

Genome 540 Discussion

March 5th 2026



HW8 Overview

HW8 Overview

Start with:

Initiation probabilities

Transition probabilities

Emission probabilities

(Arbitrary)

Nucleotide sequence

(Data)

Goal:

Find *better estimates* of said probabilities!

HW8 Overview

Goal:

Find *better estimates* of HMM probabilities!

Viterbi strategy:

Find the *most likely* path through the graph (remember: what does a path correspond to?)

Using that *most likely path*, update the emission, transmission, and initiation probabilities.

HW8 Overview

Goal:

Find *better estimates* of HMM probabilities!

Another strategy:

Simulate taking *many paths* through the DAG, then using all of those possible paths, update the transmission, emission and initiation probabilities, *weighted by how likely a given path is*.

HW8 Overview

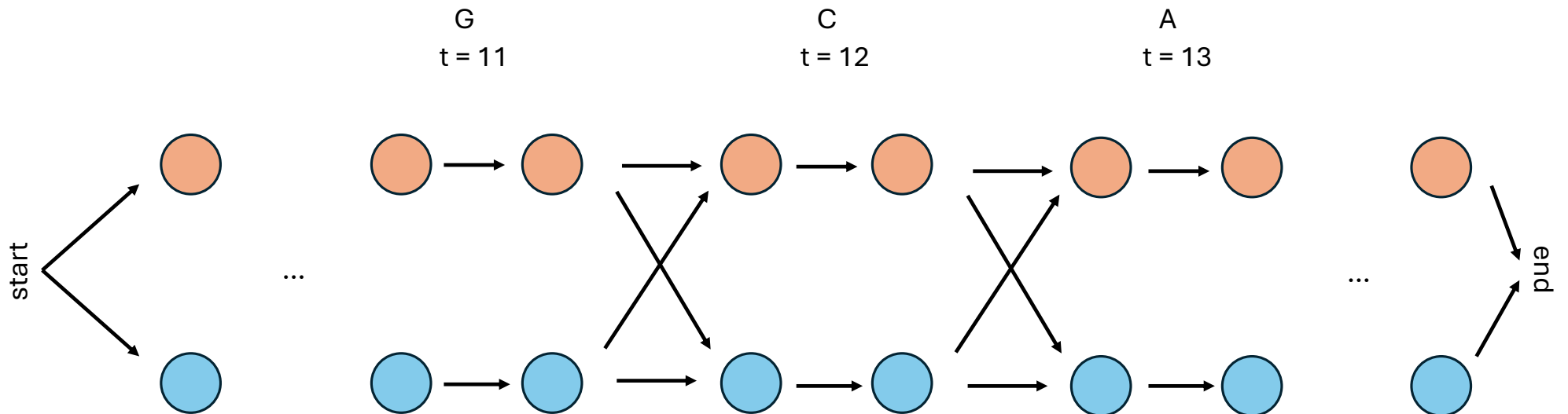
Goal:

Find *better estimates* of HMM probabilities!

Baum-Welch:

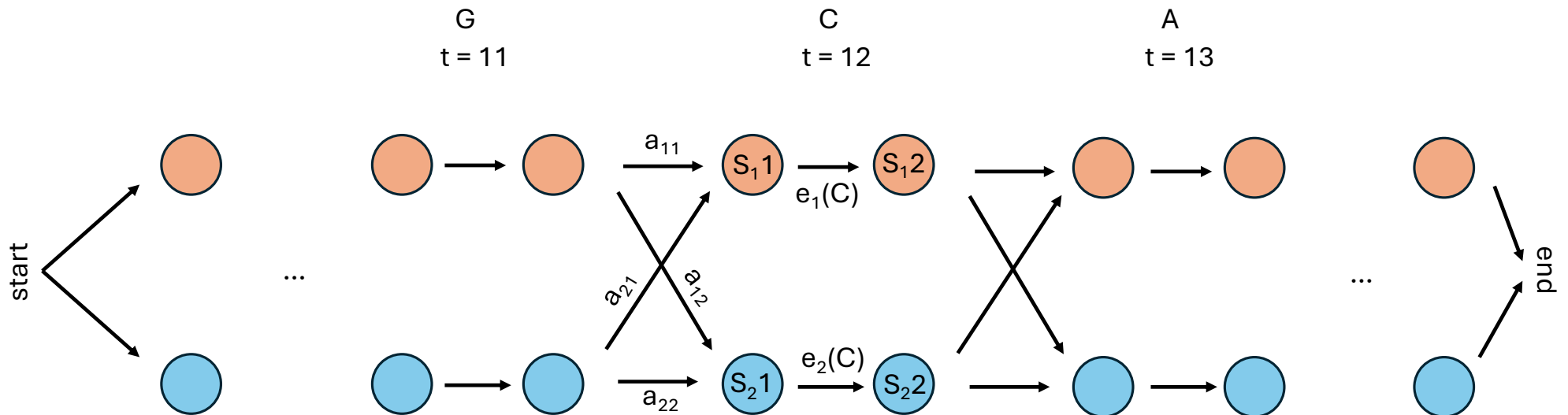
Calculate how likely I am to be in a given state at a given time. (considering all paths!) Then: update transmission and emission probabilities, *weighted by how likely you are to be in that state!*

