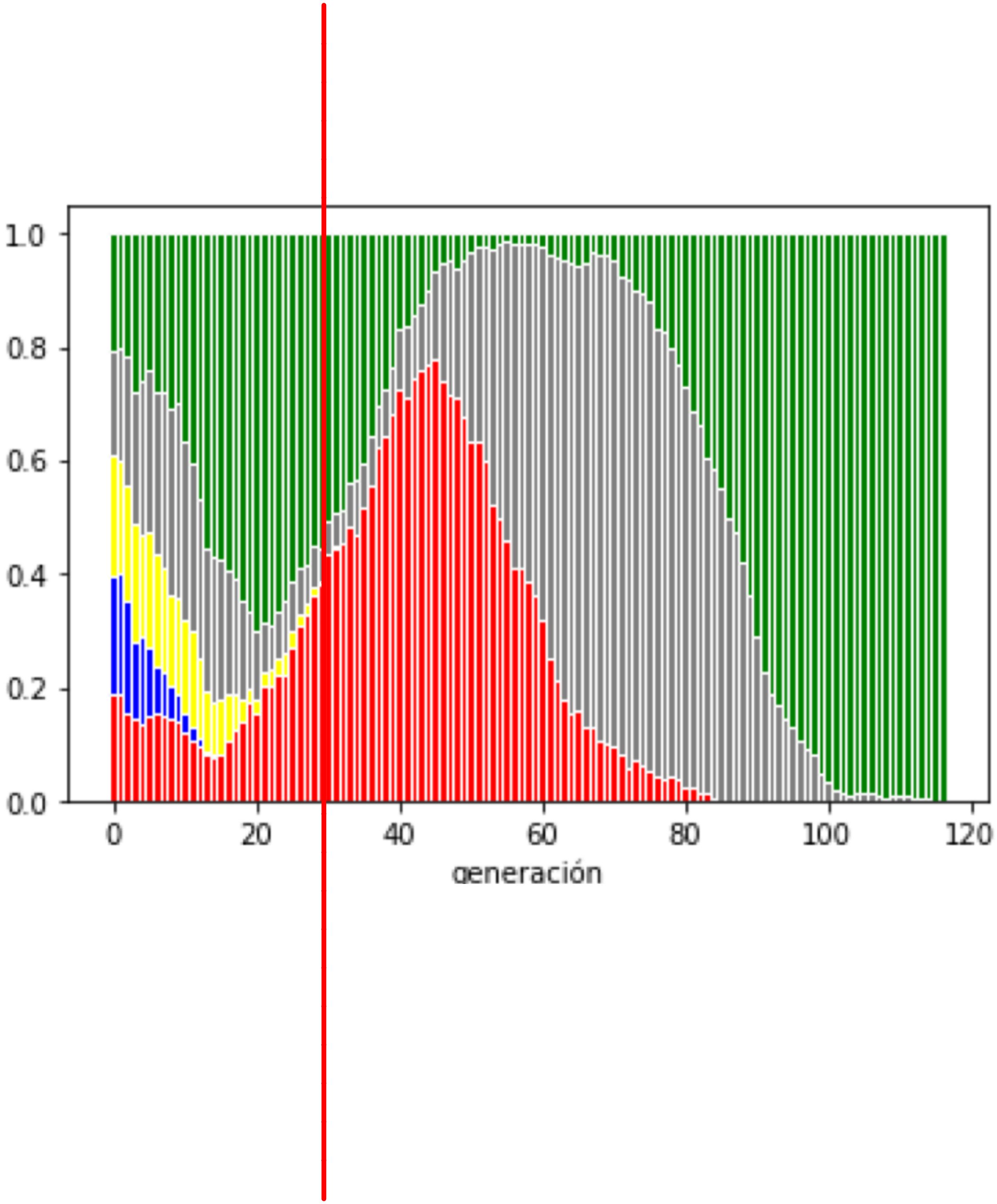




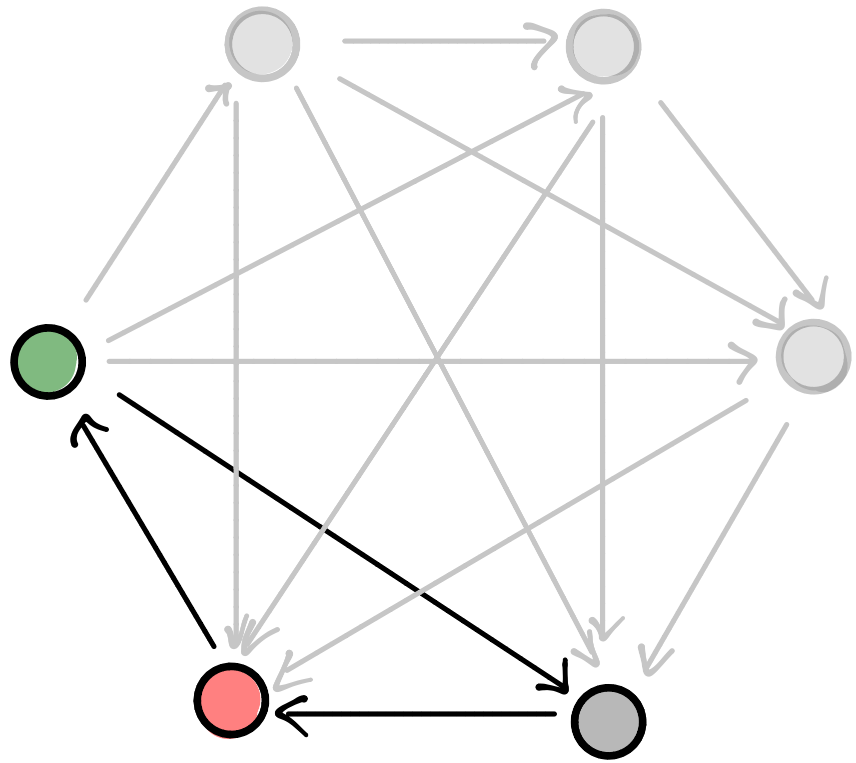



