1 Notation

$\mathcal{T}$ Phylogeny with migration history (Structured phylogenetic tree)

d Number of distinct demes

Λ Migration matrix $\Lambda = [\Lambda_{ij}]_{i,j=1}^{d}$ such that $\Lambda_{ij} = \lambda_{ij} = \lambda_{i \to j}$ is the migration rate from deme i into deme j backwards in time. Note that self-migrations are forbidden, i.e. $\lambda_{ii} = 0 \ \forall i = 1, 2, \ldots, d$.

θ Effective population vector such that θ_i is the effective population in deme i

g Mean generation length for individuals in the overall population ($g > 0$)

M Total number of migration events

n Number of leaves (tips) in the tree

2 Structured Coalescent Likelihood

(Based on Ewing et al. (2004))

$$\xcancel{L_{n,d}(\mathcal{T}) = \prod_{i=1}^{d} \prod_{\substack{j=1 \\ j \neq i}}^{d} \frac{1}{\theta_i^{c_i}} \lambda_{ij}^{m_{ij}} \prod_{r=2}^{2n+M-1} \exp\left\{ - \left(\frac{k_{ir}(k_{ir}-1)}{2\theta_i} + k_{ir}\lambda_{ij} \right) (t_{r-1} - t_r) \right\}} \tag{1}$$

$$L_{n,d}(\mathcal{T}) = \prod_{i=1}^{d} \frac{1}{\theta_i^{c_i}} \exp\left\{ - \sum_{r=1}^{2n+M-2} \frac{1}{\theta_i} \binom{k_{ir}}{2} \delta t_r \right\} \prod_{\substack{j=1 \\ j \neq i}}^{d} \lambda_{ij}^{m_{ij}} \exp\left\{ - \sum_{r=1}^{2n+M-2} k_{ir}\lambda_{ij}\delta t_r \right\} \tag{2}$$

3 Structured Coalescent Simulation Testing

Realisations of the structured coalescent process can be constructed recursively for a given set of tip data (containing a sample time and a sample deme for each leaf).

3.1 Likelihood Testing

The likelihood of a realisation of a structured coalescent process can be computed directly using (2), or alternatively computed recursively throughout the generation of the simulation. One simple verification for the simulations is then to check that the two computed likelihoods agree for the tree.

3.2 MLE for migration rates and effective population sizes

An alternative method to verify the accuracy of the simulation method is to compare the maximum likelihood estimates of the migration rates matrix and effective population sizes to the simulation parameters. Maximum likelihood estimates can be obtained in the special case

$$
\begin{aligned}
l_{n,d}(\mathcal{T}) &= \sum_{i=1}^{d}\left[-c_i \log\theta_i - \sum_{r=1}^{2n+M-2}\frac{1}{\theta_i}\binom{k_{ir}}{2}\delta t_r + \sum_{\substack{j=1\\ j\neq i}}^{d}\left(m_{ij}\log\lambda_{ij} - \sum_{r=1}^{2n+M-2} k_{ir}\lambda_{ij}\delta t_r\right)\right] \\
&=
\end{aligned}
$$

4 MCMC

Fix a structured coalescent phylogeny and the evolutionary parameters (migration matrix, effective populations). Construct a MCMC algorithm based on a subset of the proposal moves attempted by Ewing et al. (2004),

1. Migration Birth/Death Move
2. Migration Pair Birth/Death Move
3. Coalescent Node Split/Merge Move

which obtain irreducibility over the space of migration histories on a fixed phylogenetic tree with fixed evolutionary parameters.

The acceptance probabilities for a reversible jump MCMC scheme take the form

$$\alpha(\mathcal{T}'|\mathcal{T}) = \min\left(1, \frac{L(\mathcal{T}')\pi(\mathcal{T}')Q(\mathcal{T}|\mathcal{T}')}{L(\mathcal{T})\pi(\mathcal{T})Q(\mathcal{T}'|\mathcal{T})}||\mathcal{J}||\right) \tag{3}$$

where L denotes the joint likelihood of the phylogenetic tree and migration history, π denotes the prior distribution on the phylogeny, $Q(\mathcal{T}'|\mathcal{T})$ denotes the transition probability of obtaining $\mathcal{T}'$ from $\mathcal{T}$, and $\mathcal{J}$ denotes the Jacobian of the transformation.

In the three proposal methods listed above, the transformations act on the migration history by adding or removing migration events, resulting in the Jacobian term $||\mathcal{J}||$ remaining equal to 1 throughout.

4.1 Migration Birth/Death Proposal

The migration birth/death proposal aims to add or remove a single migration node from the migration history, reassigning demes as necessary to maintain a consistent migration history. With probability $\frac{1}{2}$ a birth proposal is attempted, and otherwise a death proposal is attempted.

4.1.1 Ewing et al. (2004) Version

If a migration birth event is attempted, a new migration node $\hat{\imath}$ is proposed to be born at a location uniformly distributed along edge $\langle i, ip \rangle$. Otherwise, if a migration death event is attempted, set $\hat{\imath}$ to be the parent of ip and remove node ip from the edge $\langle i, \hat{\imath} \rangle$

For any pair of vertices $i, j \in \mathcal{V}$, let $\mathcal{T}_{\langle i,j \rangle}$ denote the maximal subtree of $\mathcal{T}$ containing edge $\langle i, j \rangle$ such that the terminal nodes of $\mathcal{T}_{\langle i,j \rangle}$ are either migration or leaf nodes.

Then the proposal is completed by selecting a deme $\hat{d}$ over the subtree $\mathcal{T}_{\langle i,\hat{\imath} \rangle}$ such that $\hat{d}$ is not equal to the deme of any edge connected to $\mathcal{T}_{\langle i,\hat{\imath} \rangle}$ or the deme of edge $\langle i, \hat{\imath} \rangle$ itself. If no deme $\hat{d}$ exists then the proposal is rejected with probability 1. Note that a migration birth/death move is not possible when there are only 2 demes as there is no way to choose a new deme consistent with the surrounding demes.

The proposal ratio for the migration birth move is then given by

$$\begin{aligned}
Q(\mathcal{T}'|\mathcal{T}) &= \mathbb{P}\left[\begin{smallmatrix}\text{Selected node}\\ \text{from } \mathcal{T}\end{smallmatrix}\right] \cdot \mathbb{P}\left[\begin{smallmatrix}\text{Selected}\\ \text{deme}\end{smallmatrix}\right] \cdot \mathbb{P}\left[\begin{smallmatrix}\text{New node}\\ \text{time}\end{smallmatrix}\right] \\
&= \frac{1}{n-1+M} \cdot \frac{1}{c_b} \frac{1}{t_i - t_{ip}} \\
&= \frac{1}{(n+M-1)c_b(t_i - t_{ip})}; \\
Q(\mathcal{T}|\mathcal{T}') &= \mathbb{P}\left[\begin{smallmatrix}\text{Selected migration}\\ \text{node from } \mathcal{T}'\end{smallmatrix}\right] \cdot \mathbb{P}\left[\begin{smallmatrix}\text{Selected}\\ \text{deme}\end{smallmatrix}\right] \\
&= \frac{1}{M+1} \cdot \frac{1}{c_d}; \\
\frac{Q(\mathcal{T}|\mathcal{T}')}{Q(\mathcal{T}'|\mathcal{T})} &= \frac{c_b(n+M-1)(t_i - t_{ip})}{c_d(M+1)}
\end{aligned}$$

where c_b and c_d are the number of demes which could be proposed for the birth and death proposals respectively.