HW8 Overview



initiation probs: π_1, π_2
 transition probs: $a_{11}, a_{12}, a_{21}, a_{22}$
 emission probs:
 $e_1(A), e_1(C), e_1(G), e_1(T)$
 $e_2(A), e_2(C), e_2(G), e_2(T)$

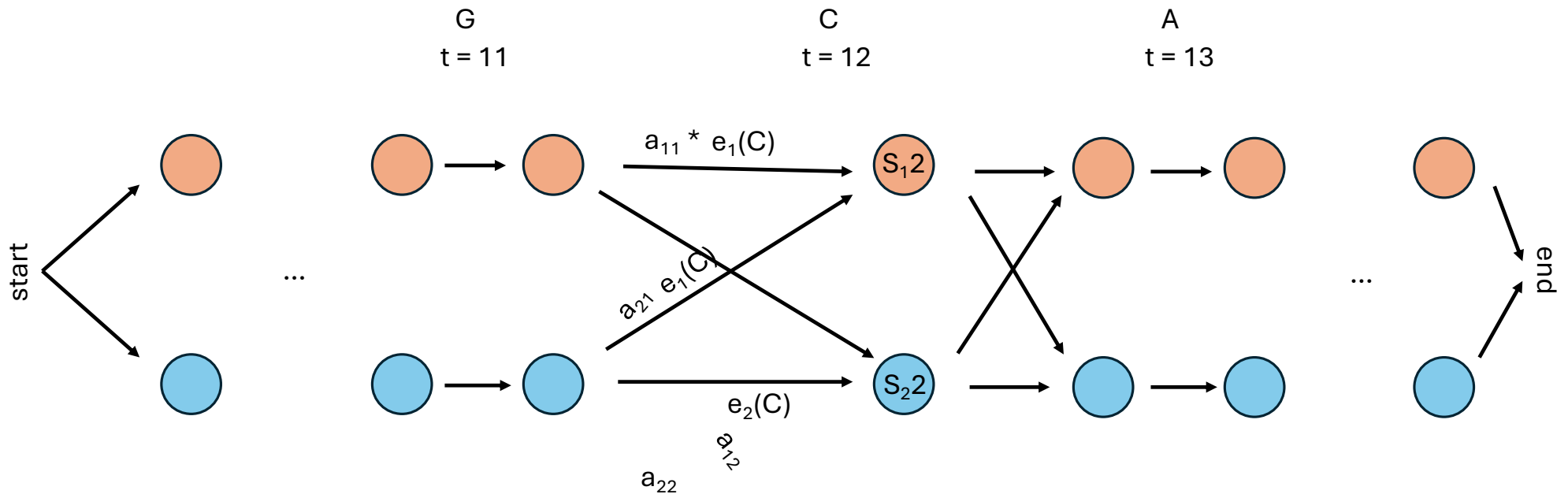
HW8 Overview



$F(\text{node}) = \text{sum of all paths that end at that node}$

$$F(S_{12}(t=12)) = (F(S_{11}(t=11)) * a_{11}) + (F(S_{21}(t=11)) * a_{21} * e_1(C))$$