1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
generación

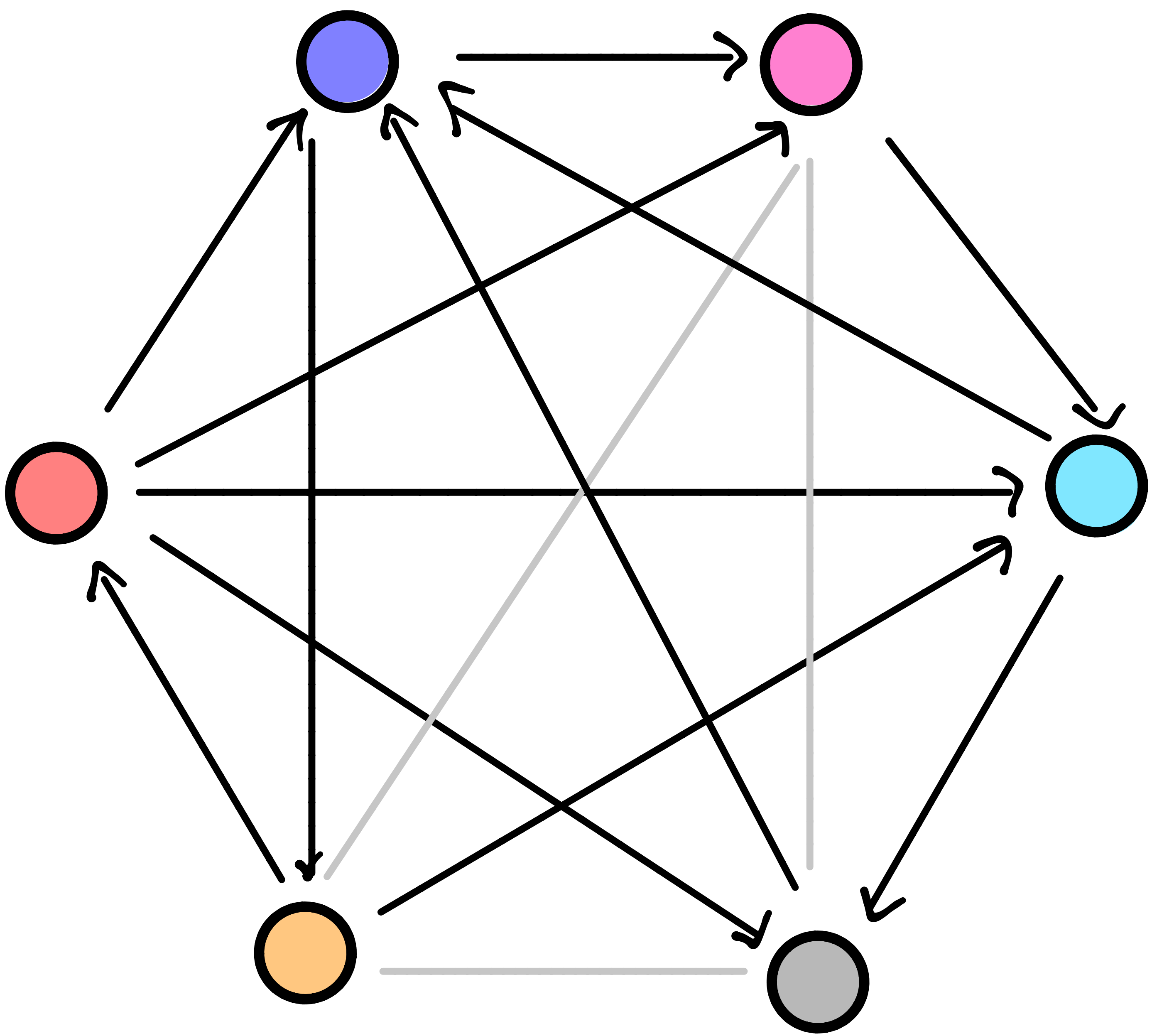




