universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Alternative Evolutionary Processes and Measures of Phylogenetic Information“

verfasst von / submitted by

Cassius Manuel Pérez de los Cobos Hermosa

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 794 620 490
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Molekulare Biologie / Molecular Biosciences
Betreut von / Supervisor:	Univ.-Prof. Dr. Arndt von Haeseler

This page intentionally left blank.

Geduld,
wenn mich falsche Zungen stechen.

Matthäus-Passion (BWV 244), 35

Acknowledgements

Along my research I encountered many obstacles that made my days unfruitful. I sincerely thank Arndt for giving me time and trust in those frustrating periods.

In general, the kind support of the CIBIV group has been fundamental for my scientific education. In a vague chronological order, I explicitly thank Minh, Olga, Heiko, Christiane and Carme. More implicitly, I am very grateful to the colleagues and staff who made this work possible.

Abstract

Phylogenetic reconstruction requires making assumptions about the evolutionary process underwent by the observed sequences. Since model misspecification can impede a correct reconstruction, it is important to choose a realistic model and test its adequacy.

In the first part of this thesis, we analyze the common assumption that all sequences are possible along evolution. This assumption may be violated due to restriction enzymes that cleave DNA at specific recognition sites, motivating our description of the set of strings over a finite alphabet with taboos, that is, with prohibited substrings. We consider the Hamming graph whose vertices are taboo-free strings, and whose edges connect any two strings differing at a single site. Any walk on this graph describes the evolution of a taboo-free sequence. We characterize when the taboo-free Hamming graph and its suffix subgraphs are connected, concluding that the existence of disconnected evolutionary paths in nature is possible, although unlikely.

The second part of this thesis proposes new measures of phylogenetic information to assess the reliability of conclusions drawn from phylogenetic inference. These measures are the coherence of a branch, quantifying the dependence between two adjacent clades, and the memory of a clade, which quantifies the identification of the parent node of a clade. We explain the relationship of these measures with the underlying tree structure of the phylogeny, and then apply them to describe two problems of phylogenetics. First, we use the coherence to construct a powerful test for saturation along a branch of a phylogeny. Secondly, the memory is used to bound the information flow from children to parent node on a d-ary tree during the reconstruction of the root identity.

Kurzfasssung

Die phylogenetische Rekonstruktion erfordert Annahmen über den evolutionären Prozess, den die beobachteten Sequenzen durchlaufen haben. Da ein falsches Modell eine korrekte Rekonstruktion verhindern kann, sind die Auswahl eines realistischen Modells und die Prüfung seiner Eignung wichtige Schritte.

Im ersten Teil dieser Arbeit analysieren wir die übliche Annahme, dass alle Sequenzen entlang der Evolution möglich sind. Diese Annahme kann durch Restriktionsenzyme verletzt werden, die die DNA an bestimmten Erkennungsstellen schneiden. Dies motiviert unsere Beschreibung der Menge von Zeichenketten mit Tabus, d.h. mit verbotenen Teilketten. Wir beschreiben den Hamming-Graphen, dessen Knoten tabufreie Zeichenketten sind und dessen Kanten jede zwei Zeichenketten verbinden, die sich an einer einzigen Stelle unterscheiden. Jede Irrfahrt auf diesem Graphen räpresentiert die Entwicklung einer tabufreien Sequenz. Wir charakterisieren, wann der tabufreie Hamming-Graph und seine Suffix-Teilgraphen zusammenhängend sind. Unser Schluss ist, dass die Existenz von unverbundenen Evolutionspfaden in der Natur möglich, wenn auch unwahrscheinlich ist.

Im zweiten Teil dieser Arbeit werden neue Maße der phylogenetische Informationen vorgeschlagen, um die Zuverlässigkeit eines bestimmten Evolutionsprozesses zu bewerten. Diese Maße sind die Kohärenz eines Astes, die die Abhängigkeit zwischen zwei benachbarten Gruppen quantifiziert, und das Gedächtnis einer Gruppe, das die Identifizierung des Elternknotens einer Gruppe quantifiziert. Wir zeigen die Beziehung zwischen diesen Maßen und der latenten Baumstruktur der Phylogenie. Dann wenden wir diese Maße an, um zwei phylogenetische Probleme zu beschreiben. Erstens verwenden wir die Kohärenz, um einen trennschärfen Test auf Sättigung entlang eines Astes einer Phylogenie zu konstruieren. Zweitens wird das Gedächtnis verwendet, um den Informationsfluss von Kindernknoten zu Elternknoten in einem d-Weg Baum während der Rekonstruktion der Stammidentität zu begrenzen.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The reconstruction of past events given some observed data is a familiar task. If a friend arrives to a meeting with some delay and wet clothes, we may infer that the rain caused both the delay and the wet clothes. This inference becomes superfluous if, for example, our friend explains that the delay was due to some urgent and unexpected duty.

In phylogenetics, we aim to reconstruct the evolutionary history of organisms or species. In modern times, the observed data are either nucleotide or amino acid sequences of currently living organisms. There are immensely many ways to explain these observed data, and therefore we are obliged to make restrictive assumptions about the possible reconstructions. As an example, we will always assume that all nucleotide mutations are substitutions of one nucleotide by another, even though insertions and deletions of nucleotides are not rare.

The different topics of this work share the focus on assessing whether a phylogenetic assumption is sensible or not. We could imagine a practitioner who, before or after the reconstruction protocol, is unsure about the truth of the assumptions made.

In Chapter 2, we study the assumption that all sequences are possible during evolution. First we note that some prokaryotes have restriction enzymes which forbid some substrings in their genomes, which we call taboos. We model the evolution of a sequence affected by taboos as a path on a graph. The nodes of this graph are the allowed sequences, while edges represent the substitution of one nucleotide by another. Interestingly, we can construct examples where just a few taboos suffice to disconnect a taboo-free graph, questioning the general assumption that all allowed sequences can be reached along evolution. The main purpose of Chapter 2 is characterizing the connectivity of all taboo-free graphs.

Chapter 3 has an abstract origin. We start by proposing new measures of phylo-

genetic information, namely the memory of a clade and the coherence of a branch, arguing that they adequately describe the structure of the latent phylogenetic tree. To apply these measures, we formalize the concept of substitution saturation from a statistical point of view. In essence, saturation occurs if we cannot reject the null hypothesis that too many substitutions occurred as to provide any information. Armed with this theoretical framework, we use the coherence of a branch to test the assumption that two clades of a reconstructed tree have a detectable evolutionary history in common. Chapter 3 likely contains the most powerful results of this thesis.

In Chapter 4, we analyze the following problem: Given a phylogenetic tree, how much do we know about the identity of the ancestral sequences? We start by quantifying identification using the norm of the memory vector, which is a very similar measure to the memory of a clade studied in Chapter 3. Then we bound the flow of the "amount of identification" from observed to unobserved sequences following the Pruning algorithm to compute likelihoods in a phylogeny. Finally, using this upper bound, we give sufficient conditions under which we cannot identify the ancestral sequence at all, no matter how many descendant species are observed.

Chapter 2

Structure of the space of taboo-free sequences

Publication history and status

Submitted on October 29th 2019 to the Journal of Mathematical Biology, accepted for publication on August 19th 2020, published online on September 17th 2020.

C. Manuel and A. von Haeseler. Structure of the space of taboo-free sequences. *Journal of Mathematical Biology*, 81(4):1029–1057, 2020

Abstract

Models of sequence evolution typically assume that all sequences are possible. However, restriction enzymes that cut DNA at specific recognition sites provide an example where carrying a recognition site can be lethal. Motivated by this observation, we studied the set of strings over a finite alphabet with **taboos**, that is, with prohibited substrings. The taboo-set is referred to as $\mathbb{T}$ and any allowed string as a taboo-free string. We consider the so-called Hamming graph $\Gamma_n(\mathbb{T})$, whose vertices are taboo-free strings of length n and whose edges connect two taboo-free strings if their Hamming distance equals one. Any (random) walk on this graph describes the evolution of a DNA sequence that avoids taboos. We describe the construction of the vertex set of $\Gamma_n(\mathbb{T})$. Then we state conditions under which $\Gamma_n(\mathbb{T})$ and its suffix subgraphs are connected. Moreover, we provide an algorithm that determines if all these graphs are connected for an arbitrary $\mathbb{T}$.

As an application of the algorithm, we show that about 87% of bacteria listed in REBASE have a taboo-set that induces connected taboo-free Hamming graphs, because they have less than four type II restriction enzymes. On the other hand,

four properly chosen taboos are enough to disconnect one suffix subgraph, and consequently connectivity of taboo-free Hamming graphs could change depending on the composition of restriction sites.

2.1 Introduction

In bacteria, restriction enzymes cleave foreign DNA to stop its propagation. To do so, a double-stranded cut is induced by a so-called recognition site, a DNA sequence of length $4-8$ base pairs [Alberts et al., 2004]. As part of their restriction-modification (R-M) system, bacteria can escape the lethal effect of their own restriction enzymes by modifying recognition sites in their own DNA [Kommireddy and Nagaraja, 2013]. Nevertheless, Gelfand and Koonin [1997] and Rocha et al. [2001] found a significant avoidance of recognition sites in bacterial DNA, and Rusinov et al. [2015] showed that this avoidance was characteristic of type II R-M systems. Also in bacteriophages, the avoidance of the recognition sites is evolutionary advantageous [Rocha et al., 2001], mainly for non-temperate bacteriophages affected by orthodox type II R-M systems [Rusinov et al., 2018a]. Therefore in those instances the recognition site is, as we call it, a **taboo** for host and foreign DNA.

Although avoidance of recognition sites is well studied, e.g. by Rusinov et al. [2018b], taboo free DNA evolution has not yet been modelled. To initiate models of sequence evolution with taboos, we studied the Hamming graph $\Gamma_n(\mathbb{T})$, whose vertices are strings of length n over a finite alphabet Σ not containing any taboos of the set $\mathbb{T}$ as subsequence. Two vertices of the Hamming graph are adjacent if the corresponding taboo-free strings have Hamming distance equal to one. In biological terms, the sequences differ by a single substitution.

We note that, for a binary alphabet $\Sigma = \{0, 1\}$ and taboo-set $\mathbb{T} = \{11\}$, the corresponding Hamming graphs $\Gamma_n(\mathbb{T})$ are known as Fibonacci cubes. Some properties of the Fibonacci cubes like the Wiener Index or the degree distribution were surveyed by Klavžar [2013]. Further results have been obtained for taboo-sets forbidding arbitrary numbers of consecutive "1"s, $\mathbb{T} = \{1 \dots 1\}$, by Hsu and Chung [1993], or when $\mathbb{T} = \{s\}$ for an arbitrary binary string s by Ilić et al. [2012]. Recently, the equivalent problem of lattice paths that avoid some patterns has been described using automata and generating functions by Asinowski et al. [2018, 2019].