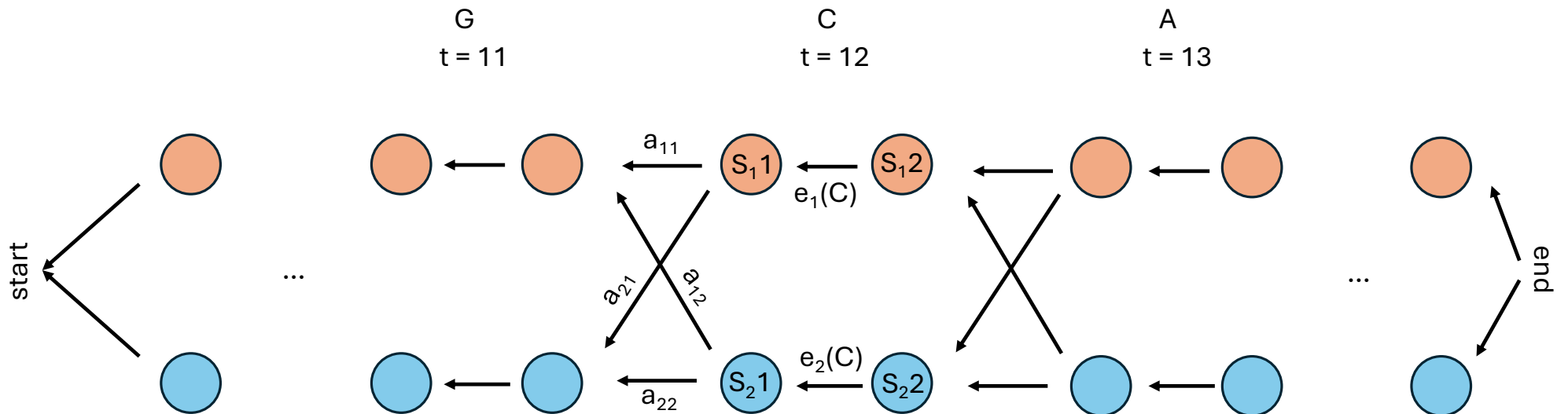
HW8 Overview



$F(\text{node}) = \text{sum of all paths that end at that node}$

$$F(S_12(t = 12)) = F(S_12(t = 11)) * a_{11} * e_1("C") + F(S_22(t = 11)) * a_{21} * e_1("C")$$

