

## General way for a set of trees

Let  $S = (s_1, \dots, s_m)$  be a set of trees and  $\mathcal{L} = (l_1(\theta), \dots, l_m(\theta))$  the set of log-likelihood functions of  $S$ .

Then

$$l_j(\theta) = \sum_{i=1}^{N_j} -\sigma_{i,j} t_{i,j} + \log(\rho_{i,j})$$

where  $N_j$  is the number of branching times of the  $j$ -tree,  $t_{i,j}$  is the  $i^{th}$  branching time of the  $j$ -tree and  $\sigma_{i,j}$  and  $\rho_{i,j}$  are functions of  $\lambda_{i,j}(\theta)$  and  $\mu_{i,j}(\theta)$ , which are the speciation and extinction rates of the species of the tree  $j$  at time  $t_{i,j}$  as described in previous reports.

In order to solve the E-step on the EM routine, under the monte-carlo approach, we need to calculate

$$l(\theta) = \sum_{j=1}^m l_j(\theta)$$

the M-step corresponds to find  $\max_{\theta} l(\theta)$ .

## Diversity dependence model

As described previously, we define

$$l_j(\theta) = \text{log-likelihood}(\theta | s_j)$$

in the case of diversity-dependence, where we have

$$\begin{aligned} \sigma_{i,j} &= n_{i,j}\lambda - \beta n_{i,j}^2 + n_{i,j}\mu \\ \rho_{i,j} &= E_{i,j}(\lambda - \beta n_{i,j}) + (1 - E_{i,j})\mu \end{aligned}$$

Thus,

$$l_j((\lambda, \beta, \mu)) = \sum_{i=1}^{N_j} -t_{i,j} [n_{i,j}\lambda - n_{i,j}^2\beta - n_{i,j}\mu] + \log(E_{i,j}(\lambda - \beta n_{i,j}) + (1 - E_{i,j})\mu)$$

Moreover,

$$l(\theta) = \sum_{j=1}^m \sum_{i=1}^{N_j} -t_{i,j} [n_{i,j}\lambda - n_{i,j}^2\beta + n_{i,j}\mu] + \log(E_{i,j}(\lambda - \beta n_{i,j}) + (1 - E_{i,j})\mu)$$

where, as in the case of 1 single tree, we can find an analytical solution for the parameter  $\mu$

$$\mu = \frac{\sum_{j=1}^m \sum_{i=1}^{N_j} (1 - E_{i,j})}{\sum_{j=1}^m \sum_{i=1}^{N_j} t_{i,j} n_{i,j}}$$

and the other two parameters might be calculated under the same framework of the single tree case.