We are not so much interested in enumerative properties of Hamming graphs. We want to define conditions under which the Hamming graphs stay connected for arbitrary finite alphabets and arbitrary finite taboo-sets. From an evolutionary point

of view, connectivity guarantees that any taboo-free sequence can be generated by point mutations from any initial taboo-free sequence without containing a taboo-string during evolution. To include further biological realism, we will also study the connectivity of subgraphs $\Gamma_n^s(\mathbb{T})$ of the Hamming graph, where s is a taboo-free suffix. Suffix s can be viewed as a conserved DNA fragment, that is, a sequence that remained invariable during evolution [Shoemaker and Fitch, 1989, Fitch and Margoliash, 1967].

The inclusion of Hamming graphs with a constant suffix provides more general results, because $\Gamma_n^e(\mathbb{T}) = \Gamma_n(\mathbb{T})$, where e is the empty string. Given a taboo-set $\mathbb{T}$, if for every taboo-free string s and integer n the Hamming graph $\Gamma_n^s(\mathbb{T})$ is connected, then evolution can explore the space of taboo-free sequences by simple point mutation, no matter which DNA suffix fragments remain invariable, as long as the taboo-set $\mathbb{T}$ does not change in the course of evolution.

2.2 Motivating examples and non-technical presentation of key results

Here, we give a non-technical description of the essential results to determine connectivity. The subsequent sections provide a more technical and precise description of the central results.

Consider an alphabet Σ, for example $\Sigma = \{0, 1\}$. In a **Hamming graph of length** n, all possible words of length n are vertices, and two of these vertices are joined by an edge if they differ in exactly one position. A taboo-set is a set of forbidden subwords, such as $\mathbb{T} = \{11, 000\}$. Then, to construct a **taboo-free Hamming graph** $\Gamma_n(\mathbb{T})$, we simply have to erase all words of the Hamming graph of length n containing those taboos. Fig. 2.1 provides an example where $\Gamma_n(\mathbb{T})$ is disconnected for $n \geq 3$.

Given some alphabet and some taboo-set, deciding whether graph $\Gamma_n(\mathbb{T})$ is connected is not a trivial task. To see this, consider the four-nucleotide alphabet $\Sigma = \{A, C, G, T\}$, which is our main object of interest. Figure 2.2 shows the connected graph $\Gamma_3(\mathbb{T})$ for taboo-set $\mathbb{T} = \{AA, AC, AG, CA, CC, CG, GA, GC, GG\}$. The word TTT is a cut vertex, meaning that taboo-set $\mathbb{T}^* = \mathbb{T} \bigcup \{TTT\}$ yields the disconnected graph $\Gamma_3(\mathbb{T}^*)$.

Since the addition or deletion of one single taboo can have such an impact on connectivity, we need a tool to determine the structure of the taboo-free Hamming

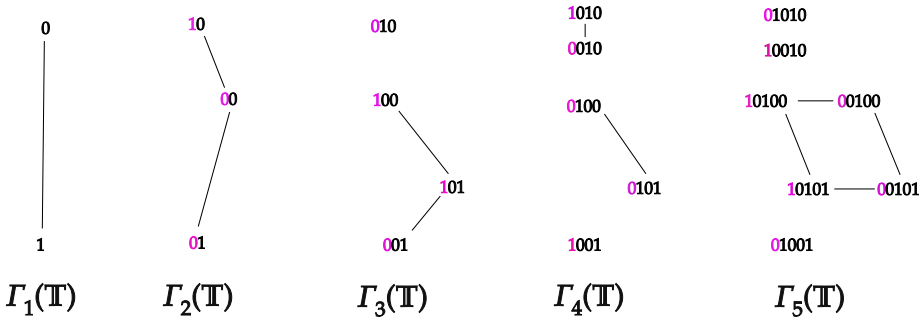

Figure 2.1: Graph $\Gamma_n(\mathbb{T})$ for $n \in [1,5]$ for binary alphabet $\Sigma = \{0,1\}$ and $\mathbb{T} = \{11,\ 000\}$. Set $V_{n+1}(\mathbb{T})$ is constructed by adding every allowed letter at the beginning of each string in $V_n(\mathbb{T})$.

graphs. This tool is described in full generality at the end of Section 2.8. In the particular case when $\Sigma = \{A, C, G, T\}$, our results can be simplified as follows.

1) If the number of taboos is smaller than the size of the alphabet, that is if $|\mathbb{T}| < 4$, then all graphs $\Gamma_n^s(\mathbb{T})$ are connected (Corollary 2.25.b). For example, given $\mathbb{T} = \{AATT, CCGG\}$, all taboo-free Hamming graphs are connected.

   Similarly, if the size of the set of all starting letters of taboos is smaller than the size of the alphabet, then all taboo-free Hamming graphs are connected (Corollary 2.25.a). This applies for taboo-set $\mathbb{T} = \{AA, AC, AG, CA, CC, CG, GA, GC, GG\}$, because the set of initial letters is $\{A, C, G\}$ and $|\{A, C, G\}| = 3 < 4$.

2) Prop. 2.24 describes a slightly more complex sufficient condition to determine connectivity. Given $\mathbb{T}$, delete the first letter of each taboo to construct the set $\Psi(\mathbb{T})$. For example, if $\mathbb{T} = \{AAA, CCA, GGA, TTT\}$, then $\Psi(\mathbb{T}) = \{AA, CA, GA, TT\}$.

   In set $\Psi(\mathbb{T})$, consider every pair of strings with Hamming distances 1 or 0. For example, the pair (AA, AA) has distance 0; the pair (AA, CA) has distance 1; and the pair (AA, TT) has distance 2. If every pair with Hamming distance 1 or 0 can be taboo-free extended to the left by the same letter, then all graphs $\Gamma_n^s(\mathbb{T})$ are connected.

   For example, the pair (AA, AA) can be extended by C, because CAA is taboo-free, and the pair (AA, CA) can be extended by T, because TAA and TCA are taboo-free. After checking all possible pairs with Hamming distance 0 or 1, we see that all such pairs in $\Psi(\mathbb{T})$ are extendable to the left, and thus taboo-set $\mathbb{T}$ generates connected taboo-free Hamming graphs.

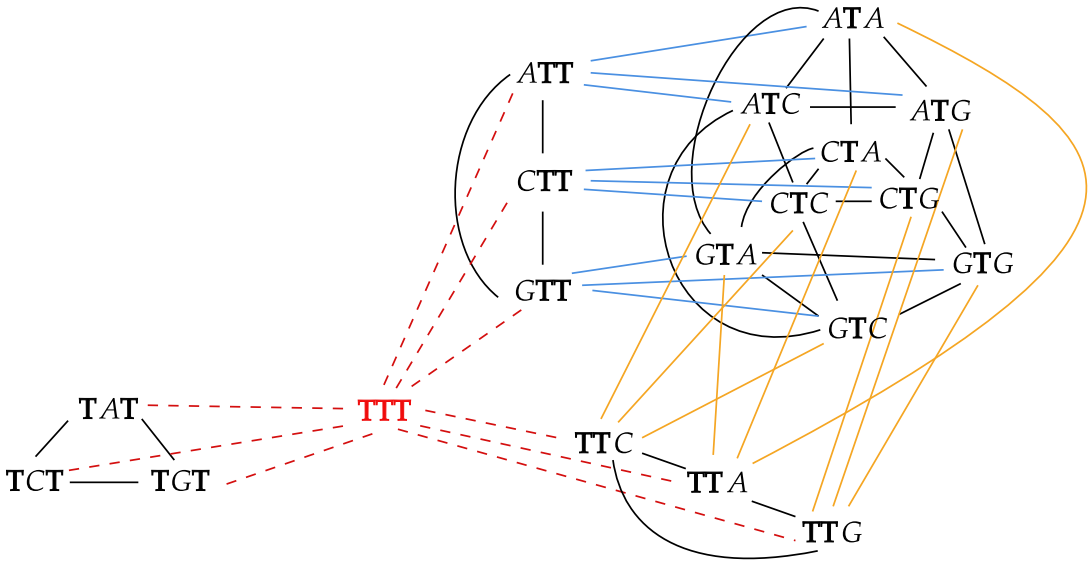

Figure 2.2: Graph $\Gamma_3(\mathbb{T})$, where $\Sigma = \{A, C, G, T\}$ and $\mathbb{T} = \{AA, AC, AG, CA, CC, CG, GA, GC, GG\}$. Vertex TTT is a cut vertex, because if we remove TTT and its incident edges (dashed lines, coloured red), then the resulting graph is disconnected. Consequently, graph $\Gamma_3(\mathbb{T}^*)$ induced by taboo-set $\mathbb{T}^* = \mathbb{T} \bigcup \{TTT\}$ is disconnected. Red, blue and yellow edges connect vertices with a different distribution of letter T.

3) If Prop. 2.24 cannot be applied, then we apply the characterization of Theorem 2.22. Assume for example that $\mathbb{T} = \{AAA, CCA, TAA, GAA\}$. Since the pair $\{AA, CA\} \subset \Psi(\mathbb{T})$ with Hamming distance one is not taboo-free extendable to the left by any letter, we proceed as follows. First we construct $\text{suf}(\mathbb{T})$, the set of all proper suffixes of $\mathbb{T}$. In our example, $\text{suf}(\mathbb{T}) = \{AA, CA, A, e\}$, where e is the string with no letters. Now we consider, for every suffix $r \in \text{suf}(\mathbb{T})$ the graph $\Gamma^r_{|r|+M}(\mathbb{T})$, where $|r|$ is the length of r and M is the length of the longest taboo(s) in $\mathbb{T}$. If all graphs $\Gamma^r_{|r|+M}(\mathbb{T})$ are connected, then every graph $\Gamma^s_n(\mathbb{T})$ is connected. In our example, graphs $\Gamma^{AA}_5(\mathbb{T})$, $\Gamma^{CA}_5(\mathbb{T})$, $\Gamma^A_4(\mathbb{T})$ and $\Gamma_3(\mathbb{T})$ are connected, implying that all taboo-free Hamming graphs are connected.

   When graph $\Gamma^r_{|r|+M}(\mathbb{T})$ is disconnected for some $r \in \text{suf}(\mathbb{T})$, then suffix r induces disconnected taboo-free Hamming graphs of the form $\Gamma^r_n(\mathbb{T})$ for $n \geq |r| + M$. Therefore evolution cannot explore the whole space of taboo-free sequences. This is the case for taboo-set $\mathbb{T}^*$ of Figure 2.2, where $r = e$ yields the disconnected graph $\Gamma_3(\mathbb{T}^*)$.

2.3 Outline

We will characterize taboo-sets $\mathbb{T}$ such that every Hamming graph of the form $\Gamma_n^s(\mathbb{T})$ is connected. To this end, we describe in Section 2.5 basic properties of taboo-sets . In Section 2.6, we introduce a very general type of taboo-sets, called **left proper** (Def. 2.4), which are our main object of study. In Prop. 2.11.b we show that, to construct graph $\Gamma_n^s(\mathbb{T})$, we only need the longest prefix of s which is a suffix of a taboo, which we call $s[1, k_s]$. In Section 2.7 we state the graph isomorphism $\Gamma_n^s(\mathbb{T}) \simeq \Gamma_n^{s[1,k_s]}(\mathbb{T})$ (Theorem 2.16). In Section 2.8 we explain how the edges of a quotient graph are related to the structure of graph $\Gamma_n^n(\mathbb{T})$ (Prop. 2.17).