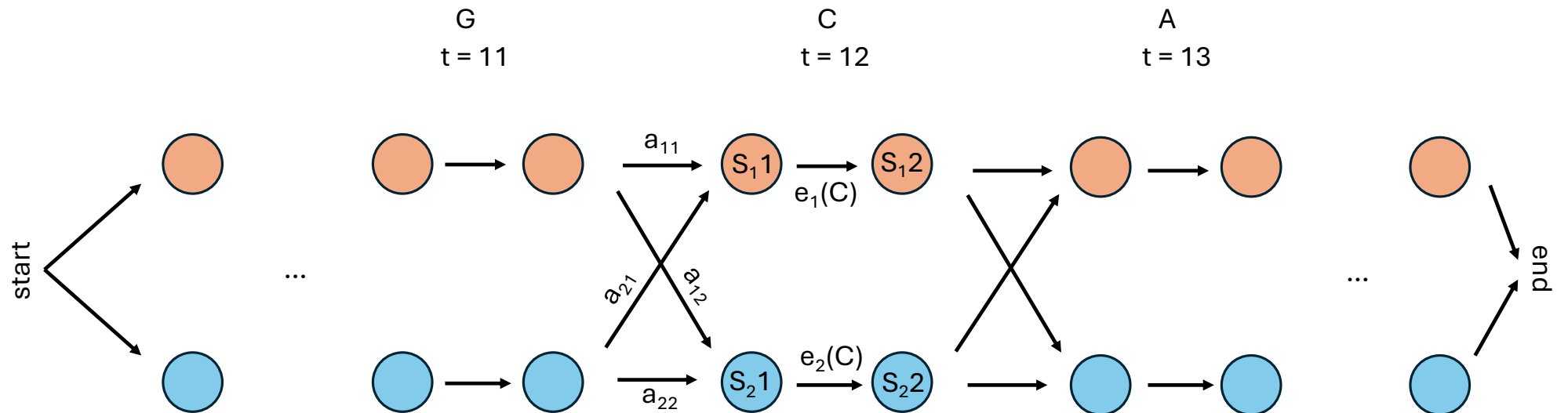
HW8 Overview



$B(\text{node}) = \text{sum of all paths that start at that node}$

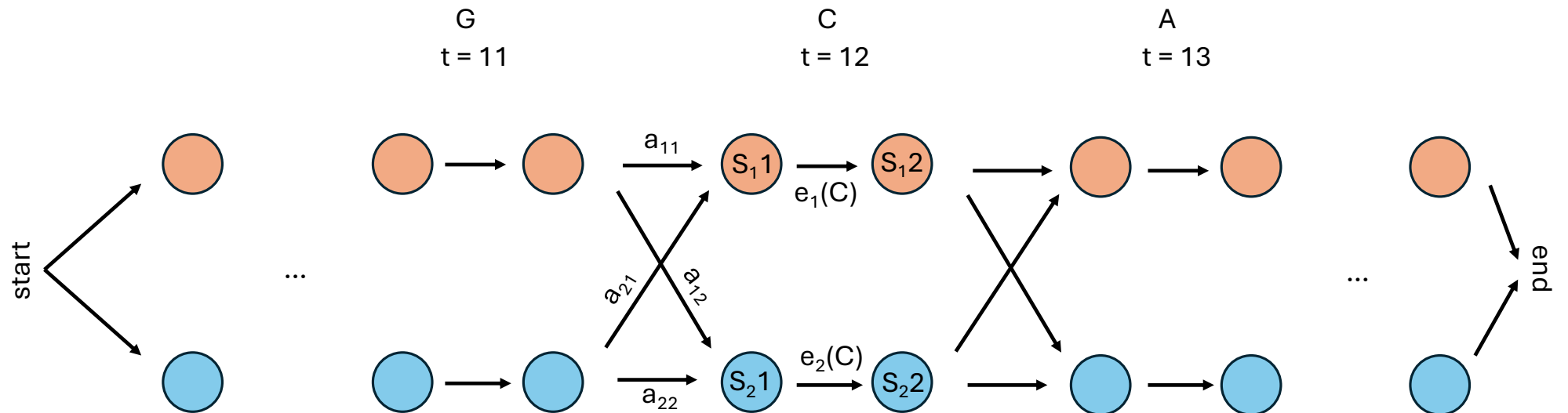
$$B(S_{12}(t=12)) = ?$$

HW8 Overview



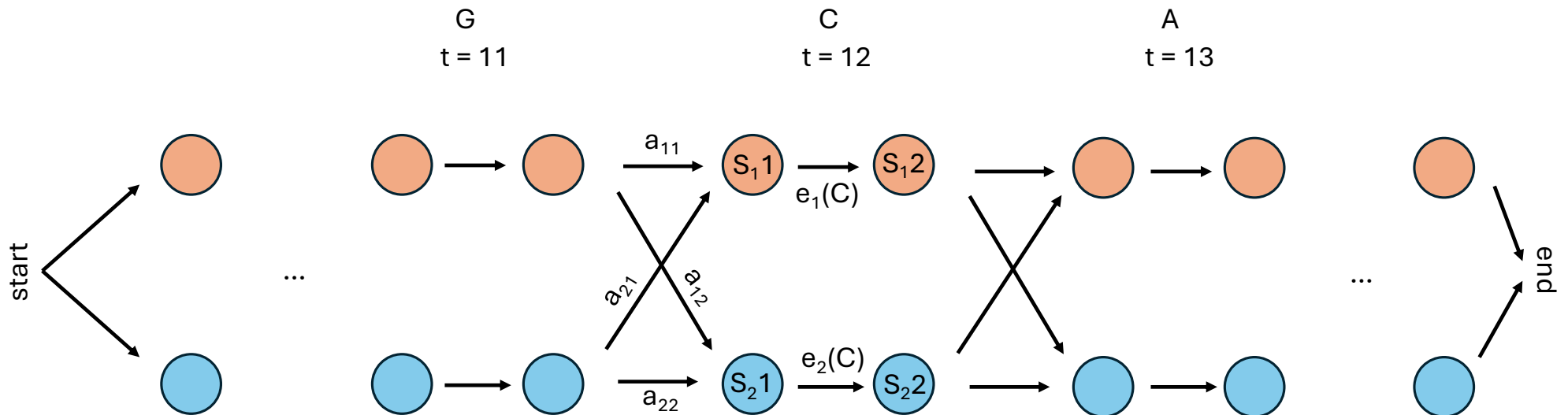
$$F(S_{12}(t = 12)) * B(S_{12}(t = 12)) = ? \text{ (with words?)}$$