Similarly, the proposal ratio for the migration death move is given by

$$\frac{Q(\mathcal{T}|\mathcal{T}')}{Q(\mathcal{T}'|\mathcal{T})} = \frac{c_d M}{c_b(n+M-2)(t_i - t_{\hat{\imath}})}.$$

4.1.2 Updated Version 1 (Uniform location on tree for birth, deterministic deme update for death)

If a migration birth proposal is attempted, a location is selected uniformly on the tree for a new migration node $\hat{\imath}$ to be born. Otherwise, a migration death proposal is attempted similarly to the migration death proposal of Ewing et al. The key difference for the migration death proposal is that the updated deme is selected deterministically to be the same as the deme above the selected migration event.

This results in the modified proposal ratios

$$\frac{Q_B(\mathcal{T}|\mathcal{T}')}{Q_B(\mathcal{T}'|\mathcal{T})} = \frac{c_b \mathcal{L}}{M+1}$$

for the birth proposal, and

$$\frac{Q_D(\mathcal{T}|\mathcal{T}')}{Q_D(\mathcal{T}'|\mathcal{T})} = \frac{M}{c_b \mathcal{L}}$$

for the death proposal.

4.1.3 Updated Version 2 (Simplified deme selection on birth)

If a migration birth proposal is attempted, continue selecting the location uniformly on the tree, but now select the proposal deme for the newly modified edge from all other demes except the current deme of the edge. This results in always selecting from a set of $d-1$ demes, and may also propose inconsistent deme labellings. To avoid this problem, now reject proposals via the likelihood so that inconsistent configurations are assigned likelihood 0.

This new change results in the modified proposal ratios

$$\frac{Q_B(\mathcal{T}|\mathcal{T}')}{Q_B(\mathcal{T}'|\mathcal{T})} = \frac{(d-1)\mathcal{L}}{M+1}$$

for the birth proposal, and

$$\frac{Q_D(\mathcal{T}|\mathcal{T}')}{Q_D(\mathcal{T}'|\mathcal{T})} = \frac{M}{(d-1)\mathcal{L}}$$

for the death proposal.

4.2 Migration Pair Birth/Death Proposal

The migration pair birth/death proposal aims to add or remove a pair of migration nodes from the migration history, reassigning demes as necessary to require a consistent migration history. With probability $\frac{1}{2}$ a pair birth proposal is attempted, and otherwise a pair death proposal is attempted.

4.2.1 Ewing et al. (2004) Version

Begin by selecting an edge $e = \langle i, j \rangle \in \mathcal{E}$ is selected uniformly at random. If a migration pair birth event is attempted, two new nodes $\hat{\imath}_1, \hat{\imath}_2$ are placed at locations uniformly along e, hence splitting edge e into 3 new edges. The demes of these 3 edges are then updated so that the outer edges (ending on either i or j) have the same deme as edge e, and the central edge $\langle \hat{\imath}_1, \hat{\imath}_2 \rangle$ has a deme not equal to that of edge e.

Otherwise, if a migration pair death event is attempted, the nodes i and j are removed from the tree subject to both i and j being migration nodes, and all of the edges adjacent to e arising from the same deme. If either of these conditions fail, the proposal is assigned a proposal ratio of 0.

The proposal ratio for the migration pair birth proposal is then given by

$$\begin{aligned}
Q_B(\mathcal{T}'|\mathcal{T}) &= \mathbb{P}[{}^{\text{Selected}}_{\text{edge}}] \cdot \mathbb{P}[{}^{\text{Selected}}_{\text{deme}}] \cdot \mathbb{P}[{}^{\text{New node}}_{\text{times}}] \\
&= \frac{1}{2n+M-2} \cdot \frac{1}{d-1} \cdot \frac{2}{\delta t^2} \\
&= \frac{2}{(2n+M-2)(d-1)\delta t^2}, \\
Q_B(\mathcal{T}|\mathcal{T}') &= \mathbb{P}[{}^{\text{Selected}}_{\text{edge}}] \\
&= \frac{1}{2n+M}, \\
\frac{Q_B(\mathcal{T}|\mathcal{T}')}{Q_B(\mathcal{T}'|\mathcal{T})} &= \frac{(2n+M-2)(d-1)\delta t^2}{2(2n+M)}
\end{aligned}$$

where δt denotes the time between the targetted pair of migration nodes.

Similarly, the proposal ratio for the migration death proposal is given by

$$\frac{Q_D(\mathcal{T}|\mathcal{T}')}{Q_D(\mathcal{T}'|\mathcal{T})} = \frac{2(2n+M-2)}{(2n+M-4)(d-1)\delta t^2}.$$

4.2.2 Updated Version 1

Begin by selecting a branch of the **genealogy**, with probability proportional to its length (sufficient to sample a point uniformly along the tree and take the edge of the genealogy where it lands). If a migration pair birth is proposal is attempted, sample two locations uniformly along the selected branch and a deme selected uniformly from the d possible demes. The proposal will fail if

(i) There exists a migration event between the two proposed times

(ii) The deme selected leads to an incompatible deme labelling (rejected via the likelihood not the proposal ratio).

If a migration death proposal is attempted, let M_b denote the number of migration nodes on the selected branch. If $M_b \geq 2$, select a pair of consecutive migration nodes along the branch (if $M_b < 2$, reject the proposal). Then the proposal is completed by deleting the two selected migration nodes.The proposal will fail if

(i) There are fewer than 2 migration nodes along the branch ($M_b < 2$)

(ii) The exterior demes lead to an incompatible deme labelling (rejected via the likelihood not the proposal ratio).