Combining all these results, in Section 2.8 we characterize the connectivity of Hamming graphs $\Gamma_n^s(\mathbb{T})$. We prove by induction that the connectivity of a small number of quotient graphs implies the connectivity of all Hamming graphs with long suffixes (Prop. 2.20). This result can be used to prove connectivity of Hamming graphs with short suffixes (Prop. 2.21). These two results yield the characterization of the connectivity of every suffix Hamming graph in Theorem 2.22. Section 2.9 provides examples of bacterial taboo-sets and their connectivity.

2.4 Basic notations

We will introduce some standard notations concerning strings as well as some relevant terms from graph theory.

2.4.1 Strings

We will use the term **string** to refer to a sequence of symbols over an arbitrary finite alphabet $\Sigma = \{a_1, \cdots, a_m\}$, where $m \geq 2$, while **(DNA) sequence** is reserved for biological contexts, where the alphabet consists of the four nucleotides $\Sigma = \{A, C, G, T\}$.

We denote the set of strings of length n over the alphabet Σ by Σ^n. The length of a string s is denoted by $|s|$. The empty string will be denoted by e, and satisfies $|e| = 0$ and $\{e\} = \Sigma^0$.

Given a string $s = b_1 \cdots b_n \in \Sigma^n$, the expression

$$s[i, j] := \begin{cases} b_i \cdots b_j & \text{if } 1 \leq i \leq j \leq n \\ e & \text{otherwise} \end{cases}$$

denotes the **substring** of s starting at the i-th position and ending at the j-th

position, and e when this substring is not well-defined (for example if $j = 0$). In particular $s[1, j]$ is a **prefix** of s that ends at position j and $s[i, n]$ is a **suffix** of s that starts at position i. A substring, prefix or suffix is called **proper** if it is not the entire string s. For a set of strings S, we define **the substrings from the ith to the jth position of** S as

$$S[i, j] := \{s[i, j] \mid s \in S\}.$$

We also need **the set of proper suffixes of** S, defined as

$$\mathrm{suf}(S) := \Big(\bigcup_{s\in S}\bigcup_{i\in[2,|s|]} s[i, |s|]\Big)\bigcup\{e\}.$$

Where $i \in [2, |s|]$ refers to all integers i within the interval $[2, |s|]$. It should not be confused with substring $s[2, |s|]$ of s.

Example 2.1. *If* $S = \{ACG,\ GGG,\ TTC,\ CC\}$ *then*

$$\mathrm{suf}(S) = \{CG,\ G,\ GG,\ TC,\ C,\ e\}.$$

If string s_1 is substring of string s_2, we write $s_1 \prec s_2$, while $s_1 \not\prec s_2$ denotes that s_1 is **not** a substring of s_2. By convention, $e \prec s$ for any string s. For strings s_1 and s_2, we define $s_1 s_2$ as the **concatenation** of s_1 and s_2. Note that $es = se = s$ for any s. For a string s and a set of strings $S = \{s_1, \cdots s_k\}$, the concatenation of s with all elements in S is denoted by $s \circ S := \{ss_1, \cdots ss_k\}$. If S_1 and S_2 are disjoint sets, then the disjoint union of S_1 and S_2 will be denoted by $S_1 \bigsqcup S_2$.

Finally, given two strings s_1, s_2 of equal length, $d(s_1, s_2)$ denotes their **Hamming distance**, that is, the number of positions at which the corresponding symbols differ.

2.4.2 Graph theory

We will use common graph theory terminology following Wilson [1986]. Let $\mathcal{G} = (V, E)$ denote a simple, undirected graph with vertex set V and edge set E. We say that graph $\mathcal{G}_1 = (V_1, E_1)$ is **subgraph** of $\mathcal{G}_2 = (V_2, E_2)$ if $V_1 \subseteq V_2$ and $E_1 \subseteq E_2$, and we denote this as $\mathcal{G}_1 \subseteq \mathcal{G}_2$.

Given a graph $\mathcal{G} = (V, E)$ and a subset $V_1 \subseteq V$, then the **subgraph induced by V_1 in $\mathcal{G}$**, $\mathcal{G}(V_1) = (V_1, E_{V_1})$, has vertex set V_1 and, for any $u, v \in V_1$, $\{u, v\} \in E_{V_1}$ iff $\{u, v\} \in E$.

Two graphs $\mathcal{G}_1 = (V_1, E_1)$ and $\mathcal{G}_2 = (V_2, E_2)$ are **isomorphic**, denoted by $\mathcal{G}_1 \simeq$

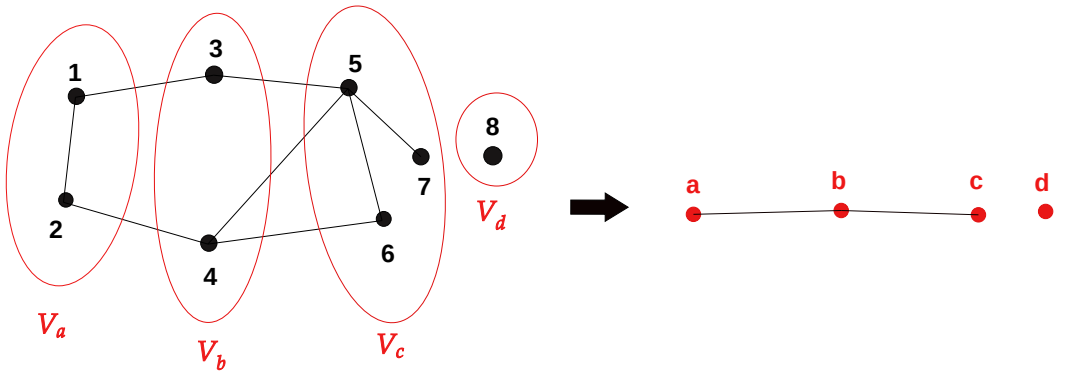

Figure 2.3: Example of a quotient graph. For $\mathcal{G} = (V, E)$ on the left hand side, with $V = \{1, 2, 3, 4, 5, 6, 7, 8\}$ and partition $V = V_a \bigsqcup V_b \bigsqcup V_c \bigsqcup V_d$, we obtain the quotient graph $\mathcal{Q}[\mathcal{G}]$ on the right hand side.

$\mathcal{G}_2$, if there exists a bijection $f : V_1 \to V_2$ such that, for every $u, v \in V_1$, $\{u, v\} \in E_1$ iff $\{f(u), f(v)\} \in E_2$. That is, $\mathcal{G}_1$ and $\mathcal{G}_2$ are isomorphic if there exists an edge-preserving bijection between their vertex sets.

We will also need the **quotient graph**, as defined by Sanders and Schulz [2013], to study the connectivity of Hamming graphs. To define it, consider a graph $\mathcal{G} = (V, E)$ and a partition of its vertex set V, namely $V = \bigsqcup_{b \in J} V_b$ for some index set J. The **quotient graph of** $\mathcal{G}$, denoted as $\mathcal{Q}[\mathcal{G}] = (J, E_J)$, is the graph whose vertices are J and such that $\{b_1, b_2\} \in E_J$ iff an edge connects a vertex in V_{b_1} with a vertex in V_{b_2}. Figure 2.3 gives an example of a quotient graph.

Our strategy to prove connectivity of taboo-free Hamming graphs will use the following propositions, whose proof is simple enough to be omitted

Proposition 2.1. *Consider graph $\mathcal{G} = (V, E)$ and partition $V = \bigsqcup_{b \in J} V_b$.*

If every induced subgraph $\mathcal{G}(V_b)$ for $b \in J$ is connected and the quotient graph $\mathcal{Q}[\mathcal{G}]$ is connected, then $\mathcal{G}$ is connected.

Proposition 2.2. *For graph $\mathcal{G} = (V, E)$, the following statements are equivalent:*

- *$\mathcal{G}$ is connected.*
- *For every partition of V, the quotient graph $\mathcal{Q}[\mathcal{G}]$ is connected.*

2.5 Properties of taboo-sets

We will repetadly use of the following terminology.

Definition 2.1.

- *A finite set of strings* $\mathbb{T}$ *such that every* $t \in \mathbb{T}$ *satisfies* $|t| \geq 2$ *is called a* **taboo-set**.
- *Strings in* $\mathbb{T}$ *are called* **taboos**.
- *The* **length of the longest taboo(s)** *in* $\mathbb{T}$ *will be denoted by* $M := \max\{|t|\}_{t\in\mathbb{T}}$.
- *A string is* **taboo-free** *if it does not contain any taboo of* $\mathbb{T}$ *as substring.*
- $V_n(\mathbb{T})$ *denotes the* **set of taboo-free strings of length** n.
- $V_n^s(\mathbb{T})$ *denotes the* ***set of strings in*** $V_n(\mathbb{T})$ ***with suffix*** s.
- *Similarly,* ${}^sV_n(\mathbb{T})$ *denotes all strings in* $V_n(\mathbb{T})$ *with prefix* s.

With Definition 2.1 in mind, we can prove some simple properties of taboo-sets.

Proposition 2.3. *Given taboo-sets* $\mathbb{T}_1$ *and* $\mathbb{T}_2$*, it holds that:*

a) Set $\mathbb{T}_1 \bigcup \mathbb{T}_2$ *is a taboo-set*

b) For $n \in \mathbb{N}$, $V_n(\mathbb{T}_1) \bigcap V_n(\mathbb{T}_2) = V_n(\mathbb{T}_1 \bigcup \mathbb{T}_2)$.

c) If for every $t_1 \in \mathbb{T}_1$ *there exists* $t_2 \in \mathbb{T}_2$ *such that* $t_2 \prec t_1$*, then for any* $n \in \mathbb{N}$, $V_n(\mathbb{T}_2) \subseteq V_n(\mathbb{T}_1)$.

Proof.

a) Every $t \in \mathbb{T}_1 \bigcup \mathbb{T}_2$ has length at least 2, and thus $\mathbb{T}_1 \bigcup \mathbb{T}_2$ is a taboo-set.

b) All strings $s \in V_n(\mathbb{T}_1) \bigcap V_n(\mathbb{T}_2)$ satisfy $t_1 \not\prec s$ for all $t_1 \in \mathbb{T}_1$ and $t_2 \not\prec s$ for all $t_2 \in \mathbb{T}_2$ this is equivalent to s satisfying $t \not\prec s$ for all $t \in \mathbb{T}_1 \bigcup \mathbb{T}_2$.

c) Consider $s \in V_n(\mathbb{T}_2)$. Assume that $s \notin V_n(\mathbb{T}_1)$; then there exists $t_1 \in \mathbb{T}_1$ such that $t_1 \prec s$. But there also exists a $t_2 \in \mathbb{T}_2$ such that $t_2 \prec t_1$, and thus $t_2 \prec s$, a contradiction. Hence $s \in V_n(\mathbb{T}_1)$. ■

□

For a given n and $\mathbb{T}$, we can find a taboo-set $\mathbb{T}' \neq \mathbb{T}$ such that $V_n(\mathbb{T}) = V_n(\mathbb{T}')$. In this sense, taboo-sets are not unique, as we illustrate in the following proposition.