HW8 Overview



$$\sum_{t=1, n="C"}^N F(S_{12}(t)) * B(S_{12}(t)) = ? \quad (\text{with words?})$$

HW8 Overview



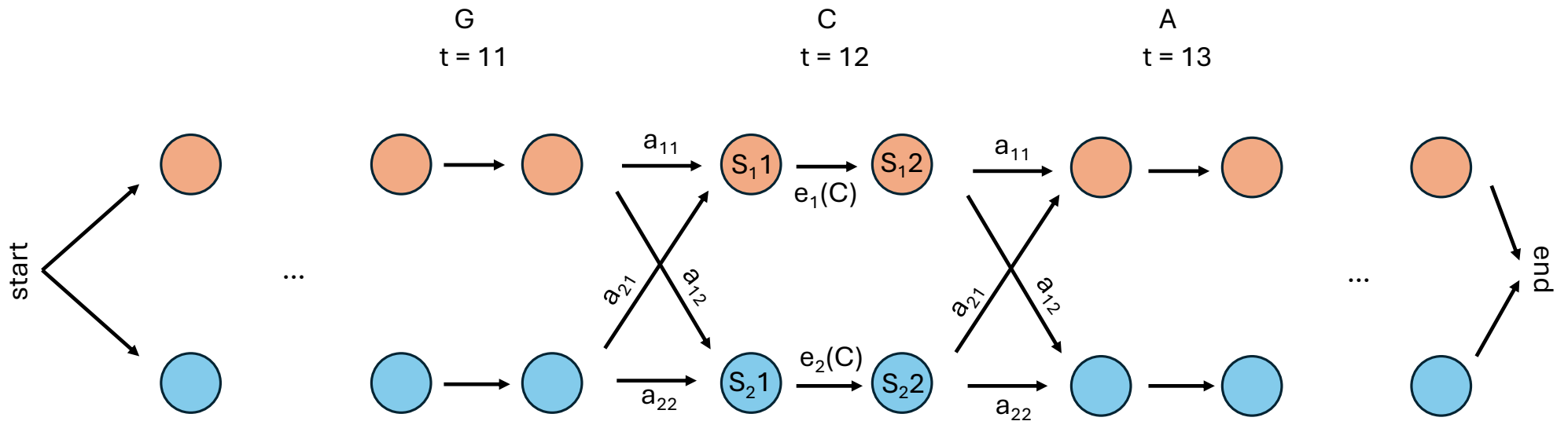
$$\frac{\sum_{t=1, n="C"}^N F(S_{12}(t)) * B(S_{12}(t))}{\sum_{t=1}^N F(S_{12}(t)) * B(S_{12}(t))} = ?$$

updated probability
of emitting a "C"
from state 1

$$\frac{\sum_{t=1, n="T"}^N F(S_{22}(t)) * B(S_{22}(t))}{\sum_{t=1}^N F(S_{22}(t)) * B(S_{22}(t))} = ?$$

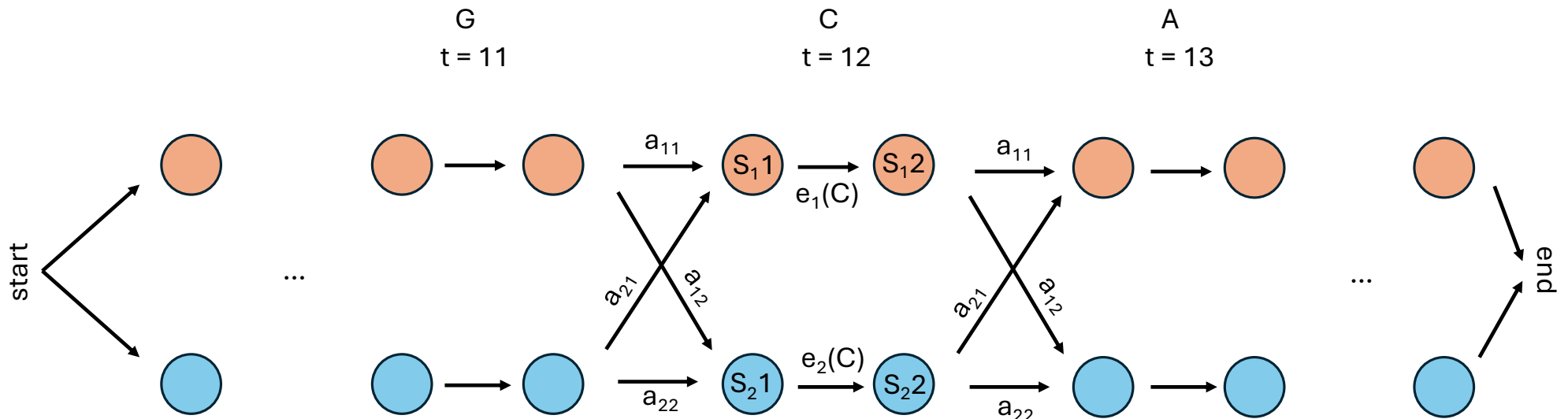
updated probability
of emitting a "T" from
state 2