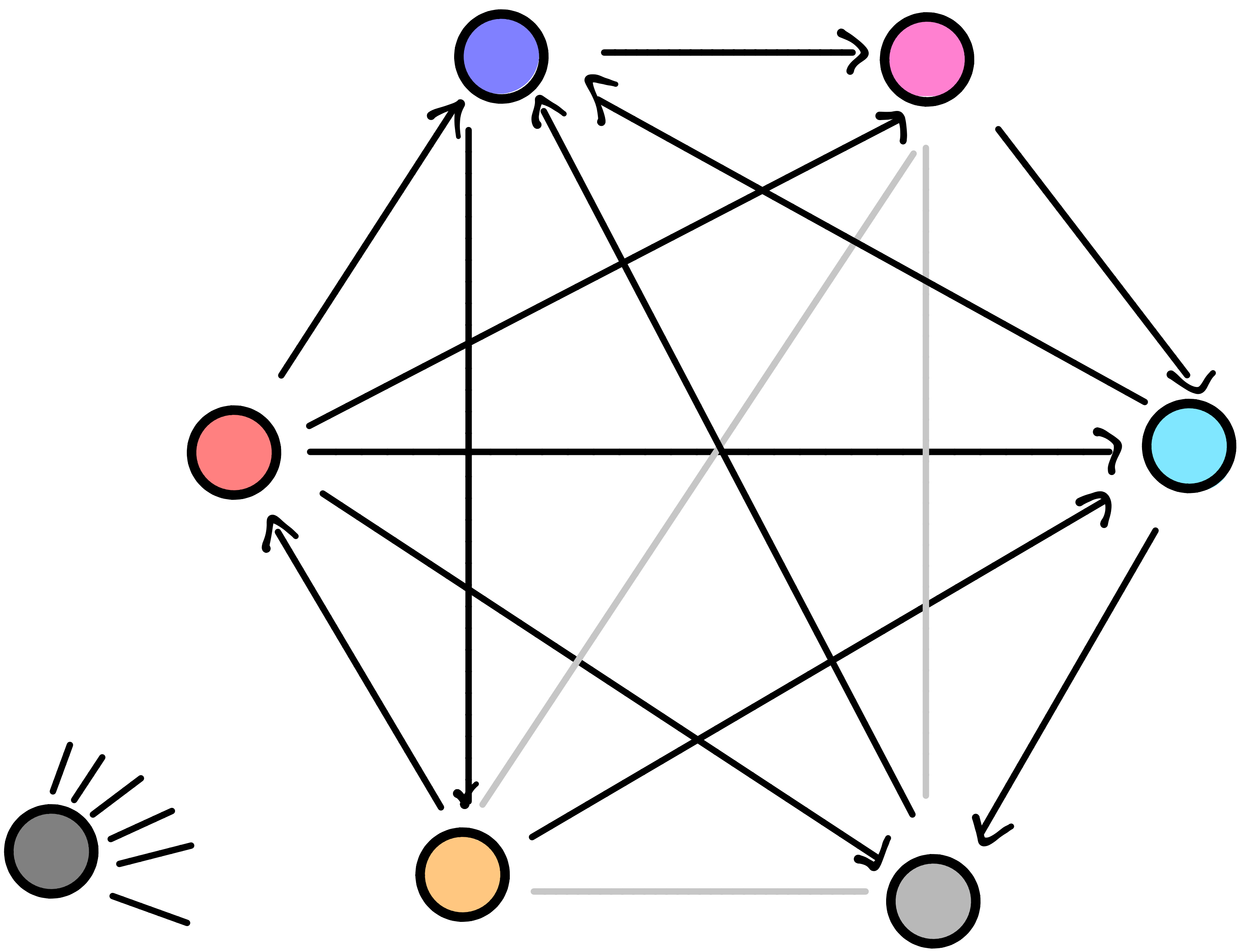
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
generación