Proposition 2.4. *For a string* t *and* $n \geq |t| + 1$*, it holds that*

$$V_n(\{t\}) = V_n\Big((t \circ \Sigma) \bigcup (\Sigma \circ t)\Big).$$

Proof.

- $\subseteq$: Any taboo in $\mathbb{T}_1 := (t \circ \Sigma) \bigcup (\Sigma \circ t)$ has $t \in \mathbb{T}_2 := \{t\}$ as substring, and thus Prop. 2.3.*c* implies $V_n(\{t\}) \subseteq V_n((t \circ \Sigma) \bigcup (\Sigma \circ t))$.

- $\supseteq$: Assume that there exists an $s \in V_n((t \circ \Sigma) \bigcup (\Sigma \circ t))$ with $t \prec s$. Since $|s| = n$ and $n \geq |t| + 1$, the substring t is either preceded or followed by some symbol $a \in \Sigma$. This contradicts $\{at, ta\} \subseteq (t \circ \Sigma) \cup (\Sigma \circ t)$. ■

□

Prop. 2.4 implies that, for any $\mathbb{T}$, we can construct many taboo-sets $\mathbb{T}'$ such that $V_n(\mathbb{T}) = V_n(\mathbb{T}')$ as long as $n \geq \max(M, M')$, where M and M' denote the length of the longest taboo in $\mathbb{T}$ and $\mathbb{T}'$, respectively.

Example 2.2. *If* $\mathbb{T} = \mathbb{T}_1 \bigsqcup \mathbb{T}_2$ *with* $\mathbb{T}_2 = (t \circ \Sigma) \cup (\Sigma \circ t)$*, Prop. 2.3.a and 2.4 imply that* $\mathbb{T}' := \mathbb{T}_1 \bigsqcup \{t\}$ *satisfies* $V_n(\mathbb{T}) = V_n(\mathbb{T}')$ *for any* $n \geq M$*. Repeating this process, we can construct a taboo-set* $\mathbb{T}'$ *such that* $(t \circ \Sigma) \bigcup (\Sigma \circ t) \not\subseteq \mathbb{T}'$ *for any string* t *and satisfying* $V_n(\mathbb{T}) = V_n(\mathbb{T}')$ *for any* $n \geq M$*.*

Example 2.2 and Prop. 2.4 motivate the following definition.

Definition 2.2. *A taboo-set* $\mathbb{T}$ *is* ***minimal*** *if the following conditions hold:*

a For every different $t_1, t_2 \in \mathbb{T}$*, it holds that* $t_1 \not\prec t_2$*.*

b For every $j \in [0, M-1]$ *and* $s \in V_j(\mathbb{T})$*, set* $(s \circ \Sigma) \bigcup (\Sigma \circ s)$ *is not a subset of* $\mathbb{T}$*.*

Condition *a*) is easy to justify: If string AA is a taboo, it is redundant that AAA be a taboo. Condition *b*) avoids unnecessarily complicated taboo-sets. For example, using the four-nucleotide alphabet, taboo-set $\mathbb{T} = \{AAA, AAC, AAG, AAT, CAA, GAA, TAA\}$ can be minimized as $\mathbb{T}' = \{AA\}$. In general, one can minimize a taboo-set according to Example 2.2.

Since we want to study taboo-free strings of arbitrary lengths, we need conditions to concatenate taboo-free strings such that the concatenated sequence is taboo-free. The following result gives such a condition.

Proposition 2.5. *Given taboo-set* $\mathbb{T}$*, consider three strings* s_1, s_2, s_3 *such that* $s_1 s_2$ *and* $s_2 s_3$ *are taboo-free and* $|s_2] \geq M - 1$*. Then* $s := s_1 s_2 s_3$ *is taboo-free.*

Proof. If $|s_1| = 0$ and $|s_3| = 0$, then $s = s_2$ is taboo-free, as desired. Now assume either $|s_1| > 0$ or $|s_3| > 0$, yielding $n := |s_1| + |s_2| + |s_3| \geq M$. For each $i \in [1, n - (M + 1)]$, the fact that $|s_2| \geq M - 1$ implies that either $s[i, i + M - 1] \prec s_1 s_2$ or $s[i, i + M - 1] \prec s_2 s_3$, hence each $s[i, i + M - 1]$ is taboo-free and the result follows. ■

□

2.6 Prefixes and suffixes of a taboo-free string

Given a taboo-free string s, the construction of set $V_n^s(\mathbb{T})$ for $n > |s|$ depends on which string w can be concatenated to the left side of s, such that $ws \in V_n(\mathbb{T})$. This motivates the following definition.

Definition 2.3. *Given a taboo-set $\mathbb{T}$, consider a taboo-free string s and $k \in \mathbb{N}_0$. The k-**prefixes** of s are the elements of the set $L^k(s)$, defined as*

$$L^k(s) := \{w \in \Sigma^k \text{ such that } ws \text{ is taboo-free } \} = V^s_{|s|+k}(\mathbb{T})[1, k].$$

*If $L^k(s) \neq \emptyset$, then we will say that s is k-**prefixable**.*
*Similarly, the k-**suffixes** of s, denoted $R^k(s)$, are the strings $w \in \Sigma^k$ such that $sw \in V_{|s|+k}(\mathbb{T})$, that is, $R^k(s) := {}^sV_{|s|+k}(\mathbb{T})[|s| + 1, |s| + k]$. When $R^k(s) \neq \emptyset$, we say that s is k-**suffixable**.*

Example 2.3. *If $\Sigma = \{A, C, G, T\}$ and $\mathbb{T} = \{CAA, GAA, TAA\}$, then $L^1(AA) = \{A\}$ and $L^2(AA) = \{AA\}$. Hence string AA is 1-prefixable and 2-prefixable. Moreover, $R^1(AA) = \{A, C, G, T\}$, hence string AA is 1-suffixable.*

By construction, given $s \in V_{|s|}(\mathbb{T})$, for any $k \in \mathbb{N}_0$ it holds that

$$V^s_{k+|s|}(\mathbb{T}) = L^k(s) \circ s. \tag{2.1}$$

That is, $V^s_{k+|s|}(\mathbb{T})$ is $L^k(s)$ with s concatenated. Moreover, the following proposition shows that the k-prefixes of a string s induce a disjoint partition of the set $V_n^s(\mathbb{T})$.

Proposition 2.6. *Given a taboo-set $\mathbb{T}$ and a taboo-free string s, consider integers $k \in \mathbb{N}_0$ and $n \geq k + |s|$. It holds that*

$$V_n^s(\mathbb{T}) = \bigsqcup_{w \in L^k(s)} V_n^{ws}(\mathbb{T}).$$

That is, the set $V_n^s(\mathbb{T})$ can be partitioned into the disjoint sets of taboo-free strings of length n with suffix ws, where $w \in L^k(s)$.

Proof. If s is not k-prefixable, then $L^k(s) = \emptyset$ and $V_n^s(\mathbb{T}) = \emptyset$, hence the equation holds. Otherwise, the inclusion $\supseteq$ is clear, while the $\subseteq$ follows from the fact that, for any string $w \in \Sigma^k$ preceding the suffix s, this w must necessarily belong to $L^k(s)$. ■ □

Clearly, if a taboo-free string s is k^*-prefixable, then it is also k-prefixable for any integer $k < k^*$, while nothing can be said *a priori* about the case $k > k^*$. Consequently, we need to find conditions under which one can concatenate at least one symbol to the left of a taboo-free string. We will first introduce such taboo-sets in Def. 2.4 and then characterize prefixability in Prop. 2.7.

Definition 2.4. *A taboo-set $\mathbb{T}$ is called **left proper** if every $s \in V_M(\mathbb{T})$ is 1-prefixable. Analogously, $\mathbb{T}$ is **right proper** if every $s \in V_M(\mathbb{T})$ is 1-suffixable.*

Example 2.4. *If $\Sigma = \{A, C, G, T\}$ and $\mathbb{T} = \Sigma \circ A$, then $AC \in V_2(\mathbb{T})$ and AC is not 1-suffixable. Thus, $\mathbb{T}$ is not left proper.*

Proposition 2.7. *Consider a left proper taboo-set $\mathbb{T}$ and a taboo-free string s such that one of the following conditions holds:*

a) $|s| \geq M$

b) $|s| \leq M - 1$ and s is $(M - |s|)$-prefixable

Then s is k-prefixable for every $k \in \mathbb{N}$.

Proof. If condition a) applies, then the prefix $s[1, M] \in V_M(\mathbb{T})$ is 1-prefixable, because $\mathbb{T}$ is left proper. That is, there exists $a \in \Sigma$ with $as \in V_{1+|s|}(\mathbb{T})$. Proceeding analogously with $(as)[1, M]$, we infer that s is 2-prefixable. Continuing with this process, we deduce that s is k-prefixable for any $k \in \mathbb{N}$.

If condition b) holds, then we can take any string in $V_M^s(\mathbb{T})$ and proceed as we did assuming a). ■ □

We mainly study left proper taboo-sets due to Prop. 2.7, because the existence of arbitrary k-prefixes is necessary in many of our proofs. Analogous results for right proper taboo-sets are obtained by reversing the order of the symbols composing the string.

According to Prop. 2.7, if the length of a taboo-free string is at least M, then the taboo-free string can prefixed for arbitrary lengths. Otherwise, one needs to check the $(M - |s|)$-prefixability of this string. To that end, the following result comes in handy.

Proposition 2.8. *Consider a left proper taboo-set $\mathbb{T}$ and a taboo-free string s.*

a) If $|s| \leq M-1$ and $s \notin \mathrm{suf}(V_M(\mathbb{T}))$, then $V_n^s(\mathbb{T}) = \emptyset$ for $n \geq M$.

b) If either $|s| \geq M$ or $s \in \mathrm{suf}(V_M(\mathbb{T}))$, then $V_n^s(\mathbb{T}) \neq \emptyset$ for $n \geq \max(|s|, M)$.

Proof.

a) If $0 \leq |s| \leq M-1$ and $s \notin \mathrm{suf}(V_M(\mathbb{T}))$, since $\mathrm{suf}(V_M^s(\mathbb{T})) \subseteq \mathrm{suf}(V_M(\mathbb{T}))$, it holds that $V_M^s(\mathbb{T}) = \emptyset$. This implies that $V_n^s(\mathbb{T}) = \emptyset$ for every $n \geq M$, because otherwise

$$\emptyset \subsetneq V_n^s(\mathbb{T})[n-M+1, n] \subseteq V_M^s(\mathbb{T}),$$

which contradicts $V_M^s(\mathbb{T}) = \emptyset$.

b) If $|s| \geq M$, since $\mathbb{T}$ is left proper, Prop. 2.7.a implies that s is k-prefixable for every $k \in \mathbb{N}$. Thus, $V_n^s(\mathbb{T}) \neq \emptyset$. Similarly, if $s \in \mathrm{suf}(V_M(\mathbb{T}))$, then s is $(M-|s|)$-prefixable, and thus Prop. 2.7.b implies that s is k-prefixable for every $k \in \mathbb{N}$. ■

□

Note that, since the assumptions of Prop. 2.8.a are the negation of the assumptions of Prop. 2.8.b, in Prop. 2.8 we have proved that $V_n^s(\mathbb{T}) = \emptyset$ for $n \geq M$ **iff** string s satisfies $|s| \leq M-1$ and $s \notin \mathrm{suf}(V_M(\mathbb{T}))$.