HW8 Overview



$$\overline{a_{21}} = F(S_{22}(t=12)) * a_{21} * B(S_{11}(t=13))$$

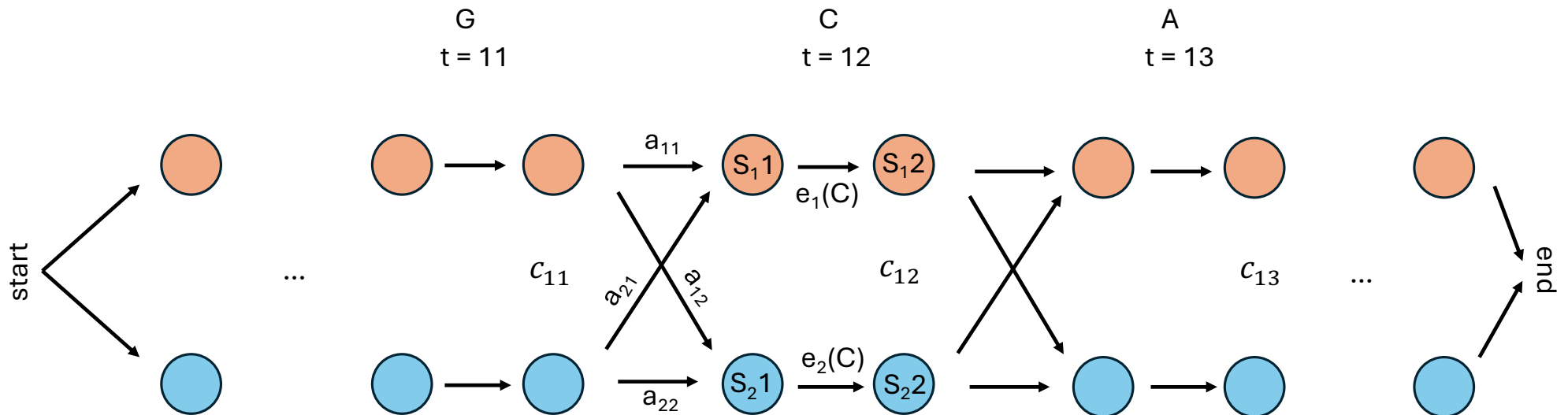
HW8 Overview



how do we update transition edges?