This change to the proposal mechanism results in the new proposal ratios

$$
\begin{aligned}
Q_B(\mathcal{T}'|\mathcal{T}) &= \mathbb{P}\left[\substack{\text{Selected}\\\text{branch}}\right] \cdot \mathbb{P}\left[\substack{\text{New node}\\\text{times}}\right] \cdot \mathbb{P}\left[\substack{\text{Selected}\\\text{deme}}\right] \\
&= \frac{\mathcal{L}_b}{\mathcal{L}} \cdot \frac{2}{\mathcal{L}_b^2} \cdot \frac{1}{d} \\
&= \frac{2}{\mathcal{L}\mathcal{L}_b d}, \\
Q_B(\mathcal{T}|\mathcal{T}') &= \mathbb{P}\left[\substack{\text{Selected}\\\text{branch}}\right] \cdot \mathbb{P}\left[\substack{\text{Selected pair}\\\text{of nodes}}\right] \\
&= \frac{\mathcal{L}_b}{\mathcal{L}} \cdot \frac{1}{M_b - 1} \\
&= \frac{\mathcal{L}_b}{\mathcal{L}(M_b - 1)}, \\
\frac{Q_B(\mathcal{T}|\mathcal{T}')}{Q_B(\mathcal{T}'|\mathcal{T})} &= \frac{\mathcal{L}_b^2 d}{2(M_b - 1)}
\end{aligned}
$$

4.2.3 Updated Version 2 (Jere's version?)

Pair birth:

(i) Sample branch from the genealogy proportional to its length

(ii) Sample two locations uniformly along the branch

(iii) Sample 2 demes to be pulled inwards from the proposed locations (can sample deme uniformly from set of demes not equal to current deme at that location)

(iv) Pull proposed demes inwards from proposed locations towards centre of branch until hitting a migration node (or the other sampled location)

$$
\begin{aligned}
Q_B(\mathcal{T}'|\mathcal{T}) &= \mathbb{P}\left[\substack{\text{Selected}\\\text{branch}}\right] \cdot \mathbb{P}\left[\substack{\text{New node}\\\text{times}}\right] \cdot \mathbb{P}\left[\substack{\text{Selected}\\\text{demes}}\right] \\
&= \frac{\mathcal{L}_b}{\mathcal{L}} \cdot \frac{2}{\mathcal{L}_b^2} \cdot \frac{1}{(d-1)^2}.
\end{aligned}
$$

$$
\begin{aligned}
\frac{Q_D(\mathcal{T}|\mathcal{T}')}{Q_B(\mathcal{T}'|\mathcal{T})} &= \frac{\frac{\mathcal{L}_b}{\mathcal{L}} \cdot \frac{2}{(M_b+1)(M_b+2)} \cdot \frac{1}{(d-1)^2}}{\frac{\mathcal{L}_b}{\mathcal{L}} \cdot \frac{2}{\mathcal{L}_b^2} \cdot \frac{1}{(d-1)^2}} \\
&= \frac{\mathcal{L}_b^2}{(M_b+1)(M_b+2)}
\end{aligned}
$$

Pair death:

(i) Sample branch from the genealogy proportional to its length

(ii) If there are at least 2 migration events on the branch, sample two migration nodes

(iii) Sample 2 demes to be pulled inwards from the proposed locations (if the locations are consecutive, special case recovers Ewing pair death move if both proposed demes are the same)

(iv) Pull proposed demes inwards from proposed locations towards centre of branch until hitting a migration node (including the other sampled migration node)

$$\begin{aligned} Q_D(\mathcal{T}'|\mathcal{T}) &= \mathbb{P}\left[\substack{\text{Selected}\\\text{branch}}\right] \cdot \mathbb{P}\left[\substack{\text{Selected}\\\text{nodes}}\right] \cdot \mathbb{P}\left[\substack{\text{Selected}\\\text{demes}}\right] \\ &= \frac{\mathcal{L}_b}{\mathcal{L}} \cdot \frac{2}{M_b(M_b-1)} \cdot \frac{1}{(d-1)^2} \end{aligned}$$

$$\begin{aligned} \frac{Q_B(\mathcal{T}|\mathcal{T}')}{Q_D(\mathcal{T}'|\mathcal{T})} &= \frac{\frac{\mathcal{L}_b}{\mathcal{L}} \cdot \frac{2}{\mathcal{L}_b^2} \cdot \frac{1}{(d-1)^2}}{\frac{\mathcal{L}_b}{\mathcal{L}} \cdot \frac{2}{M_b(M_b-1)} \cdot \frac{1}{(d-1)^2}} \\ &= \frac{M_b(M_b-1)}{\mathcal{L}_b^2}. \end{aligned}$$

4.3 Coalescent Node Split/Merge Proposal

4.3.1 Ewing et al. (2004) Version

NOTE THAT THE ROOT OF THE TREE IS NOT CLASSED AS A COALESCENT NODE FOR THIS TO WORK

Sample a coalescent node c uniformly at random. With probability $\frac{1}{2}$, a coalescent node split proposal is attempted, and otherwise a coalescent node merge proposal is attempted.

If a split proposal is attempted, a coalescent node i is sampled. Denote by ip the parent of i and by i_1 and i_2 denote the children of i. If ip is a migration node, the proposal proceeds by removing node ip and adding nodes $\hat{\imath}_1$, $\hat{\imath}_2$ at locations uniformly along edges $\langle i_1, i\rangle$ and $\langle i_2, i\rangle$ respectively. Then the demes on edges $\langle i_1, \hat{\imath}_1\rangle$ and $\langle i_2, \hat{\imath}_2\rangle$ are assigned to be the same as the initial demes on edges $\langle i_1, i\rangle$ and $\langle i_2, i\rangle$, and the demes on edges $\langle \hat{\imath}_1, i\rangle$ and $\langle \hat{\imath}_2, i\rangle$ are assigned to be the same as the initial deme on edge $\langle i, ip\rangle$. If ip is not a migration node, the proposed step is rejected without the MCMC algorithm recording a proposal.

If a merge move is attempted, a coalescent node i is sampled. Denote by ip the parent of i and by i_1 and i_2 denote the children of i. If both i_1 and i_2 are migration nodes, let ic_1 and ic_2 be the children of i_1 and i_2. If the demes of edges $\langle ic_1, i_1\rangle$ and $\langle ic_2, i_2\rangle$ are not the same, the proposed step is rejected without the MCMC algorithm recording a proposal. Otherwise, the proposal proceeds by removing nodes i_1 and i_2 and adding node $\hat{\imath}$ uniformly along edge $\langle i, \hat{\imath}\rangle$. The demes on edges $\langle ic_1, i\rangle$, $\langle ic_2, i\rangle$ and $\langle i, \hat{\imath}\rangle$ are set the same as the initial deme on edge $\langle ic_1, i_1\rangle$. Finally, the deme on edge $\langle \hat{\imath}, ip\rangle$ is updated to be same as the initial deme on edge $\langle i_1, i\rangle$. If either i_1 or i_2 is not a migration node, the proposed step is rejected without the MCMC algorithm recording a proposal.

