5 Discrete coalescent

Given: Cannings model with constant population size N as in Section 3.

The gen. are now labeled backwards in time, i.e. $r = 0$ is the (current) gen. of the offspring, $r = 1$ the gen. of the parents, $r = 2$ the gen. of the grandparents and so on. Sample n ($\le N$) individuals from gen. 0 and consider their ancestors.

5.1 Def.
For $r \in \mathbb{N}_0$ define a random relation $\mathcal{R}_r$ on $[n] := \{1, \ldots, n\}$ via

$$(i,j) \in \mathcal{R}_r \;:\Longleftrightarrow\; \begin{array}{l}\textit{The individuals } i \textit{ and } j \textit{ have a common ancestor}\\ \textit{in gen. } r\textit{, i.e. } r \textit{ gen. backwards in the past.}\end{array}$$

Clearly, $\mathcal{R}_r$ is a random equivalence relation on $[n]$, i.e. $\mathcal{R}_r$ is a r.v. taking values in $\mathcal{E}_n$, the (finite) set of equivalence relations on $[n]$. Note that $\mathcal{R}_r = \mathcal{R}_r^{(N,n)}$ depends on N and n.

5.2 Theorem and Def. (Discrete Coalescent)
$\mathcal{R} := (\mathcal{R}_r)_{r\in\mathbb{N}_0}$ is a HMC with state space $\mathcal{E}_n$ and initial state $\mathcal{R}_0 \equiv \Delta_n := \{(i,i)\,|\,1 \le i \le n\}$ (diagonal relation). The transition probabilities $p_{\xi\eta} := P(\mathcal{R}_r = \eta\,|\,\mathcal{R}_{r-1} = \xi)$, $\xi, \eta \in \mathcal{E}_n$, are $p_{\xi\eta} = 0$ for $\xi \not\subseteq \eta$ and

$$\boxed{p_{\xi\eta} \;=\; \frac{(N)_a}{(N)_b}\mathrm{E}((\nu_1)_{b_1}\cdots(\nu_a)_{b_a})} \quad \textit{for } \xi \subseteq \eta,$$

where $(x)_0 := 1$ and $(x)_k := x(x-1)\cdots(x-k+1)$, $k \in \mathbb{N}$. Here $a = |\eta|$ and $b = |\xi|$ are the number of (equivalence) classes of η and ξ respectively, and $b_1, \ldots, b_a$ are the group sizes of merging classes of ξ ($\Rightarrow b_1 + \cdots + b_a = b$).

$\mathcal{R}$ is called (discrete) n-coalescent.

(Latin: coalescere = to coalesce, to union)

Proof.
Clearly, $\mathcal{R}_0 = \Delta_n$ and $p_{\xi\eta} = 0$ for $\xi \not\subseteq \eta$. Assume now that $\xi \subseteq \eta$. $\mathcal{R}_{r-1}, \nu^{(r)}$ independent. $\Rightarrow$

$$P(\mathcal{R}_r = \eta\,|\,\mathcal{R}_{r-1} = \xi) \;=\; \sum_k P(\mathcal{R}_r = \eta\,|\,\mathcal{R}_{r-1} = \xi, \nu^{(r)} = k)P(\nu^{(r)} = k),$$

where the sum extends over all $k = (k_1, \ldots, k_N)$ with $P(\nu^{(r)} = k) > 0$. In order to calculate $P(\mathcal{R}_r = \eta\,|\,\mathcal{R}_{r-1} = \xi, \nu^{(r)} = k)$ consider the following experiment: Assume that N colors are given. A box contains N balls (the children), more precisely, k_i balls of color i, $i \in [N]$. Sample b balls without replacement. The probability that among the b drawn balls there are exactly b_j balls of the same color, $j \in \{1, \dots, a\}$, is

$$P(\mathcal{R}_r = \eta \,|\, \mathcal{R}_{r-1} = \xi, \nu^{(r)} = k) \;=\; \sum_{\substack{i_1,\dots,i_a=1 \\ \text{all distinct}}}^{N} \frac{(k_{i_1})_{b_1} \cdots (k_{i_a})_{b_a}}{(N)_b}.$$