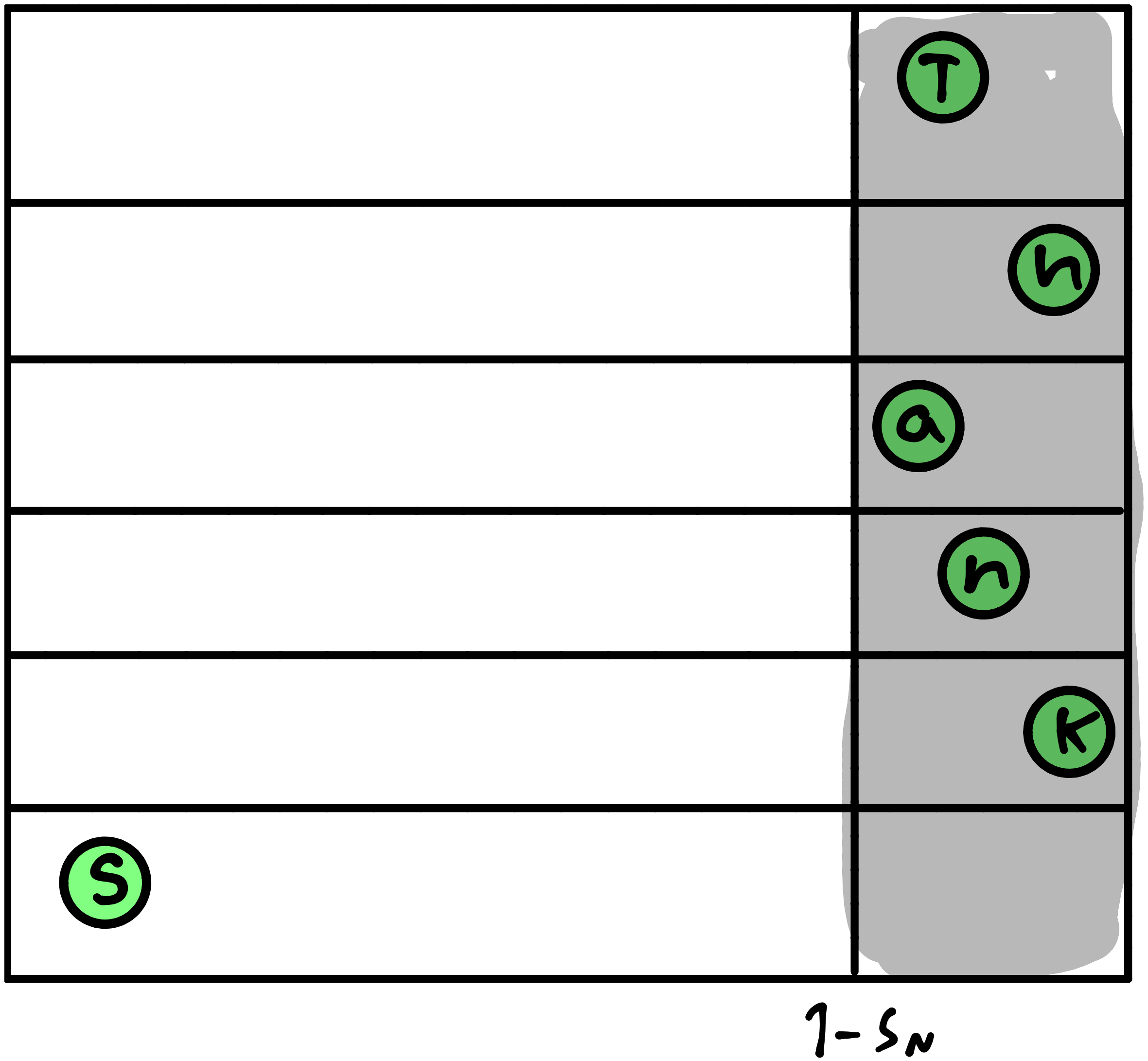


What is the most stable configuration?



Distribution of the number of types





T
h
a
n
k
S
1 - s_N

Wow!