The proposal ratio for the coalescent node split proposal is then given by

$$\begin{aligned}
Q_S(\mathcal{T}'|\mathcal{T}) &= \mathbb{P}[\substack{\text{Selected}\\\text{node}}] \cdot \mathbb{P}[\substack{\text{New node}\\\text{times}}] \\
&= \frac{1}{n-2} \cdot \frac{1}{\delta t_1 \delta t_2} \\
Q_M(\mathcal{T}|\mathcal{T}') &= \mathbb{P}[\substack{\text{Selected}\\\text{node}}] \cdot \mathbb{P}[\substack{\text{New node}\\\text{time}}] \\
&= \frac{1}{n-2} \cdot \frac{1}{\delta t} \\
\frac{Q_M(\mathcal{T}|\mathcal{T}')}{Q_S(\mathcal{T}'|\mathcal{T})} &= \frac{\delta t_1 \delta t_2}{\delta t}
\end{aligned}$$

where δt_j denotes the time between the selected node and its j^{th} child, and δt denotes the time between the selected node and its parent. Similarly, for the coalescent node merge proposal, the proposal ratio is given by

$$\frac{Q_S(\mathcal{T}|\mathcal{T}')}{Q_M(\mathcal{T}'|\mathcal{T})} = \frac{\delta t}{\delta t_1 \delta t_2}.$$

4.3.2 Updated Version 1 (Treat root as a coalescent node):

Could also treat the root of the tree as a coalescent node. For the split procedure, would need to sample a deme to add two new migration nodes below the root. For the merge procedure, would need to merge two migration nodes from below the root without adding a node above the root.

Now need cases for the proposal ratios. If a non-root coalescent node is sampled, they remain unchanged (except $\frac{1}{n-1}$ instead of $\frac{1}{n-2}$) from the Ewing et al. version. If the root is the sampled node, the proposal ratios become for the split proposal

$$\begin{aligned}
Q_S(\mathcal{T}'|\mathcal{T}) &= \frac{1}{n-1} \cdot \frac{1}{d-1} \cdot \frac{1}{\delta t_1 \delta t_2} \\
Q_M(\mathcal{T}|\mathcal{T}') &= \frac{1}{n-1} \\
\frac{Q_M(\mathcal{T}|\mathcal{T}')}{Q_S(\mathcal{T}'|\mathcal{T})} &= (d-1)\delta t_1 \delta t_2,
\end{aligned}$$

and for the merge proposal when sampling the root node,

$$\frac{Q_S(\mathcal{T}|\mathcal{T}')}{Q_M(\mathcal{T}'|\mathcal{T})} = \frac{1}{(d-1)\delta t_1 \delta t_2}.$$

4.4 Block Recolouring Proposal

Let a block of a structured coalescent genealogy be a maximal subtree of the migration history contained in a single deme. Then a block recolouring proposal aims to select a block within the migration history and resample the deme over that block.

A block of the migration history is selected with probability proportional to the length of the edges contained in the block, and a new deme is proposed over the block (not equal to the current deme). If the proposed deme does not provide a consistent deme labelling, reject the move.

The proposal ratio for the block recolouring proposal is then given by

$$\begin{aligned} Q_{BR}(\mathcal{T}'|\mathcal{T}) &= \mathbb{P}[^{\text{Selected}}_{\text{block}}] \cdot \mathbb{P}[^{\text{Proposed}}_{\text{deme}}] \\ &= \frac{\mathcal{L}_{\text{block}}}{\mathcal{L}} \cdot \frac{1}{d-1}. \end{aligned}$$

This move is symmetric in that

$$Q_{BR}(\mathcal{T}|\mathcal{T}') = Q_{BR}(\mathcal{T}'|\mathcal{T}),$$

and hence the proposal ratio is simply

$$\frac{Q_{BR}(\mathcal{T}|\mathcal{T}')}{Q_{BR}(\mathcal{T}'|\mathcal{T})} = 1.$$

4.5 Migration Shift

An alternative form of move which may be useful is to shift a migration node in time without modifying any other parameters. A migration node rescaling proposal will select a migration node uniformly at random and propose an update to the time of the event. To simplify the form of the move, the updated time will be required to still remain between the initial parent and child nodes within the migration history.

4.6 Option 1 (Resample event time uniformly along edge):

The updated time of the migration node could be selected by simply resampling a uniform time between the time of the selected node's parent and child nodes. Similarly to the block recolouring move, the migration time rescaling move will be self-inverting. The proposal probability is given by

$$\begin{aligned} Q_{TR}(\mathcal{T}'|\mathcal{T}) &= \mathbb{P}[^{\text{Selected migration}}_{\text{node}}] \cdot \mathbb{P}[^{\text{New node}}_{\text{time}}] \\ &= \frac{1}{M}\frac{1}{\delta t} \end{aligned}$$

where M denotes the number of migration nodes and δt denotes the difference in times between the parent and child of the selected migration node. The proposal ratio is then given by

$$\frac{Q_{TR}(\mathcal{T}|\mathcal{T}')}{Q_{TR}(\mathcal{T}'|\mathcal{T})} = 1.$$

4.7 Option 2 (Perturb current event time with reflecting boundaries):

Alternatively, the updated time of the migration node could be selected by perturbing the current time according to a normal distribution centred on the current time. The new node time t' would hence be given by

$$t' \sim N(t, \sigma^2)$$

for some suitably chosen variance σ^2 (POSSIBLY SELECTED USING ADAPTIVE MCMC??).

To preserve the ordering of the migration history, reflecting boundaries should be introduced at the times of the parent and child nodes. So, if the proposed time would fall outside of the interval of interest, reflect the "excess difference" back into the interval. Notably, the proposed perturbations are proposed from a symmetric distribution, and so the proposal ratio will again by given by

$$\frac{Q_{TR}(\mathcal{T}|\mathcal{T}')}{Q_{TR}(\mathcal{T}'|\mathcal{T})} = 1.$$

4.8 Migration Rates Matrix Update

The migration rates update matrix can be updated element-wise according to a Gibbs move. The likelihood for a single element of the migration rates matrix with the tree topology, migration history, effective population sizes and all other elements of the migration rates matrix fixed satisfies

$$L(\lambda_{ij}|\Lambda_{-(i,j)}, \theta, \mathcal{T}) \propto \lambda_{ij}^{m_{ij}} \exp\left\{-\left(\sum_{r=1}^{2n+M-2} k_{ir}\delta t_r\right)\lambda_{ij}\right\},$$

where

- $\Lambda = (\lambda_{ij})$ is the matrix of backwards-in-time migration rates,
- (δt_r) is the vector of time intervals between events,
- k_{ir} gives the number of lineages in deme i during time interval r,
- m_{ij} gives the total number of migration events $i \to j$.