Note that $i_1, \dots, i_a$ are the numbers of the existing colors. $\Rightarrow$

$$\begin{aligned} P(\mathcal{R}_r = \eta \,|\, \mathcal{R}_{r-1} = \xi) \;&=\; \sum_k P(\nu^{(r)} = k) \sum_{\substack{i_1,\dots,i_a=1 \\ \text{all distinct}}}^{N} \frac{(k_{i_1})_{b_1} \cdots (k_{i_a})_{b_a}}{(N)_b} \\ &=\; \frac{1}{(N)_b} \sum_{\substack{i_1,\dots,i_a=1 \\ \text{all distinct}}}^{N} \mathrm{E}((\nu_{i_1})_{b_1} \cdots (\nu_{i_a})_{b_a}) \;=\; \frac{(N)_a}{(N)_b} \mathrm{E}((\nu_1)_{b_1} \cdots (\nu_a)_{b_a}), \end{aligned}$$

since $\nu_1, \dots, \nu_N$ are exchangeable. In particular, this probability does not depend on r and n. The calculation does not change if $\{\mathcal{R}_{r-1} = \xi\}$ is replaced by $\{\mathcal{R}_{r-1} = \xi, \mathcal{R}_{r-2} = \xi_{r-2}, \dots, \mathcal{R}_1 = \xi_1\}$. $\Rightarrow$ $\mathcal{R}$ is a HMC. $\square$

5.3 Example
<u>*WFM:*</u> $\boxed{p_{\xi\eta} = \dfrac{(N)_a}{N^b}}$ *for* $\xi \subseteq \eta$, $p_{\xi\eta} = 0$ *otherwise.*

Proof. $(X_1, \dots, X_N) \overset{d}{=} \mathrm{Mn}(N, p_1, \dots, p_N)$ *has joint factorial moments* $\mathrm{E}((X_1)_{b_1} \cdots (X_a)_{b_a}) = (N)_b p_1^{b_1} \cdots p_a^{b_a}$ *(exercise). Since* $p_1 = \cdots = p_N = 1/N$ *the result follows from Theorem 5.2.* $\square$

<u>*MM:*</u> $\boxed{p_{\xi\eta} = \dfrac{2}{N(N-1)}}$ *for* $\xi \subseteq \eta$ *with* $|\xi| = |\eta| + 1$,

$p_{\xi\xi} = 1 - \dfrac{b(b-1)}{N(N-1)}$, $p_{\xi\eta} = 0$ *otherwise.*

Proof. For $\xi \subseteq \eta$ *with* $b = a + 1$ *and, hence,* $b_1 = 2$, $b_2 = \cdots = b_a = 1$ *(the* b_i*'s w.l.o.g. ordered by size) by Theorem 5.2*

$$\begin{aligned} p_{\xi\eta} \;&=\; \frac{(N)_a}{(N)_b} \mathrm{E}((\nu_1)_2 \nu_2 \cdots \nu_a) \;=\; \frac{(N)_a}{(N)_b} 2P(\nu_1 = 2, \nu_2 = \cdots = \nu_a = 1) \\ &=\; \frac{(N)_a}{(N)_b} 2\frac{N-a}{N(N-1)} \;=\; \frac{2}{N(N-1)}. \end{aligned}$$

In the MM only two classes coalesce and the assertion follows. $\square$

5.4 Def. (**Transition Functions**)
For $a, b_1, \dots, b_a \in \mathbb{N}$ *with* $b := b_1 + \cdots + b_a \le N$ *define*

$$\boxed{\Phi_a(b_1, \dots, b_a) \;:=\; \frac{(N)_a}{(N)_b} \mathrm{E}((\nu_1)_{b_1} \cdots (\nu_a)_{b_a})}\,.$$

By Theorem 5.2, the transition matrix of $\mathcal{R}$ is uniquely determined by $\Phi_1, \ldots, \Phi_n$. We call $\Phi_1, \ldots, \Phi_n$ the transition functions of $\mathcal{R}$.