To study the connectivity of Hamming graphs $\Gamma_n^s(\mathbb{T})$, we need to know whether two different strings have a k-prefix in common. Thus, we introduce the following.

Definition 2.5. *Given a taboo-set $\mathbb{T}$, we say that two taboo-free strings s_1 and s_2 (maybe of different length) are **left k-synchronized** if $L^k(s_1) \bigcap L^k(s_2) \neq \emptyset$. If $R^k(s) \bigcap R^k(r) \neq \emptyset$, then we say that s_1 and s_2 are **right k-synchronized**.*

In words, two taboo-free strings are left k-synchronized if they are k-prefixable by at least one string w. Clearly, two taboo-free strings s_1, s_2 that are left k^*-synchronized are also left k-synchronized for any $k \leq k^*$ (one simply has to "cut" the k symbols on the left of $L^{k^*}(s_1) \bigcap L^{k^*}(s_2)$). The following proposition states when we can also guarantee k-synchronization for $k > k^*$:

Proposition 2.9. *Consider a left proper taboo-set $\mathbb{T}$ and two taboo-free strings s_1, s_2, with length greater than zero, such that s_1 and s_2 are left $(M-1)$-synchronized. Then s_1 and s_2 are left k-synchronized for any $k \in \mathbb{N}$.*

Proof. If $k \leq M-1$, then the assertion is true since s_1 and s_2 are $(M-1)$-synchronized.

For $k > M-1$, consider a string $w \in L^{M-1}(s_1) \bigcap L^{M-1}(s_2)$. We know that ws_1 and ws_2 are taboo-free strings with length at least M. Since $\mathbb{T}$ is left proper, Prop. 2.7.a applied to ws_1 and ws_2 implies that ws_1 and ws_2 are k'-prefixable for any $k' \in \mathbb{N}$. Therefore w is k'-prefixable for any $k' \in \mathbb{N}$. For any k', take $x \in L^{k'}(w)$ and consider strings xws and xwr. The fact that $|w| = M-1$, together with the fact that xw and the pair ws_1, ws_2 are taboo-free, allows applying Prop. 2.5, hence xws_1 and xws_2 are also taboo-free.

It follows that $xw \in L^{M-1+k'}(s_1) \bigcap L^{M-1+k'}(s_2)$. With $k := M-1+k'$, the result follows for any $k > M-1$. ■ □

The following proposition provides a Hamming-distance based criterion to quickly decide whether two taboo-free strings of length M are left k-synchronized.

Proposition 2.10. *Consider a left proper taboo-set $\mathbb{T}$. If all pairs $s_1, s_2 \in V_M(\mathbb{T})$ with $d(s_1, s_2) = 1$ are left 1-synchronized, then all pairs $s_1, s_2 \in V_M(\mathbb{T})$ with $d(s_1, s_2) = 1$ are left k-synchronized for all $k \in \mathbb{N}_0$.*

Proof. Given any left 1-synchronized pair s_1, s_2 with $d(s_1, s_2) = 1$, there exists an $a \in \Sigma$ such that as_1 and as_2 are taboo-free. Since $(as_i)[1, M] \in V_M(\mathbb{T})$ for $i \in \{1, 2\}$ and the Hamming distance between these two strings is at most 1, as_1, as_2 are 1-synchronized, hence there exists a symbol $b \in \Sigma$ such that bas_1 and bas_2 are taboo-free, i.e. s_1 and s_2 are left 2-synchronized. Continuing with this process, it follows that s_1 and s_2 are k-synchronized. ■ □

We will now discuss conditions that allow increasing the string length of an entire set of taboo-free strings. To this end, consider two taboo-free strings s_1, s and the set $V^{s_1 s}_{n+|s_1|+|s|}(\mathbb{T})$. It is generally not true that $V^{s_1 s}_{n+|s_1|+|s|}(\mathbb{T}) = V^{s_1}_{n+|s_1|}(\mathbb{T}) \circ s$, because the concatenation of s to a taboo-free string from $V^{s_1}_{n+|s_1|}(\mathbb{T})$ can create a taboo string around the junction of both strings. For the remainder of this section we will discuss when the equality holds.

Definition 2.6. *For a taboo-set $\mathbb{T}$ and a taboo-free string s, we define the **length of the longest taboo suffix-prefix match** as*

$$k_s := \max\left\{ i \in [0, |s|] \;\middle|\; s[1, i] \in \mathrm{suf}(\mathbb{T}) \right\},$$

i.e. k_s denotes the length of the longest prefix of s being a proper suffix of a taboo.

Note that the length k_s is well defined, because $s[1,0] = e \in \mathrm{suf}(\mathbb{T})$, hence $k_s \in [0, \min(M-1, |s|)]$. Using this length k_s, in Prop. 2.11 we give conditions implying that equality $V^{s_1 s}_{n+|s_1|+|s|}(\mathbb{T}) = V^{s_1}_{n+|s_1|}(\mathbb{T}) \circ s$ holds.

Proposition 2.11. *For a taboo-set $\mathbb{T}$ and a taboo-free string s, the following holds:*

a) Take $w \in \Sigma^{M-1}$ such that $ws \in V_{M-1+|s|}(\mathbb{T})$. Then for any $n \geq M-1$,

$$V^{ws}_{n+|s|}(\mathbb{T}) = V^{w}_{n}(\mathbb{T}) \circ s.$$

b) For any $n \in \mathbb{N}_0$ it holds that

$$V^{s}_{n+|s|}(\mathbb{T}) = V^{s[1,k_s]}_{n+k_s}(\mathbb{T}) \circ s[k_s+1, |s|].$$

Proof.

a) The inclusion $\subseteq$ is clear. The inclusion $\supseteq$ follows from the fact that, if we are given $rw \in V^{w}_{n}(\mathbb{T})$ such that $ws \in V_{M-1+j}(\mathbb{T})$, since $|w| = M-1$, Prop. 2.5 yields that the concatenated string rws is taboo-free.

b) The result is obvious if $|s| = 0$ or $n = 0$, hence assume $|s| > 0$ and $n > 0$.

Clearly $V^{s}_{n+|s|}(\mathbb{T}) \subseteq V^{s[1,k_s]}_{n+k_s}(\mathbb{T}) \circ s[k_s+1, |s|]$. For $r \in V_n(\mathbb{T})$, consider $rs[1,k_s] \in V^{s[1,k_s]}_{n+k_s}(\mathbb{T})$. We need to prove that the string

$$rs[1,k_s]s[k_s+1,|s|] = rs$$

is taboo-free. But otherwise, since $rs[1,k_s]$ and s are taboo-free, there would exist integers c, d such that $1 \leq c \leq |r| \leq |r|+k_s < d \leq |r|+|s|$ and $(rs)[c,d] \in \mathbb{T}$. Take $k^* := d - |r| > k_s$, which yields $s[1,k^*] \in \mathrm{suf}(\mathbb{T})$, contradicting the maximality of k_s. Hence rs is taboo-free, as desired. Note that the same argument applies if $k_s = 0$. ■

□

From Prop. 2.11.b we obtain two corollaries.

Corollary 2.12. *Given a taboo-set $\mathbb{T}$ and a taboo-free string s, for any $k \in \mathbb{N}_0$ it holds that*

$$L^k(s) = L^k(s[1,k_s]).$$

Proof. By construction, $L^k(s) = V^s_{|s|+k}(\mathbb{T})[1,k]$. Prop. 2.11.b yields

$$V^s_{k+|s|}(\mathbb{T})[1,k] = \left(V^{s[1,k_s]}_{k+k_s}(\mathbb{T}) \circ s[k_s+1,|s|]\right)[1,k] =$$
$$= V^{s[1,k_s]}_{k+k_s}(\mathbb{T})[1,k] = L^k(s[1,k_s]). \blacksquare$$

□

Corollary 2.13. *For a taboo-set $\mathbb{T}$ and for any pair of taboo-free strings s_1 and s_2, the following statements are equivalent for all $k \in \mathbb{N}_0$:*

- *s_1 and s_2 are left k-synchronized*
- *$s_1[1,k_{s_1}]$ and $s_2[1,k_{s_2}]$ are left k-synchronized.*

Proof. Strings s_1 and s_2 are left k-synchronized iff $L^k(s_1) \bigcap L^k(s_2) \neq \emptyset$. We just have to apply Corollary 2.12. ■ □

Thus, the string $s[1,k_s]$, which is the longest prefix of s that matches a proper suffix of the taboos, provides all the information we need to construct $V^s_n(\mathbb{T})$ or $L^k(s)$.

2.7 Isomorphisms between taboo-free Hamming graphs

Here we will discuss isomorphism between Hamming graphs. Let us first introduce the formal definition of a taboo-free Hamming graph.

Definition 2.7. ***The taboo-free Hamming graph of length*** *n, $\Gamma_n(\mathbb{T}) := (V_n(\mathbb{T}), E_n(\mathbb{T}))$, is the graph with vertex set $V_n(\mathbb{T})$ such that two vertices $u, v \in V_n(\mathbb{T})$ are adjacent if their Hamming distance equals 1, that is, $e = \{u,v\} \in E_n(\mathbb{T})$ iff $d(u,v) = 1$. Analogously, $\Gamma^s_n(\mathbb{T})$ is the Hamming graph with vertex set $V^s_n(\mathbb{T})$.*

Examples of disconnected Hamming graphs are given in Figures 2.1 and 2.2. When dealing with taboo-free Hamming graphs, the following proposition is a simple way to establish graph isomorphisms.

Proposition 2.14. *Consider a taboo-set $\mathbb{T}$, a taboo-free string s and a taboo-free string w satisfying $ws \in V_{|w|+|s|}(\mathbb{T})$. If $V^{ws}_{n+|s|}(\mathbb{T}) = V^w_n(\mathbb{T}) \circ s$ for some $n \geq |w|$, then $\Gamma^{ws}_{n+|s|}(\mathbb{T})$ and $\Gamma^w_n(\mathbb{T})$ are isomorphic.*

Proof. By assumption, the vertex set of $\Gamma^{ws}_{n+|s|}(\mathbb{T})$ is $V^{ws}_{n+|s|}(\mathbb{T}) = V^w_n(\mathbb{T}) \circ s$. Thus, the map

$$f\colon V^w_n(\mathbb{T}) \circ s \to V^w_n(\mathbb{T})$$
$$rs \mapsto r$$

is well defined and bijective. Moreover, f is an edge-preserving bijection: Given any pair of strings $r_1, r_2 \in \Sigma^n$ and any string $s \in \Sigma^{|s|}$, then $d(r_1, r_2) = 1$ iff $d(r_1 s, r_2 s) = 1$. ■ □

Propositions 2.14 and 2.11.a imply that, for a taboo-free string s with $|s| \geq M$, the graphs $\Gamma^s_{n+|s|}(\mathbb{T})$ and $\Gamma^{s[1,M-1]}_{n+M-1}(\mathbb{T})$ are isomorphic. Furthermore Prop. 2.11.b implies that $\Gamma^s_{n+|s|}(\mathbb{T}) \simeq \Gamma^{s[1,k_s]}_{n+k_s}(\mathbb{T})$, which can be stated as follows.