Appealing to the Gamma-Gamma conjugacy, this motivates a Gamma prior on λ_{ij},

Gamma shape, rate parameterisation: $X \sim \text{Gamma}(\alpha, \beta)$ has density $f(x) \propto x^{\alpha-1}\exp\{-\beta x\}$

The posterior density then satisfies

$$\begin{aligned}\Pi(\lambda_{ij}|\Lambda_{-(i,j)}, \theta, \mathcal{T}) &\propto L(\lambda_{ij}|\Lambda_{-(i,j)}, \theta, \mathcal{T}) \cdot \pi(\lambda_{ij}) \\ &\propto \lambda_{ij}^{m_{ij}} \exp\left\{-\left(\sum_{r=1}^{2n+M-2} k_{ir}\delta t_r\right)\lambda_{ij}\right\} \cdot \lambda_{ij}^{\alpha-1}e^{-\beta\lambda_{ij}} \\ &\propto \lambda_{ij}^{m_{ij}+\alpha-1} \exp\left\{-\left(\beta + \sum_{r=1}^{2n+M-2} k_{ir}\delta t_r\right)\lambda_{ij}\right\},\end{aligned}$$

i.e. $\lambda_{ij}|... \sim \text{Gamma}\left(\alpha + m_{ij}, \beta + \sum k_{ir}\delta t_r\right)$.

4.9 Forwards-in-time Migration Rates

De Iorio and Griffiths (2004) obtain a relation between the forward-in-time and backward-in-time migration rates given by

$$\tilde{\lambda}_{ij} = \frac{\theta_j}{\theta_i}\lambda_{ji},$$

where $\tilde{\lambda}_{\cdot\cdot}$ denotes a forward-in-time migration rate. Substituting (the inverse of) this relation into the likelihood for the structured coalescent yields

$$\begin{aligned}
L_{n,d}(\mathcal{T}) &= \prod_{i=1}^{d} \frac{1}{\theta_i^{c_i}} \exp\left\{-\frac{1}{\theta_i}\sum_{r=1}^{2n+M-2}\binom{k_{ir}}{2}\delta t_r\right\} \prod_{\substack{j=1\\ j\neq i}} \tilde{\lambda}_{ji}^{m_{ij}} \left(\frac{\theta_j}{\theta_i}\right)^{m_{ij}} \exp\left\{-\frac{\theta_j}{\theta_i}\tilde{\lambda}_{ji}\sum_{r=1}^{2n+M-1} k_{ir}\delta t_r\right\} \\
&\propto_{\tilde{\lambda}_{ij}} \tilde{\lambda}_{ij}^{m_{j}i} \exp\left\{-\left(\frac{\theta_i}{\theta_j}\sum_{r=1}^{2n+M-2} k_{jr}\delta t_r\right)\tilde{\lambda}_{ij}\right\},
\end{aligned}$$

motivating a similar Gibbs update scheme to the backward-in-time migration rates, with a Gamma conjugate prior.

4.10 Effective Population Sizes Update

The effective population sizes vector can be updated similarly to the migration rates matrix with a Gibbs move. The likelihood for a single element of the effective population sizes vector with the tree topology, migration history, migration rates matrix and all other elements of the effective population sizes vector fixed satisfies

$$L(\theta_i|\theta_{-i},\lambda,\mathcal{T}) \propto \frac{1}{\theta_i^{c_i}}\exp\left\{-\left(\sum_{r=1}^{2n+M-2}\binom{k_{ir}}{2}\delta t_r\right)\frac{1}{\theta_i}\right\},$$

where c_i denotes the total number of coalescence events in deme i.

This likelihood corresponds to an inverse-gamma distribution on θ_i (equivalently a gamma distribution on $\frac{1}{\theta_i}$), hence appealing again to conjugacy, take an inverse gamma prior on θ_i. The posterior density then satisfies

$$\begin{aligned}
\Pi(\theta_i|\Lambda,\theta_{-i},\mathcal{T}) &\propto L(\theta_i|\Lambda,\theta_{-i},\mathcal{T})\cdot\pi(\theta_i) \\
&\propto \frac{1}{\theta_i^{c_i}}\exp\left\{-\left(\sum_{r=1}^{2n+M-2}\binom{k_{ir}}{2}\delta t_r\right)\frac{1}{\theta_i}\right\}\cdot\theta_i^{-\alpha-1}e^{-\frac{\beta}{\theta_i}} \\
&\propto \theta_i^{-\alpha-c_i-1}\exp\left\{-\left(\beta+\sum_{r=1}^{2n+M-2}\binom{k_{ir}}{2}\delta t_r\right)\frac{1}{\theta_i}\right\},
\end{aligned}$$

i.e. $\theta_i|... \sim$ Inverse-gamma$(\alpha + c_i, \beta + \sum\binom{k_{ir}}{2}\delta t_r)$.

4.11 Forward-in-time Migration Rates

Again substituting the relation from De Iorio and Griffiths (2004), the likelihood for the structured coalescent with forward-in-time migration rates satisfies

$$
\begin{aligned}
L_{n,d}(\mathcal{T}) &= \prod_{i=1}^{d} \frac{1}{\theta_i^{c_i}} \exp\left\{-\frac{1}{\theta_i}\sum_{r=1}^{2n+M-2}\binom{k_{ir}}{2}\delta t_r\right\} \prod_{\substack{j=1\\ j\neq i}} \tilde{\lambda}_{ji}^{m_{ij}} \left(\frac{\theta_j}{\theta_i}\right)^{m_{ij}} \exp\left\{-\frac{\theta_j}{\theta_i}\tilde{\lambda}_{ji}\sum_{r=1}^{2n+M-1} k_{ir}\delta t_r\right\} \\
&\propto_{\theta_1} \frac{1}{\theta_1^{c_1}} \exp\left\{-\frac{1}{\theta_1}\sum_{r}\binom{k_{1r}}{2}\delta t_r\right\} \times \underbrace{\left[\prod_{\alpha=2}^{d} \frac{1}{\theta_1^{m_{1\alpha}}} \exp\left\{-\frac{1}{\theta_1}\left(\theta_\alpha \tilde{\lambda}_{\alpha 1}\sum_{\gamma} k_{1\gamma}\delta t_\gamma\right)\right\}\right]}_{i=1} \\
&\qquad \times \underbrace{\left[\prod_{\eta=2}^{d} \theta_1^{m_{\eta 1}} \exp\left\{-\theta_1\left(\frac{\tilde{\lambda}_{1\eta}}{\theta_\eta}\sum_{\nu} k_{\eta\nu}\delta t_\nu\right)\right\}\right]}_{i\neq 1} \\
&\propto_{\theta_1} \theta_1^{m_{+1}-m_{1+}-c_1} \exp\left\{-\frac{1}{\theta_1}\sum_{r}\left[\binom{k_{1r}}{2} + k_{1r}\sum_{\alpha=2}^{d}\theta_\alpha\tilde{\lambda}_{\alpha 1}\right]\delta t_r - \theta_1\sum_{\beta=2}^{d}\frac{\tilde{\lambda}_{1\beta}}{\theta_\beta}\sum_{r} k_{\beta r}\delta t_r\right\} \\
&\sim_{\theta_1} \theta_1^{-a-1} \exp\left\{-\frac{b}{\theta} - c\theta\right\},
\end{aligned}
$$

where $m_{+i} = \sum_{\alpha=1}^{d} m_{\alpha i}$, $m_{i+} = \sum_{\alpha=1}^{d} m_{i\alpha}$ and a, b, c are constants satisfying

- $a \in \mathbb{Z}$
- $b \in (0, \infty)$
- $c \in (0, \infty)$.