5.5 Lemma (Consistence of the Transition Functions)
For $a, b_1, \ldots, b_a \in \mathbb{N}$ with $b := b_1 + \cdots + b_a < N$,

$$\boxed{\Phi_{a+1}(b_1, \ldots, b_a, 1) = \Phi_a(b_1, \ldots, b_a) - \sum_{i=1}^{a} \Phi_a(b_1, \ldots, b_{i-1}, b_i + 1, b_{i+1}, \ldots, b_a)}$$

Proof.
Exchangeability of $\nu_1, \ldots, \nu_N$ and $\nu_1 + \cdots + \nu_N = N$. $\Rightarrow$

$$\begin{aligned}
&(N-a)\mathrm{E}\big((\nu_1)_{b_1} \cdots (\nu_a)_{b_a} \nu_{a+1}\big) \\
&= \mathrm{E}\big((\nu_1)_{b_1} \cdots (\nu_a)_{b_a} (\nu_{a+1} + \cdots + \nu_N)\big) \\
&= \mathrm{E}\big((\nu_1)_{b_1} \cdots (\nu_a)_{b_a} (N - \nu_1 - \cdots - \nu_a)\big) \\
&= \mathrm{E}\big((\nu_1)_{b_1} \cdots (\nu_a)_{b_a} (N - b - (\nu_1 - b_1) - \cdots - (\nu_a - b_a))\big) \\
&= (N-b)\mathrm{E}\big((\nu_1)_{b_1} \cdots (\nu_a)_{b_a}\big) - \sum_{i=1}^{a} \mathrm{E}\big((\nu_1)_{b_1} \cdots (\nu_i)_{b_i+1} \cdots (\nu_a)_{b_a}\big).
\end{aligned}$$

The assertion follows by multiplying both sides with $(N)_a/(N)_{b+1}$. □

5.6 Def. (Restriction)
For $m, n \in \mathbb{N}$ with $m \leq n$ let $\varrho_{nm} : \mathcal{E}_n \to \mathcal{E}_m$ denote the natural restriction defined via $\varrho_{nm}(\xi) := \{(i,j) \in \xi \,|\, 1 \leq i, j \leq m\}\ \forall\, \xi \in \mathcal{E}_n$.

5.7 Theorem (Natural Coupling)
Let $(\mathcal{R}_r)_{r \in \mathbb{N}_0}$ be a n-coalescent with transition functions $\Phi_1, \ldots, \Phi_n$ and let $m \in [n]$. Then the process $(\varrho_{nm} \circ \mathcal{R}_r)_{r \in \mathbb{N}_0}$ is a m-coalescent with transition functions $\Phi_1, \ldots, \Phi_m$.

Proof.
$\varrho_{nm} = \varrho_{m+1,m} \circ \cdots \circ \varrho_{n,n-1}$. $\Rightarrow$ W.l.o.g. $m = n-1$. Define $f := \varrho_{n,n-1}$. Let us verify that, for $\xi \in \mathcal{E}_n$ and $\tilde{\eta} \in \mathcal{E}_m$,

$$P(f \circ \mathcal{R}_r = \tilde{\eta} \,|\, \mathcal{R}_{r-1} = \xi) \;=\; \sum_{\eta \in f^{-1}(\tilde{\eta})} p_{\xi\eta} \qquad (+)$$

only depends via $\tilde{\xi} := f(\xi)$ on ξ. If $\tilde{\xi} \not\subseteq \tilde{\eta}$, then $\xi \not\subseteq \eta$, i.e. $p_{\xi\eta} = 0$ for all $\eta \in f^{-1}(\tilde{\eta})$ and $(+)$ is 0. Assume now that $\tilde{\xi} \subseteq \tilde{\eta}$. Let $C_1, \ldots, C_a$ denote the classes of $\tilde{\eta}$ and $C_{\alpha\beta}$, $\alpha \in \{1, \ldots, a\}$, $\beta \in \{1, \ldots, b_\alpha\}$, the classes of $\tilde{\xi}$ such that $C_\alpha = \bigcup_{\beta=1}^{b_\alpha} C_{\alpha\beta}$. Let $\eta_0 \in \mathcal{E}_n$ be the equivalence relation with classes

$C_1, \dots, C_a, \{n\}$, and for $i \in \{1, \dots, a\}$ let $\eta_i \in \mathcal{E}_n$ be the equivalence relation with classes $C_1, \dots, C_{i-1}, C_i \cup \{n\}, C_{i+1}, \dots, C_a$.