$$F(S_{12}(t=12)) * a_{12} * e_2("A") * B(S_{22}(t=13))$$

Scaling

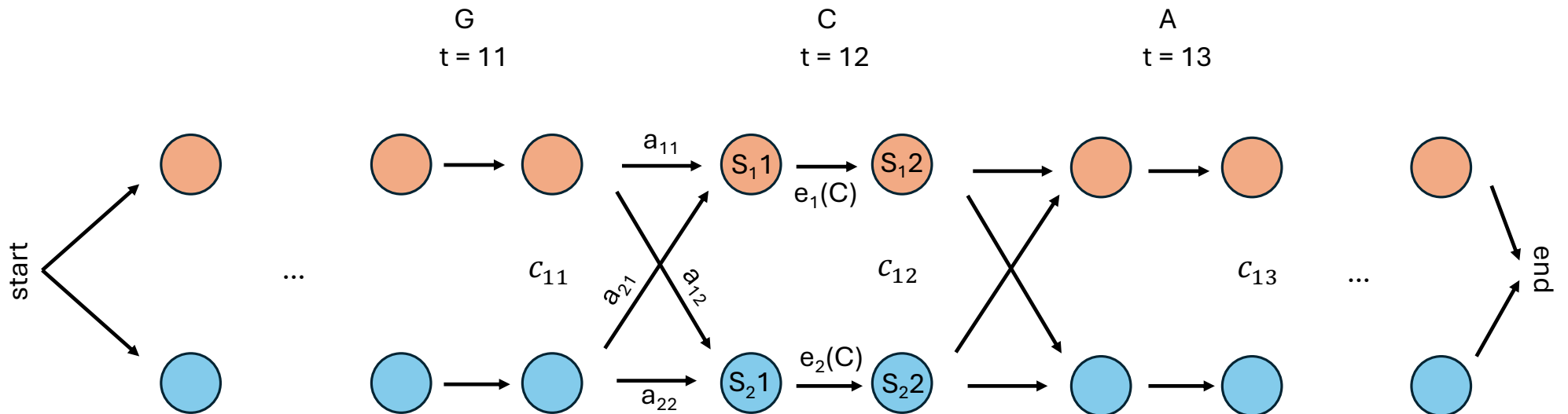


$$\hat{F}(S_{12}) = \frac{F(S_{12})}{F(S_{12}) + F(S_{22})}$$

$$\hat{F}(S_{22}) = \frac{F(S_{22})}{F(S_{12}) + F(S_{22})}$$

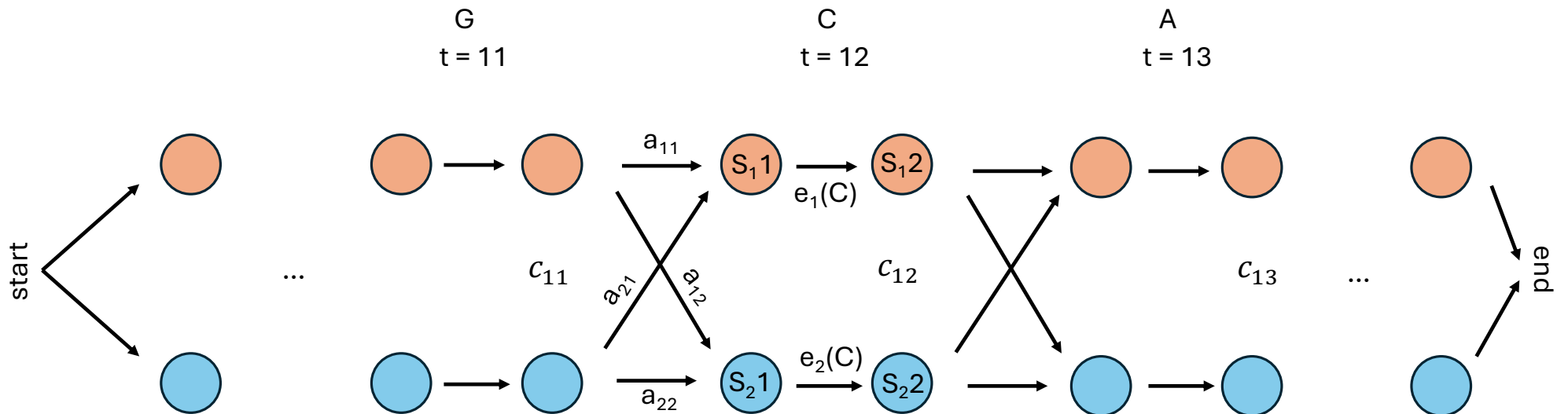
$$c_{t=12} = \frac{1}{F(S_{12}(t=12)) + F(S_{22}(t=12))}$$