This density resembles a product of the densities for a gamma distribution and inverse gamma distribution, however is not itself a common distribution. This suggests that if samples can be drawn from this distribution, either a Gamma or Inverse-gamma prior would be conjugate under this likelihood.

Samples can be drawn using a rejection sampling method, by noting that for $b, c, x \geq 0$, both $e^{-}b/x \leq 1$ and $e^{-}cx \leq 1$ hold.

Rejection Sampling:

1. If $a \geq 0$, sample $X \sim \text{Inv-Gamma}(a, b)$ [with density $f(x) \propto x^{-a-1}e^{-b/x}$]. Accept X with probability
$$\frac{x^{-a-1}e^{-\frac{b}{x}-cx}}{x^{-a-1}e^{-b/x}} = e^{-cx}.$$
2. Otherwise, if $a < 0$, sample $X \sim \text{Gamma}(a, c)$ [with density $f(x) \propto x^{-a-1}e^{-cx}$]. Accept X with probability
$$\frac{x^{-a-1}e^{-\frac{b}{x}-cx}}{x^{-a-1}e^{-cx}} = e^{-\frac{b}{x}}.$$

Do I need to do something with M - I'm not sure whether $M = 1$ is sufficient here since, or whether I need to think more about relative size of my bounding functions?

Rejection Sampling Algorithm:
Given two densities f, g with $f(x) \leq M \cdot g(x)$ for all x, a sample can be generated from f as follows:

1. Draw $X \sim g$
2. Accept X as a sample from f with probability

$$\frac{f(X)}{M \cdot g(X)},$$

otherwise return to step 1.

5 MCMC Testing

Due to the complexity of the proposals involved in a MCMC scheme on the structured coalescent model, verifying the results of the iterations can be difficult. One method to verify the iterations is to use the structured coalescent models to generate realisations from an alternative distribution. For example, this can be achieved by using a Poisson(λ) prior on the number of migration events in the tree (with additional terms in the prior to obtain a full distribution over migration histories).

Restricting attention to the case of using a Poisson prior on the number of migration events with only migration birth/death moves. The full prior for the verification model is given by

$$p(\theta) = \frac{\lambda^M e^{-\lambda}}{M!} \cdot \left(\frac{1}{\mathcal{L}}\right)^M \cdot \left(\frac{1}{d-1}\right)^M \cdot \frac{1}{d} \cdot M!,$$

where the term $\frac{\lambda^M e^{-\lambda}}{M!}$ arises as the Poisson-likelihood of having M migration events on the tree; the term $\left(\frac{1}{\mathcal{L}}\right)^M$ arises from selecting migration event locations uniformly on the tree; the term $\left(\frac{1}{d-1}\right)^M \cdot \frac{1}{d}$ arises from selecting demes on the $M+1$ disjoint subtrees of the tree with roots either the root of the tree, or migration nodes (d demes to select from for the subtree containing the root and $d-1$ to select from for each of the M subtrees below migration nodes); and the term $M!$ arises from summing over all possible labellings of the M migration events.

An algorithm can then be derived which has acceptance probabilities

$$\alpha(\mathcal{T}'|\mathcal{T}) = \min\left(1, \frac{p(\theta')Q(\mathcal{T}|\mathcal{T}')}{p(\theta)Q(\mathcal{T}'|\mathcal{T})}||\mathcal{J}||\right),$$

noting that θ' corresponds to having either added a single migration node (migration birth proposal) or removed a single migration node (migration death proposal).

5.1 Migration Birth/Death Proposals

The prior ratio for the migration birth move is given by

$$\begin{aligned}\frac{p(\theta')}{p(\theta)} &= \frac{\frac{\lambda^{M+1}e^{-\lambda}}{(M+1)!}\cdot\left(\frac{1}{\mathcal{L}}\right)^{M+1}\cdot\left(\frac{1}{d-1}\right)^{M+1}\cdot\frac{1}{d}\cdot(M+1)!}{\frac{\lambda^{M}e^{-\lambda}}{M!}\cdot\left(\frac{1}{\mathcal{L}}\right)^{M}\cdot\left(\frac{1}{d-1}\right)^{M}\cdot\frac{1}{d}\cdot M!}\\ &= \frac{\lambda}{(d-1)\mathcal{L}},\end{aligned}$$

and the acceptance probability for the second updated version of the move culminates with

$$\alpha_B(\mathcal{T}'|\mathcal{T}) = \min\left(1, \frac{\lambda}{(d-1)\mathcal{L}}\frac{(d-1)\mathcal{L}}{M+1}\right) = \min\left(1, \frac{\lambda}{M+1}\right).$$

Similarly, for the migration death move

$$\begin{aligned}\frac{p(\theta')}{p(\theta)} &= \frac{\frac{\lambda^{M-1}e^{-\lambda}}{(M-1)!}\cdot\left(\frac{1}{\mathcal{L}}\right)^{M-1}\cdot\left(\frac{1}{d-1}\right)^{M-1}\cdot\frac{1}{d}\cdot(M-1)!}{\frac{\lambda^{M}e^{-\lambda}}{M!}\cdot\left(\frac{1}{\mathcal{L}}\right)^{M}\cdot\left(\frac{1}{d-1}\right)^{M}\cdot\frac{1}{d}\cdot M!}\\ &= \frac{(d-1)\mathcal{L}}{\lambda}\end{aligned}$$

and

$$\alpha_D(\mathcal{T}'|\mathcal{T}) = \min\left(1, \frac{(d-1)\mathcal{L}}{\lambda}\frac{M}{(d-1)\mathcal{L}}\right) = \min\left(1, \frac{M}{\lambda}\right).$$

5.2 Migration Pair Birth/Death Proposals

The prior ratio for the migration birth move is given by

$$\begin{aligned}\frac{p(\theta')}{p(\theta)} &= \frac{\frac{\lambda^{M+2}e^{-\lambda}}{(M+2)!}\cdot\left(\frac{1}{\mathcal{L}}\right)^{M+2}\cdot\left(\frac{1}{d-1}\right)^{M+2}\cdot\frac{1}{d}\cdot(M+2)!}{\frac{\lambda^{M}e^{-\lambda}}{M!}\cdot\left(\frac{1}{\mathcal{L}}\right)^{M}\cdot\left(\frac{1}{d-1}\right)^{M}\cdot\frac{1}{d}\cdot M!}\\ &= \frac{\lambda^2}{(d-1)^2\mathcal{L}^2},\end{aligned}$$

and the prior ratio for the migration death move is given by

$$\begin{aligned}\frac{p(\theta')}{p(\theta)} &= \frac{\frac{\lambda^{M-2}e^{-\lambda}}{(M-2)!}\cdot\left(\frac{1}{\mathcal{L}}\right)^{M-2}\cdot\left(\frac{1}{d-1}\right)^{M-2}\cdot\frac{1}{d}\cdot(M-2)!}{\frac{\lambda^{M}e^{-\lambda}}{M!}\cdot\left(\frac{1}{\mathcal{L}}\right)^{M}\cdot\left(\frac{1}{d-1}\right)^{M}\cdot\frac{1}{d}\cdot M!}\\ &= \frac{(d-1)^2\mathcal{L}^2}{\lambda^2}.\end{aligned}$$