Proposition 2.15. *Consider a taboo-set $\mathbb{T}$ and a taboo-free string s. There exists a unique $w \in \mathrm{suf}(\mathbb{T})$ such that $w = s[1, k_s]$. Moreover, for any $n \geq 0$,*

$$\Gamma^s_{n+|s|}(\mathbb{T}) \simeq \Gamma^w_{n+|w|}(\mathbb{T}).$$

Prop. 2.15 does not describe in which cases $V^s_{n+|s|}(\mathbb{T}) = \emptyset$. However, if $\mathbb{T}$ is left proper, Prop. 2.8 implies that this happens iff $|s| \leq M - 1$ and $s \notin \mathrm{suf}(V_M(\mathbb{T}))$. This suggests that we can state a version of Prop. 2.15 for left proper $\mathbb{T}$. But first, due to our interest in taboo-free strings of length M, we introduce the following.

Definition 2.8. *Given a left proper taboo-set $\mathbb{T}$,* ***the long suffix classification*** $\mathrm{lsc}(\mathbb{T})$ *is defined as*

$$\mathrm{lsc}(\mathbb{T}) := \{w \in \mathrm{suf}(\mathbb{T}) \text{ such that } \exists s \in V_M(\mathbb{T}) \text{ satisfying } s[1, k_s] = w\},$$

that is, $\mathrm{lsc}(\mathbb{T})$ *is the set of all suffixes of taboos that are the longest prefix of at least one taboo-free string of length M.*

Example 2.5. *If $\Sigma_1 = \{A, C, G, T\}$ and $\mathbb{T}_1 = \{AA,\ CC,\ GG,\ TT\}$, then*

$$\mathrm{lsc}(\mathbb{T}_1) \subseteq \mathrm{suf}(\mathbb{T}_1) = \{A,\ C,\ G,\ T,\ e\} = \Sigma_1 \bigcup \{e\}.$$

For any $s \in V_2(\mathbb{T}_1)$, we see $k_s > 0$, hence $e \notin \mathrm{lsc}(\mathbb{T}_1)$. Moreover,

$$\{AC, CG, GT, TA\} \subseteq V_2(\mathbb{T}_1),$$

yielding $\mathrm{lsc}(\mathbb{T}_1) = \Sigma_1$. If we consider $\Sigma_2 := \{A, C, G, T, C'\}$, where C' could represent a 5-methylcytosine, and $\mathbb{T}_2 := \mathbb{T}_1$, then string $s = C'A$ satisfies $k_s = 0$, hence

$\mathrm{lsc}(\mathbb{T}_2) = \mathrm{suf}(\mathbb{T}_2)$.

The following theorem classifies graphs $\Gamma_n^s(\mathbb{T})$ for left proper $\mathbb{T}$.

Theorem 2.16. *Consider a left proper taboo-set $\mathbb{T}$ and a taboo-free string s such that either $|s| \geq M$ or $s \in \mathrm{suf}(V_M(\mathbb{T}))$. Then a unique $w \in \mathrm{suf}(V_M(\mathbb{T})) \bigcap \mathrm{suf}(\mathbb{T})$ exists such that $w = s[1, k_s]$, which satisfies $\Gamma^s_{n+|s|}(\mathbb{T}) \simeq \Gamma^w_{n+|w|}(\mathbb{T})$ for $n \geq 0$. Moreover, if $|s| \geq M$, then $w \in \mathrm{lsc}(\mathbb{T})$.*

Proof. Prop. 2.8.b yields $V^s_{n+|s|}(\mathbb{T}) \neq \emptyset$ for $n \geq 0$, while $\Gamma^s_{n+|s|}(\mathbb{T}) \simeq \Gamma^{s[1,k_s]}_{n+k_s}(\mathbb{T})$ for $n \geq 0$ follows from Prop. 2.15. Hence we can set $w := s[1, k_s]$, which by definition belongs to $\mathrm{suf}(\mathbb{T})$. Since by assumption either $|s| \geq M$ or $s \in \mathrm{suf}(V_M(\mathbb{T}))$, it follows from Prop. 2.7 that s is k-prefixable for any k, and thus also $w := s[1, k_s]$ is k-prefixable. We consider $x \in L^{M-k_s}(w)$, which satisfies $xw \in V_M(\mathbb{T})$. Therefore $w = (xw)[M - k_s + 1, M] \in \mathrm{suf}(V_M(\mathbb{T}))$. All in all, $w \in \mathrm{suf}(V_M(\mathbb{T})) \bigcap \mathrm{suf}(\mathbb{T})$. This w is trivially unique since k_s is uniquely determined given s.
As for the case $|s| \geq M$, the fact that $s[1, M] \in V_M(\mathbb{T})$ and the definition of $\mathrm{lsc}(\mathbb{T})$ implies that $w \in \mathrm{lsc}(\mathbb{T})$. ■

□

In formal terms, Theorem 2.16 states that the equivalence relation "being isomorphic" divides all graphs $\Gamma^s_{n+|s|}(\mathbb{T})$ into equivalence classes. The representative of each class is a graph $\Gamma^w_{n+|w|}(\mathbb{T})$, where $w \in \mathrm{suf}(V_M(\mathbb{T})) \bigcap \mathrm{suf}(\mathbb{T})$. When $|s| \geq M$, string w belongs to $\mathrm{lsc}(\mathbb{T})$. This is why $\mathrm{lsc}(\mathbb{T})$ is called the long suffix classification.

To efficiently compute $\mathrm{lsc}(\mathbb{T})$, we recommend that $\mathbb{T}$ be minimal. Theorem 2.16 implies that

$$\mathrm{lsc}(\mathbb{T}) \subseteq \mathrm{suf}(V_M(\mathbb{T})) \bigcap \mathrm{suf}(\mathbb{T}), \tag{2.2}$$

and thus we define the **short suffix classification** as

$$\mathrm{ssc}(\mathbb{T}) := \Big(\mathrm{suf}(V_M(\mathbb{T})) \bigcap \mathrm{suf}(\mathbb{T}) \Big) - \mathrm{lsc}(\mathbb{T}). \tag{2.3}$$

The set $\mathrm{ssc}(\mathbb{T})$ is called short suffix classification because only when $|s| < M$ it can happen that a graph $\Gamma^s_{n+|s|}(\mathbb{T})$ is represented by a graph $\Gamma^w_{n+|w|}(\mathbb{T})$ with $w \in \mathrm{ssc}(\mathbb{T})$. Note that, if a string w satisfies the condition $|w| < M - 1$ and $w \circ R^i(w) \subseteq \mathrm{suf}(\mathbb{T})$ for some $i \in [1, M - 1 - |w|]$, then any $s \in w \circ R^i(w)$ satisfies $s[1, k_s + i] \in \mathrm{suf}(\mathbb{T})$, hence $w \notin \mathrm{lsc}(\mathbb{T})$. This property is used in the following example.

Example 2.6. *If* $\Sigma_1 = \{A, C, G, T\}$ *and* $\mathbb{T}_1 = \{AA,\ CC,\ GG,\ TT\}$*, then it is clear that* $e \in \mathrm{suf}(V_M(\mathbb{T})) \bigcap \mathrm{suf}(\mathbb{T})$*, because the empty string* e *belongs to both sets. Moreover,* $e \notin \mathrm{lsc}(\mathbb{T}_1)$ *due to* $e \circ \Sigma \subseteq \mathrm{suf}(\mathbb{T}_1)$*. Therefore* $e \in \mathrm{ssc}(\mathbb{T})$*.*

2.8 Connectivity of taboo-free Hamming graphs

We will make extensive use of the quotient graph to study the connectivity of taboo-free Hamming graphs. Before we start with the technicalities, we briefly describe our initial strategy.

For a Hamming graph $\Gamma_{n+j}(\mathbb{T})$, let us consider two different subsets of its vertex set, namely $V_{n+j}^{s_b}(\mathbb{T})$ and $V_{n+j}^{s_c}(\mathbb{T})$, where $s_b, s_c \in V_j(\mathbb{T})$. These two subsets are disjoint, so we can use the quotient graph $\mathcal{Q}[\Gamma_{n+j}(\mathbb{T})]$ to make each of them collapse in a single vertex, represented respectively by s_b and s_c. We will prove in Prop. 2.17 that s_b and s_c are adjacent in $\mathcal{Q}[\Gamma_{n+j}(\mathbb{T})]$ iff strings s_b and s_c have Hamming distance 1 and are left n-synchronized. This is specially interesting, because we know from Prop. 2.9 that two left $(M-1)$-synchronized strings are left n-synchronized for any $n \in \mathbb{N}$. Thus, it is enough to know that s_b, s_c are adjacent in $\mathcal{Q}[\Gamma_{(M-1)+j}(\mathbb{T})]$ to claim that s_b, s_c are adjacent in all partition graphs $\mathcal{Q}[\Gamma_{n+j}(\mathbb{T})]$ for $n \in \mathbb{N}$ (that is the essential content of Lemma 2.18). More formally, we have the following results.

Proposition 2.17. *Given taboo-set* $\mathbb{T}$*,* $j \in \mathbb{N}_0$ *and* $n \in \mathbb{N}_0$*, consider graph* $\Gamma_{n+j}(\mathbb{T})$ *and a subset* $S \subseteq V_{n+j}(\mathbb{T})$ *partitioned as* $S = \bigsqcup_{b \in J} V_{n+j}^{s_b}(\mathbb{T})$*, where* s_b *are taboo-free strings of length* j*. Consider moreover the quotient graph* $\mathcal{Q}[\Gamma_{n+j}(\mathbb{T})(S)] = \{J, E_J\}$*, where* $\Gamma_{n+j}(\mathbb{T})(S)$ *denotes the graph induced by* S *in* $\Gamma_{n+j}(\mathbb{T})$*.*

In these conditions, a pair of vertices $b, c \in J$ *is connected by an edge* $\{b, c\} \in E_J$ *iff the pair* s_b*,* s_c *is left* n*-synchronized and* $d(s_b, s_c) = 1$*.*

Proof. By definition, b and c are adjacent in $\mathcal{Q}[\Gamma_{n+j}(\mathbb{T})(S)]$ iff in graph $\Gamma_{n+j}(\mathbb{T})$ an edge connects a vertex in $V_{n+j}^{s_b}(\mathbb{T})$ with a vertex in $V_{n+j}^{s_c}(\mathbb{T})$. Since $d(s_b, s_c) \geq 1$, this edge exists iff $d(s_b, s_c) = 1$ and there exists $s \in V_n(\mathbb{T})$ such that $ss_b, ss_c \in V_{n+j}(\mathbb{T})$. The last condition is the definition of s_b and s_c being left n-synchronized. ■

□

The combination of Prop. 2.17 and Prop. 2.9 gives the following lemma.