Assume first that $\{n\}$ is not a class of ξ. Then there exists $i \in \{1, \dots, a\}$ and $j \in \{1, \dots, b_i\}$ such that $C_{ij} \cup \{n\}$ is a class of ξ. The relation $\eta = \eta_i$ is then the only one satisfying the conditions $\xi \subseteq \eta$ and $f(\eta) = \tilde{\eta}$. Hence, (+) is in this case equal to $p_{\xi\eta_i} = \Phi_a(b_1, \dots, b_a)$.

Assume now that $\{n\}$ is a class of ξ. Then, exactly the relations $\eta \in \{\eta_0, \dots, \eta_a\}$ satisfy the conditions $\xi \subseteq \eta$ and $f(\eta) = \tilde{\eta}$. $\Rightarrow$

$$\begin{aligned}\sum_{\eta \in f^{-1}(\tilde{\eta})} p_{\xi\eta} &= \sum_{i=0}^{a} p_{\xi\eta_i} \\ &= \Phi_{a+1}(b_1, \dots, b_a, 1) + \sum_{i=1}^{a} \Phi_a(b_1, \dots, b_{i-1}, b_i + 1, b_{i+1}, \dots, b_a) \\ &= \Phi_a(b_1, \dots, b_a),\end{aligned}$$

by Lemma 5.5. $\Rightarrow$ (+) only depends via $\tilde{\xi}$ on ξ. We have

$$\{f \circ \mathcal{R}_{r-1} = \tilde{\xi}\} \ = \bigcup_{\xi \in f^{-1}(\tilde{\xi})} \{\mathcal{R}_{r-1} = \xi\}.$$

Thus, by Exercise 2.1.4, the transition probabilities $p_{\tilde{\xi}\tilde{\eta}} := P(f \circ \mathcal{R}_r = \tilde{\eta} \mid f \circ \mathcal{R}_{r-1} = \tilde{\xi})$ are equal to 0 for $\tilde{\xi} \not\subseteq \tilde{\eta}$ and for $\tilde{\xi} \subseteq \tilde{\eta}$ equal to

$$p_{\tilde{\xi}\tilde{\eta}} \ = \ \Phi_a(b_1, \dots, b_a).$$

$\mathcal{R}$ is a MC. $\Rightarrow$ The calculation does not change if $\{f \circ \mathcal{R}_{r-1} = \tilde{\xi}\}$ is replaced by $\{f \circ \mathcal{R}_{r-1} = \tilde{\xi}, f \circ \mathcal{R}_{r-2} = \tilde{\xi}_{r-2}, \dots, f \circ \mathcal{R}_0 = \tilde{\xi}_0\}$. $\Rightarrow$ $(f \circ \mathcal{R}_r)_{r \in \mathbb{N}_0}$ is a HMC. □

5.8 Lemma (Monotonicity of the Transition Functions)
For $1 \le l \le j \le N$ and $k_1, \dots, k_j, m_1, \dots, m_l \in \mathbb{N}$ with $k_1 \ge m_1, \dots, k_l \ge m_l$ and $k_1 + \dots + k_j \le N$,

$$\boxed{\Phi_j(k_1, \dots, k_j) \ \le \ \Phi_l(m_1, \dots, m_l)}\,, \qquad (*)$$

i.e. as more classes coalesce, as less likely this is.
In particular, $\Phi_j(k_1, \dots, k_j) \le \Phi_1(1) = \mathrm{E}(\nu_1) = 1$.