The migration pair birth/death proposal MCMC scheme can only update the number of migrations by incrementing it by 2, thus the sample of M will return only odd or even values depending on the initial configuration. Thus, given an $X \sim \text{Poi}(\lambda)$, the sample should be compared with either $X|(X\,\text{even})$ or $X|(X\,\text{odd})$.

In the even case,

$$\mathbb{P}[X = k|X\,\text{even}] = \frac{\mathbb{P}[X = k]}{\mathbb{P}[X\,\text{even}]},$$

where

$$\begin{aligned}\mathbb{P}[X\,\text{even}] &= \frac{1}{2}\sum_{j=0}^{\infty}\frac{\lambda^j e^{-\lambda}}{j!}(1+(-1)^j)\\ &= \frac{1}{2}+\frac{e^{-\lambda}}{2}\sum_{j=0}^{\infty}\frac{(-\lambda)^j}{j!}\\ &= \frac{1}{2}(1+e^{-2\lambda}),\end{aligned}$$

thus

$$\mathbb{P}[X=k|X\,\text{even}] = \frac{2\lambda^k e^{-\lambda}}{(1+e^{-2\lambda})k!}\mathbb{1}\{k\,\text{even}\},$$

and

$$\mathbb{P}[X=k|X\,\text{odd}] = \frac{2\lambda^k e^{-\lambda}}{(1-e^{-2\lambda})k!}\mathbb{1}\{k\,\text{odd}\}.$$

6 Detailed Balance

Let $\pi(\cdot)$ denote the stationary distribution of the RJMCMC algorithm (for testing purposes, $\pi(\mathcal{T}) = (\frac{\lambda}{\mathcal{L}(d-1)})^M\frac{e^{-\lambda}}{d}$) and $p(\cdot|\cdot)$ denote the one-step transition probabilities. Then the MCMC algorithm satisfies detailed balance if

$$p(\mathcal{T}_{k+1}|\mathcal{T}_k)\pi(\mathcal{T}_k) = p(\mathcal{T}_k|\mathcal{T}_{k+1})\pi(\mathcal{T}_{k+1}).$$

Assuming that $\mathcal{T}_{k+1}\neq\mathcal{T}_k$, the transition $\mathcal{T}_k\mapsto\mathcal{T}_{k+1}$ can happen only by accepting a proposal, with transition probability

$$p(\mathcal{T}_{k+1}|\mathcal{T}_k) = \alpha(\mathcal{T}_{k+1}|\mathcal{T}_k)Q(\mathcal{T}_{k+1}|\mathcal{T}_k).$$

6.1 Migration birth/death

Let $\mathcal{T}_{k+1} = \text{migration.birth}(\mathcal{T}_k)$, then

$$\begin{aligned}\frac{p(\mathcal{T}_{k+1}|\mathcal{T}_k)}{p(\mathcal{T}_k|\mathcal{T}_{k+1})} &= \frac{\alpha(\mathcal{T}_{k+1}|\mathcal{T}_k)Q(\mathcal{T}_{k+1}|\mathcal{T}_k)}{\alpha(\mathcal{T}_k|\mathcal{T}_{k+1})Q(\mathcal{T}_k|\mathcal{T}_{k+1})}\\ &= \frac{\min\left(1,\frac{\lambda}{M+1}\right)\cdot\frac{1}{(d-1)\mathcal{L}}}{\min\left(1,\frac{M+1}{\lambda}\right)\cdot\frac{1}{M+1}}\\ &= \frac{\lambda}{(d-1)\mathcal{L}}\cdot\frac{\min(\lambda,M+1)}{\min(\lambda,M+1)}\\ &= \frac{\pi(\mathcal{T}_{k+1})}{\pi(\mathcal{T}_k)}.\end{aligned}$$

Otherwise, let $\mathcal{T}_{k+1} = \text{migration.death}(\mathcal{T}_k)$,

$$\begin{aligned}
\frac{p(\mathcal{T}_{k+1}|\mathcal{T}_k)}{p(\mathcal{T}_k|\mathcal{T}_{k+1})} &= \frac{\alpha(\mathcal{T}_{k+1}|\mathcal{T}_k)Q(\mathcal{T}_{k+1}|\mathcal{T}_k)}{\alpha(\mathcal{T}_k|\mathcal{T}_{k+1})Q(\mathcal{T}_k|\mathcal{T}_{k+1})} \\
&= \frac{\min\left(1, \frac{M}{\lambda}\right) \cdot \frac{1}{M}}{\min\left(1, \frac{\lambda}{M}\right) \cdot \frac{1}{(d-1)\mathcal{L}}} \\
&= \frac{(d-1)\mathcal{L}}{\lambda} \cdot \frac{\min(\lambda, M)}{\min(\lambda.M)} \\
&= \frac{\pi(\mathcal{T}_{k+1})}{\pi(\mathcal{T}_k)}.
\end{aligned}$$

Hence, detailed balance holds for the migration birth/death moves.