Lemma 2.18. *Given a left proper taboo-set* $\mathbb{T}$*, a taboo-free string* s *and* $k \in \mathbb{N}$*, consider, for any* $n \geq |s| + k$*, partition* $V_n^s(\mathbb{T}) = \bigsqcup_{w \in L^k(s)} V_n^{ws}(\mathbb{T})$ *and quotient graph*

$\mathcal{Q}[\Gamma_n^s(\mathbb{T})] = (L^k(s), E_{L^k(s)})$. *Then it holds that*

$$\begin{aligned}\mathcal{Q}[\Gamma^s_{|s|+k}(\mathbb{T})] \supseteq \mathcal{Q}[\Gamma^s_{|s|+k+1}(\mathbb{T})] \supseteq \cdots \supseteq \mathcal{Q}[\Gamma^s_{|s|+k+M-1}(\mathbb{T})] = \\ = \mathcal{Q}[\Gamma^s_{|s|+k+M}(\mathbb{T})] = \mathcal{Q}[\Gamma^s_{|s|+k+M+1}(\mathbb{T})] = \cdots .\end{aligned}$$

If $\mathcal{Q}[\Gamma^s_{|s|+k+M-1}(\mathbb{T})]$ *is connected, then* $\mathcal{Q}[\Gamma^s_n(\mathbb{T})]$ *is connected for* $n \geq |s| + k$.

Proof. For some $n_0 \geq |s| + k$, consider an edge $\{w_b, w_c\}$ of graph $\mathcal{Q}[\Gamma^s_{n_0}(\mathbb{T})]$, where $w_b, w_c \in L^k(s)$. We set $s_b := w_b s$ and $s_c := w_c s$. Prop. 2.17 implies that w_b and w_c are adjacent in $\mathcal{Q}[\Gamma^s_{n_0}(\mathbb{T})]$ iff s_b and s_c are are left $(n_0 - |s| - k)$-synchronized and $d(w_b, w_c) = 1$. Since s_b and s_c are left $(n_0 - |s| - k)$-synchronized, they are also left $(n - |s| - k)$-synchronized for any $n \leq n_0$, and thus w_b and w_c are adjacent in $\mathcal{Q}[\Gamma^s_n(\mathbb{T})]$ for $|s| + k \leq n \leq n_0$. Hence the decreasing chain of quotient graphs is proven.

Now we will prove that this chain stabilizes after $n = |s| + k + M - 1$. If $n_0 - |s| - k = M - 1$, then, according to Proposition 2.9, w_b and w_c are left k-synchronized for arbitrary k, and thus Prop. 2.17 implies that w_b and w_c are adjacent in $\mathcal{Q}[\Gamma^s_n(\mathbb{T})]$ for arbitrary $n \geq |s| + k$. All in all, $\mathcal{Q}[\Gamma^s_{n_0}(\mathbb{T})]$ and $\mathcal{Q}[\Gamma^s_n(\mathbb{T})]$ have the same edges, as desired.

Regarding connectivity, given graphs G_1 and G_2 with the same vertex set $V_1 = V_2$ such that $G_1 \subseteq G_2$, if subgraph G_1 is connected, then G_2 is connected. ■ □

Figure 2.4 visualizes Lemma 2.18 for alphabet $\Sigma = \{a, b, c\}$, taboo-set $\mathbb{T} = \{ba, aa, ac, cc\}$ (which is left proper), suffix $s = b$ and $k = 1$.

We are finally ready to study the connectivity of graphs $\Gamma^s_n(\mathbb{T})$ for $|s| \geq M$. Let us begin with the following lemma.

Lemma 2.19. *Given a left proper* $\mathbb{T}$*, for any* $w \in V_M(\mathbb{T})$ *consider the set* $V^w_{2M}(\mathbb{T})$ *and partition*

$$V^w_{2M}(\mathbb{T}) = \bigsqcup_{a \in L^1(w)} V^{aw}_{2M}(\mathbb{T}),$$

inducing the quotient graph $\mathcal{Q}[\Gamma^w_{2M}(\mathbb{T})] = (L^1(w), E_{L^1(w)})$. *Then the following statements are equivalent:*

a) For every $w \in V_M(\mathbb{T})$, $\mathcal{Q}[\Gamma^w_{2M}(\mathbb{T})]$ *is connected.*

b) For every $w \in V_M(\mathbb{T})$ *and integer* $n \geq M$, $\Gamma^w_n(\mathbb{T})$ *is connected.*

Proof. Prop. 2.2 states that, in a connected graph, every quotient graph is connected, and thus *b*) implies *a*) by considering $n = 2M$.

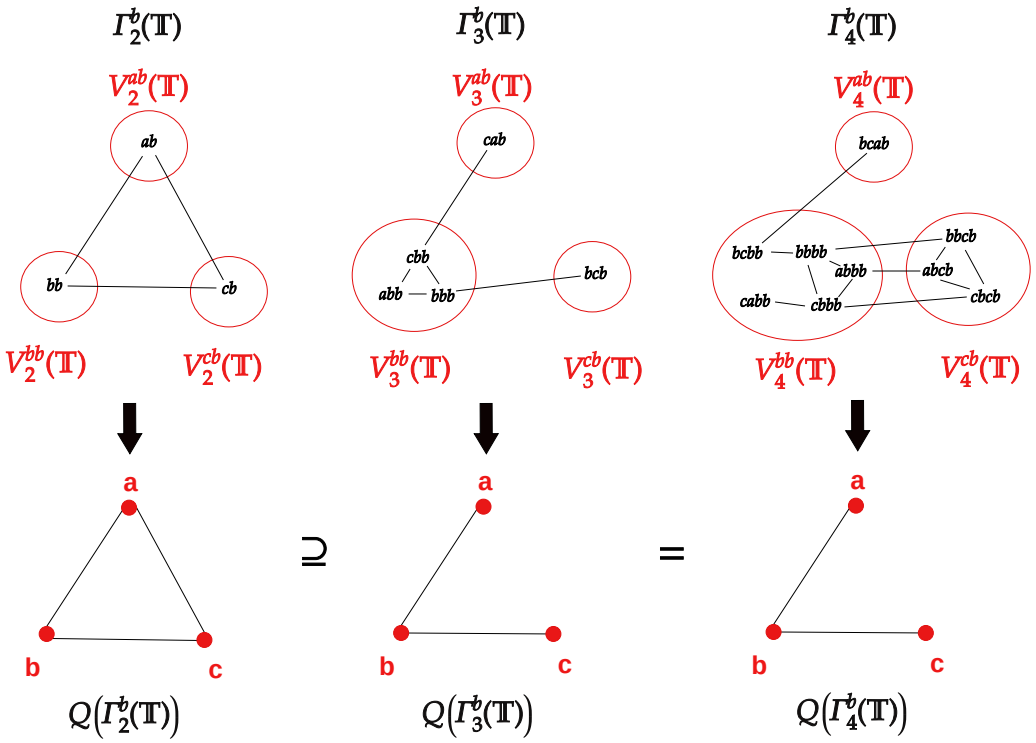

Figure 2.4: Visualization of Lemma 2.18 for $\Sigma = \{a, b, c\}$, $\mathbb{T} = \{ba, aa, ac, cc\}$, $s = b$ and $k = 1$. It holds that $L^1(b) = \{a, b, c\}$.

Now we prove by induction that $a)$ implies $b)$. For $n = M$ and $w \in V_M(\mathbb{T})$, we have that

$$V_M^w(\mathbb{T}) = \{w\},$$

hence $\Gamma_M^w(\mathbb{T})$ is connected. For the inductive step, assume that $\Gamma_n^w(\mathbb{T})$ is connected for every $w \in V_M(\mathbb{T})$ and up to an integer $n \geq M$. We will prove that also every $\Gamma_{n+1}^w(\mathbb{T})$ is connected. Consider

$$V_{n+1}^w(\mathbb{T}) = \bigsqcup_{a \in L^1(w)} V_{n+1}^{aw}(\mathbb{T}).$$

Let us write w separating the first $M-1$ symbols from the last one, that is $w = rc$ for $r \in \Sigma^{M-1}$ and $c \in \Sigma$. Then for any $a \in L^1(w)$, $V_{n+1}^{aw}(\mathbb{T}) = V_{n+1}^{arc}(\mathbb{T})$. Since $|r| = M-1$, Prop. 2.11.a implies $V_{n+1}^{arc}(\mathbb{T}) = V_n^{ar}(\mathbb{T}) \circ c$, while the isomorphism established in Prop. 2.14 yields

$$\Gamma_{n+1}^{aw}(\mathbb{T}) = \Gamma_{n+1}^{arc}(\mathbb{T}) \simeq \Gamma_n^{ar}(\mathbb{T}).$$

Thus, every $\Gamma_{n+1}^{aw}(\mathbb{T})$ is connected, because the induction hypothesis implies that $\Gamma_n^{ar}(\mathbb{T})$ is connected since $ar \in V_M(\mathbb{T})$. To prove that graph $\Gamma_{n+1}^w(\mathbb{T})$ is connected,

it remains to apply Prop. 2.1, so we need to prove that the quotient graph induced by partition $V_{n+1}^w(\mathbb{T}) = \bigsqcup_{a\in L^1(w)} V_{n+1}^{aw}(\mathbb{T})$, namely $\mathcal{Q}[\Gamma_{n+1}^w(\mathbb{T})]$, is connected.

We know that, given partition $V_{2M}^w(\mathbb{T}) = \bigsqcup_{a\in L^1(w)} V_{2M}^{aw}(\mathbb{T})$, the quotient graph $\mathcal{Q}[\Gamma_{2M}^w(\mathbb{T})]$ is connected. Applying Lemma 2.18 with $s = w$ and $k = 1$, we get the following chain of inclusions:

$$\begin{aligned}\mathcal{Q}[\Gamma_{M+1}^w(\mathbb{T})] \supseteq \mathcal{Q}[\Gamma_{M+2}^w(\mathbb{T})] \supseteq \cdots \supseteq \mathcal{Q}[\Gamma_{2M}^w(\mathbb{T})] = \\ = \mathcal{Q}[\Gamma_{2M+1}^w(\mathbb{T})] = \mathcal{Q}[\Gamma_{2M+2}^w(\mathbb{T})] = \cdots .\end{aligned}$$

Since $\mathcal{Q}[\Gamma_{2M}^w(\mathbb{T})]$ is connected, every quotient graph of the chain of inclusions is connected, as shown in Lemma 2.18. In particular, graph $\mathcal{Q}[\Gamma_{n+1}^w(\mathbb{T})]$ is an element of the chain of inclusions because $n + 1 \geq M + 1$, so it is connected, as desired. ■

□

Lemma 2.19 is very interesting: We wanted to characterize the connectivity of graphs $\Gamma_n^s(\mathbb{T})$ for $s \in V_M(\mathbb{T})$ and $n \geq M$. We have proved that it is enough to study a finite number of graphs, namely $\mathcal{Q}[\Gamma_{2M}^w(\mathbb{T})]$ for $s \in V_M(\mathbb{T})$, that is, $|V_M(\mathbb{T})|$ graphs. Let us summarize the connectivity results that follow from Lemma 2.19 and Theorem 2.16.