Scaling



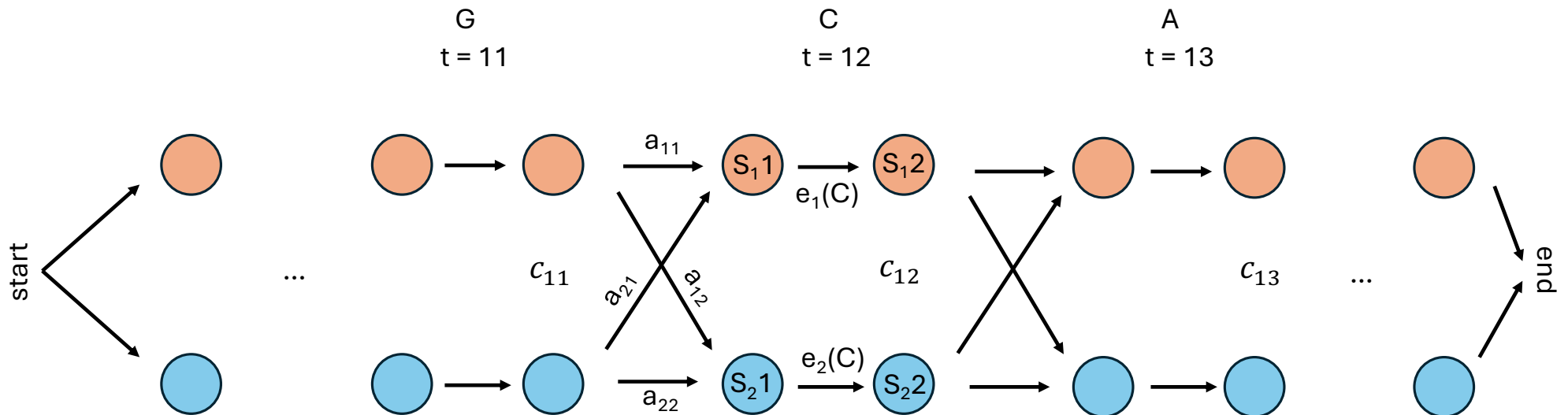
$$\hat{F}(S_{12}(t)) = \left(\prod_{\tau=1}^t c_{\tau} \right) * F(S_{12}(t))$$

Scaling



$$c_{t=12} = \frac{1}{F(S_{12}(t=12)) + F(S_{22}(t=12))} \quad \hat{B}(S_{12}(t)) = B(S_{12}(t)) * c_t$$

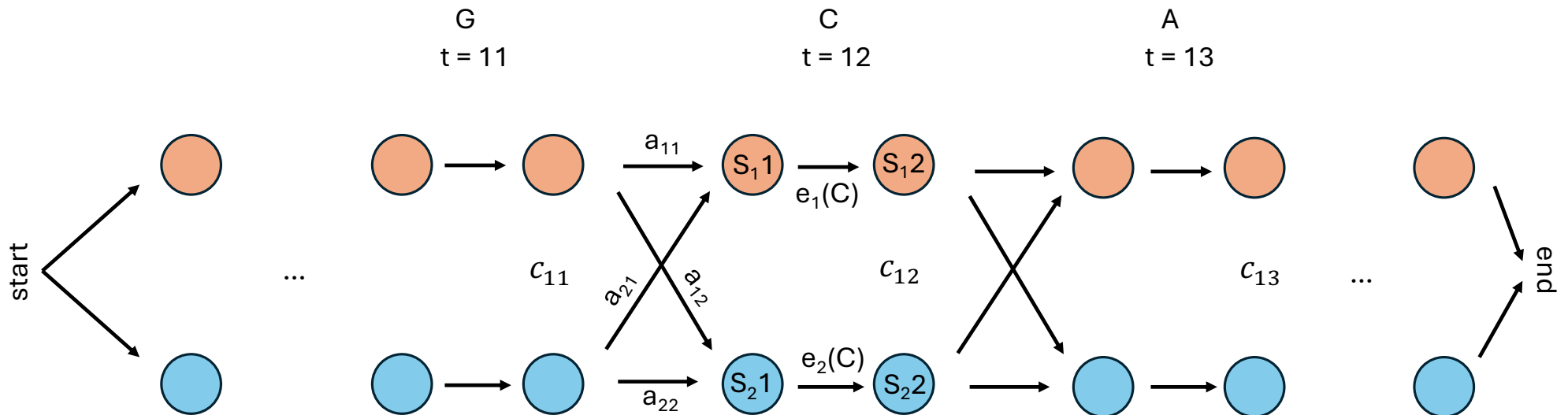
Scaling



updated probability
of emitting a "A" from
state 1

$$\overline{e_1("A")} = \frac{\sum_{t=1, n="A"}^N \frac{\hat{F}(S_12(t)) * \hat{B}(S_12(t))}{c_t}}{\sum_{t=1}^N \frac{\hat{F}(S_12(t)) * \hat{B}(S_12(t))}{c_t}}$$

Scaling



$$\overline{a_{12}} = \frac{\sum_{t=1}^{N-1} \hat{F}(S_1 2(t)) * a_{12} * e_2(seq[t]) * \hat{B}(S_2 2(t+1))}{\sum_{t=1}^{N-1} \frac{\hat{F}(S_1 2(t)) * \hat{B}(S_1 2(t))}{c_t}}$$