6.2 Migration pair birth/death

Let $\mathcal{T}_{k+1} = \text{pair.birth}(\mathcal{T}_k)$, then

$$\begin{aligned}
\frac{p(\mathcal{T}_{k+1}|\mathcal{T}_k)}{p(\mathcal{T}_k|\mathcal{T}_{k+1})} &= \frac{\alpha(\mathcal{T}_{k+1}|\mathcal{T}_k)Q(\mathcal{T}_{k+1}|\mathcal{T}_k)}{\alpha(\mathcal{T}_k|\mathcal{T}_{k+1})Q(\mathcal{T}_k|\mathcal{T}_{k+1})} \\
&= \frac{\min\left\{1, \left(\frac{\lambda}{(d-1)\mathcal{L}}\right)^2 \frac{(2n+M-2)(d-1)\delta t^2}{2(2n+M)}\right\} \cdot \frac{2}{(2n+M-2)(d-1)\delta t^2}}{\min\left\{1, \left(\frac{(d-1)\mathcal{L}}{\lambda}\right)^2 \frac{2(2n+M)}{(2n+M-2)(d-1)\delta t^2}\right\} \cdot \frac{1}{2n+M}} \\
&= \frac{\min\left\{2(2n+M), \left(\frac{\lambda}{(d-1)\mathcal{L}}\right)^2 (2n+M-2)(d-1)\delta t^2\right\}}{\min\left\{(2n+M-2)(d-1)\delta t^2, \left(\frac{(d-1)\mathcal{L}}{\lambda}\right)^2 \cdot 2(2n+M)\right\}} \\
&= \left(\frac{\lambda}{(d-1)\mathcal{L}}\right)^2 \cdot \frac{\min\left\{2(2n+M)\left(\frac{(d-1)\mathcal{L}}{\lambda}\right)^2, (2n+M-2)(d-1)\delta t^2\right\}}{\min\left\{2(2n+M)\left(\frac{(d-1)\mathcal{L}}{\lambda}\right)^2, (2n+M-2)(d-1)\delta t^2\right\}} \\
&= \frac{\pi(\mathcal{T}_{k+1})}{\pi(\mathcal{T}_k)}.
\end{aligned}$$

Otherwise, let $\mathcal{T}_{k+1} = \text{pair.death}(\mathcal{T}_k)$, then

$$\begin{aligned}
\frac{p(\mathcal{T}_{k+1}|\mathcal{T}_k)}{p(\mathcal{T}_k|\mathcal{T}_{k+1})} &= \frac{\alpha(\mathcal{T}_{k+1}|\mathcal{T}_k)Q(\mathcal{T}_{k+1}|\mathcal{T}_k)}{\alpha(\mathcal{T}_k|\mathcal{T}_{k+1})Q(\mathcal{T}_k|\mathcal{T}_{k+1})} \\
&= \frac{\min\left\{1, \left(\frac{(d-1)\mathcal{L}}{\lambda}\right)^2 \frac{2(2n+M-2)}{(2n+M-4)(d-1)\delta t^2}\right\} \cdot \frac{1}{2n+M-2}}{\min\left\{1, \left(\frac{\lambda}{(d-1)\mathcal{L}}\right)^2 \frac{(2n+M-4)(d-1)\delta t^2}{2(2n+M-2)}\right\} \cdot \frac{2}{(2n+M-4)(d-1)\delta t^2}} \\
&= \frac{\min\left\{(2n+M-4)(d-1)\delta t^2, 2\left(\frac{(d-1)\mathcal{L}}{\lambda}\right)^2(2n+M-2)\right\}}{\min\left\{2(2n+M-2), \left(\frac{\lambda}{(d-1)\mathcal{L}}\right)^2(2n+M-4)(d-1)\delta t^2\right\}} \\
&= \left(\frac{(d-1)\mathcal{L}}{\lambda}\right)^2 \cdot \frac{\min\left\{(2n+M-4)(d-1)\delta t^2\left(\frac{\lambda}{(d-1)\mathcal{L}}\right)^2, 2(2n+M-2)\right\}}{\min\left\{(2n+M-4)(d-1)\delta t^2\left(\frac{\lambda}{(d-1)\mathcal{L}}\right)^2, 2(2n+M-2)\right\}} \\
&= \frac{\pi(\mathcal{T}_{k+1})}{\pi(\mathcal{T}_k)}.
\end{aligned}$$

6.3 Coalescent Node Split/Merge

Let $\mathcal{T}_{k+1} = \text{coalescent.split}(\mathcal{T}_k)$ and assume the split event does not target the root node, then

$$\begin{aligned}
\frac{p(\mathcal{T}_{k+1}|\mathcal{T}_k)}{p(\mathcal{T}_k|\mathcal{T}_{k+1})} &= \frac{\alpha(\mathcal{T}_{k+1}|\mathcal{T}_k)Q(\mathcal{T}_{k+1}|\mathcal{T}_k)}{\alpha(\mathcal{T}_k|\mathcal{T}_{k+1})Q(\mathcal{T}_k|\mathcal{T}_{k+1})} \\
&= \frac{\min\left\{1, \frac{\delta t_1 \delta t_2}{\delta t} \cdot \frac{\lambda}{(d-1)\mathcal{L}}\right\} \frac{1}{(n-1)\delta t_1 \delta t_2}}{\min\left\{1, \frac{\delta t}{\delta t_1 \delta t_2} \cdot \frac{(d-1)\mathcal{L}}{\lambda}\right\} \frac{1}{(n-1)\delta t}} \\
&= \left(\frac{\lambda}{(d-1)\mathcal{L}}\right) \cdot \frac{\min\left\{\frac{(d-1)\mathcal{L}}{\lambda}\delta t, \delta t_1 \delta t_2\right\}}{\min\left\{\frac{(d-1)\mathcal{L}}{\lambda}\delta t, \delta t_1 \delta t_2\right\}} \\
&= \frac{\pi(\mathcal{T}_{k+1})}{\pi(\mathcal{T}_k)}.
\end{aligned}$$

The corresponding merge event (assuming it does not target the root node) is similar and also satisfies detailed balance.

Now, let $\mathcal{T}_{k+1} = \text{coalescent.split}(\mathcal{T}_k)$ and assume the split event does target the root node, then

$$
\begin{aligned}
\frac{p(\mathcal{T}_{k+1}|\mathcal{T}_k)}{p(\mathcal{T}_k|\mathcal{T}_{k+1})} &= \frac{\alpha(\mathcal{T}_{k+1}|\mathcal{T}_k)Q(\mathcal{T}_{k+1}|\mathcal{T}_k)}{\alpha(\mathcal{T}_k|\mathcal{T}_{k+1})Q(\mathcal{T}_k|\mathcal{T}_{k+1})} \\
&= \frac{\min\left\{1, \delta t_1 \delta t_2 (d-1) \cdot \frac{\lambda}{(d-1)\mathcal{L}}\right\} \frac{1}{(n-1)(d-1)\delta t_1 \delta t_2}}{\min\left\{1, \frac{1}{(d-1)\delta t_1 \delta t_2} \cdot \frac{(d-1)\mathcal{L}}{\lambda}\right\} \frac{1}{(n-1)}} \\
&= \left(\frac{\lambda}{(d-1)\mathcal{L}}\right) \cdot \frac{\min\left\{\frac{(d-1)\mathcal{L}}{\lambda}, \delta t_1 \delta t_2 (d-1)\right\}}{\min\left\{\frac{(d-1)\mathcal{L}}{\lambda}, \delta t_1 \delta t_2 (d-1)\right\}} \\
&= \frac{\pi(\mathcal{T}_{k+1})}{\pi(\mathcal{T}_k)}.
\end{aligned}
$$

The corresponding merge event (assuming it does not target the root node) is similar and also satisfies detailed balance.

(NO DEPENDENCE ON M SO THE MERGE EVENTS ARE JUST THE RECIPROCALS OF THE SPLIT EVENTS)