Proof.
Inductively on the difference $d := j - l \in \{0, \dots, N-1\}$. Lemma 5.5. $\Rightarrow$ $\Phi_l(m_1, \dots, m_l) \ge \Phi_l(m_1, \dots, m_{i-1}, m_i + 1, m_{i+1}, \dots, m_l)$ for $i \in \{1, \dots, l\}$.

Iteratively, $(*)$ follows for $j = l$, i.e. for $d = 0$. Again by Lemma 5.5 and $(*)$ for $d = 0$,

$$\Phi_l(m_1, \ldots, m_l) \overset{\text{Lem. 5.5}}{\geq} \Phi_{l+1}(m_1, \ldots, m_l, 1) \overset{(*)\text{ for } d=0}{\geq} \Phi_{l+1}(k_1, \ldots, k_{l+1}).$$

Thus, $(*)$ holds for $j = l + 1$, i.e. for $d = 1$. Applying $(*)$ for $d = 1$ exactly $(j - l)$-times yields

$$\begin{aligned} \Phi_l(m_1, \ldots, m_l) &\geq \Phi_{l+1}(k_1, \ldots, k_{l+1}) \\ &\geq \Phi_{l+2}(k_1, \ldots, k_{l+2}) \\ &\geq \cdots \geq \Phi_{j-1}(k_1, \ldots, k_{j-1}) \geq \Phi_j(k_1, \ldots, k_j). \qquad \square \end{aligned}$$

5.9 Def. (Coalescence Probability)
Let c_N denote the probability that two individuals, randomly sampled from some gen., have a common ancestor one gen. backwards, i.e.

$$\boxed{c_N := \Phi_1(2) = \frac{N}{(N)_2} \mathrm{E}((\nu_1)_2) = \frac{\mathrm{Var}(\nu_1)}{N-1}}.$$

c_N is called the coalescence probability

5.10 Example
WFM: $c_N = \dfrac{1}{N}$. MM: $c_N = \dfrac{2}{N(N-1)}$.

5.11 Example (Time to MRCA)
Consider 2 individuals of gen. 0. Let T_N denote the number of gen. which one has to go back to the past until you find the (most recent) common ancestor (MRCA) of these two individuals.

$\Rightarrow P(T_N = k) = c_N(1 - c_N)^{k-1}$, $k \in \mathbb{N}$, i.e. $T_N - 1 \overset{d}{=} G(c_N)$ (geometric distribution).

$\Rightarrow \mathrm{E}(T_N) = \dfrac{1}{c_N}$, $\mathrm{Var}(T_N) = \dfrac{1 - c_N}{c_N^2} = \dfrac{1}{c_N}\left(\dfrac{1}{c_N} - 1\right)$.

WFM: $\mathrm{E}(T_N) = N$, $\mathrm{Var}(T_N) = N(N-1)$.

MM: $\mathrm{E}(T_N) = \dfrac{N^2 - N}{2} \sim \dfrac{N^2}{2}$, $\mathrm{Var}(T_N) = \dfrac{N^2 - N}{2}\left(\dfrac{N^2 - N}{2} - 1\right) \sim \dfrac{N^4}{4}$.

Rem.
$N_e := 1/c_N = \mathrm{E}(T_N)$ is called the effective population size. Example 5.11 indicates that N_e is a measure for the speed of the evolution. Under mild conditions, the time-scaled ancestral process $(\mathcal{R}^{(N,n)}_{[t/c_N]})_{t \geq 0}$ converges in distribution as $N \to \infty$ to a continuous-time limiting process $(R^{(n)}_t)_{t \geq 0}$, which

is called an n-coalescent. The most prominent such limiting process is the Kingman n-coalescent (Kingman, 1982), allowing only for binary mergers of ancestral lineages. In general, these limiting processes allow for multiple collisions (Λ-coalescent) or even simultaneous multiple collisions (Ξ-coalescent) of ancestral lineages (Pitman 1999, Sagitov 1999, Schweinsberg 2000, M. and Sagitov 2001).