Proposition 2.20. *Given a left proper $\mathbb{T}$, the following statements are equivalent:*

a) For any taboo-free string s with $|s| \geq M$ and any integer $n \geq |s|$, $\Gamma_n^s(\mathbb{T})$ is connected.

b) For any $w \in V_M(\mathbb{T})$ and any integer $n \geq M$, $\Gamma_n^w(\mathbb{T})$ is connected.

c) For any $r \in \mathrm{lsc}(\mathbb{T})$, $\Gamma_{M+|r|}^r(\mathbb{T})$ is connected.

d) For any $r \in \mathrm{lsc}(\mathbb{T})$, the partition $V_{M+|r|}^r(\mathbb{T}) = \bigsqcup_{a\in L^1(r)} V_{M+|r|}^{ar}(\mathbb{T})$ induces a connected partition graph $\mathcal{Q}[\Gamma_{M+|r|}^r(\mathbb{T})]$.

Proof. Implication $a) \Rightarrow b)$ is obvious, while $b) \Rightarrow a)$ is proven as follows: Given $V_n^s(\mathbb{T})$, where s is a taboo-free string with $|s| \geq M$, Prop. 2.11.a implies that $V_n^s(\mathbb{T}) = V_n^{s[1,M-1]}(\mathbb{T}) \circ s[M, j]$. Since $s[M, j] = s[M, M]s[M + 1, j]$, applying Prop. 2.11.a again we have $V_n^s(\mathbb{T}) = V_n^{s[1,M]}(\mathbb{T}) \circ s[M + 1, j]$. Prop. 2.14 yields the isomorphism $\Gamma_{n+j}^s(\mathbb{T}) \simeq \Gamma_{n+M}^{s[1,M]}(\mathbb{T})$, and $\Gamma_{n+M}^{s[1,M]}(\mathbb{T})$ is connected due to $s[1, M] \in V_M(\mathbb{T})$ and the assumption of $b)$. Thus, statements $a)$ and $b)$ are equivalent.

Implication $b) \Rightarrow c)$ is consequence of Theorem 2.16. Moreover, $c) \Rightarrow d)$ follows from Prop. 2.2. It remains to prove $d) \Rightarrow b)$, which we do as follows. Corollary 2.12 implies $L^1(w) = L^1(w[1, k_w])$. Moreover, for any $w \in V_M(\mathbb{T})$ and $a, b \in L^1(w)$, we claim that the following statements are equivalent:

i) Strings aw and bw are left k-synchronized.

ii) Strings $aw[1, k_w]$ and $bw[1, k_w]$ are left k-synchronized.

Indeed, the implication i) $\Rightarrow$ ii) is obvious, so let us prove ii) $\Leftarrow$ i. Given a taboo-free string $s \in V_j(\mathbb{T})$ such that $saw[1, k_w]$ and $sbw[1, k_w]$ are taboo-free, we want to prove that also saw and sbw are taboo-free. But if that were not the case, it would be the consequence of either $(saw)[c, d] \in \mathbb{T}$ or $(sbw)[c, d] \in \mathbb{T}$ for some integers $1 \leq c \leq j < j + 1 + k_w \leq d \leq j + 1 + M$. However, that contradicts the maximality of k_w, yielding ii) $\Leftarrow$ i.

Our previous claim and Prop. 2.17 imply that, if $r = w[1, k_w]$ for some $w \in V_M(\mathbb{T})$, given partition $V^r_{M+|r|}(\mathbb{T}) = \bigsqcup_{a \in L^1(r)} V^{ar}_{M+|r|}(\mathbb{T})$, it holds that

$$\mathcal{Q}[\Gamma^r_n(\mathbb{T})] \simeq \mathcal{Q}[\Gamma^w_n(\mathbb{T})].$$

Theorem 2.16 implies that, for every $w \in V_M(\mathbb{T})$, there exists $r = w[1, k_w] \in \mathrm{lsc}(\mathbb{T})$. Applying Lemma 2.19, finally $d) \Rightarrow b)$ follows. ■

□

It is worth noticing how simpler the connectivity problem has become. Initially, we were studying whether every $\Gamma^s_n(\mathbb{T})$ with $|s| \geq M$ is connected, obtaining in Lemma 2.19 that this is equivalent to the connectivity of graphs $\Gamma^w_{2M}(\mathbb{T})$ for $w \in V_M(\mathbb{T})$, which are $|V_M(\mathbb{T})|$ graphs. Now we see, using Prop. 2.20 and the fact that $\mathrm{lsc}(\mathbb{T}) \subseteq \mathrm{suf}(\mathbb{T})$, that we only need to prove the connectivity of $|\mathrm{lsc}(\mathbb{T})| \leq |\mathrm{suf}(\mathbb{T})| \leq (M-1)|\mathbb{T}| + 1$ graphs, namely either $\mathcal{Q}[\Gamma^r_{M+|r|}(\mathbb{T})]$ or $\Gamma^r_{M+|r|}(\mathbb{T})$ for $r \in \mathrm{lsc}(\mathbb{T})$. We give an example.

Example 2.7. *Take* $\Sigma = \{A, C, G, T\}$ *and* $\mathbb{T} = \{AA,\ CCC\}$*, which is left proper. Using Prop. 2.20, since* $M = 3$ *and* $\mathrm{lsc}(\mathbb{T}) = \mathrm{suf}(\mathbb{T}) = \{e, A,\ C,\ CC\}$*, the connectivity of graphs*

$$\Gamma^e_3(\mathbb{T}),\ \Gamma^A_4(\mathbb{T}),\ \Gamma^C_4(\mathbb{T}),\ \Gamma^{CC}_5(\mathbb{T})$$

implies that any $\Gamma^w_n(\mathbb{T})$ *with* $w \in \mathrm{suf}(\mathbb{T})$ *is connected. Proposition 2.15 implies that, for any taboo-free string* s *and* $n \geq |s|$*,* $\Gamma^s_n(\mathbb{T})$ *is connected.*

Prop. 2.20 characterizes the connectivity of every $\Gamma^s_{n+|s|}(\mathbb{T})$ for $|s| \geq M$. We know from Theorem 2.16 that there exists $r \in \mathrm{lsc}(\mathbb{T}) \subseteq \mathrm{suf}(V_M(\mathbb{T})) \bigcap \mathrm{suf}(\mathbb{T})$ such that $\Gamma^s_{n+|s|}(\mathbb{T}) \simeq \Gamma^r_{n+|r|}(\mathbb{T})$. Since $\mathrm{ssc}(\mathbb{T}) := \mathrm{suf}(V_M(\mathbb{T})) \bigcap \mathrm{suf}(\mathbb{T}) - \mathrm{lsc}(\mathbb{T})$, to complete our characterization of the connectivity of every taboo-free Hamming graph, some cases (such as Example 2.6) require considering the connectivity of graphs $\Gamma^p_n(\mathbb{T})$ for $p \in \mathrm{ssc}(\mathbb{T})$. We have the following.

Proposition 2.21. *Given a left proper* $\mathbb{T}$ *and* $p \in \mathrm{ssc}(\mathbb{T})$*, assume that, for every* $r \in \mathrm{lsc}(\mathbb{T})$*, graph* $\Gamma^r_{M+|r|}(\mathbb{T})$ *is connected. Given* $k \in \mathbb{N}$*, if partition*

$$V^p_{|p|+k+M-1}(\mathbb{T}) = \bigsqcup_{w \in L^k(p)} V^{wp}_{|p|+k+M-1}(\mathbb{T})$$

satisfies that $(wp)[1, k_{wp}] \in \mathrm{lsc}(\mathbb{T})$ *for each* $w \in L^k(p)$*, and moreover* $\mathcal{Q}[\Gamma^p_{|p|+k+M-1}(\mathbb{T})]$ *is connected, then* $\Gamma^p_n(\mathbb{T})$ *is connected for* $n \geq |p| + k$*.*

Proof. For $n \geq |p| + k$, given partition

$$V^p_n(\mathbb{T}) = \bigsqcup_{w \in L^k(p)} V^{wp}_n(\mathbb{T}),$$

subgraphs $\Gamma^{wp}_n(\mathbb{T})$ are connected due to $(wp)[1, k_{wp}] \in \mathrm{lsc}(\mathbb{T})$. Moreover, since $\mathcal{Q}[\Gamma^p_{2M-1}(\mathbb{T})]$ is connected, Lemma 2.18 with $s = p$ implies that $\mathcal{Q}[\Gamma^p_n(\mathbb{T})]$ is connected for $n \geq |p|+k$. Thus, the quotient graph $\mathcal{Q}[\Gamma^p_n(\mathbb{T})]$ and all induced subgraphs $\Gamma^{wp}_n(\mathbb{T})$ are connected. The connectivity of $\Gamma^p_n(\mathbb{T})$ follows applying Prop. 2.1. ■ □

In Prop. 2.21, one can always take $k = M - |p|$ and just check if $\mathcal{Q}[\Gamma^p_{2M-1}(\mathbb{T})]$ or $\Gamma^p_{2M-1}(\mathbb{T})$ is connected for $p \in \mathrm{ssc}(\mathbb{T})$. Otherwise one can try $k = 1$ and increase it progressively.

Example 2.8. *If* $\Sigma = \{A, C, G, T\}$ *and* $\mathbb{T} = \{AA,\ CC,\ GG,\ TT\}$*, then it holds that* $\mathrm{lsc}(\mathbb{T}) = \{A, C, G, T\}$ *and* $\mathrm{ssc}(\mathbb{T}) = \{e\}$*. For* $r \in \mathrm{lsc}(\mathbb{T})$*, it can be proven that* $\Gamma^r_3(\mathbb{T})$ *is connected. Thus, Proposition 2.20 implies that every* $\Gamma^r_n(\mathbb{T})$ *is connected for* $r \in \mathrm{lsc}(\mathbb{T})$ *and* $n \geq 1$*.*

We can combine Propositions 2.20 and 2.21 to obtain our aimed characterization of the connectivity of every suffix Hamming graph. We do so in the following theorem.

Theorem 2.22. *Given a left proper taboo-set* $\mathbb{T}$*, the following are equivalent.*

a) Consider, for every $r \in \mathrm{lsc}(\mathbb{T})$*, partition* $V^r_{M+|r|}(\mathbb{T}) = \bigsqcup_{a \in L^1(r)} V^{ar}_{M+|r|}(\mathbb{T})$*, and for every* $p \in \mathrm{ssc}(\mathbb{T})$*, partition* $V^p_{2M-1}(\mathbb{T}) = \bigsqcup_{w \in L^{M-|p|}(p)} V^{wp}_{2M-1}(\mathbb{T})$*.*
For $r \in \mathrm{lsc}(\mathbb{T})$*, every partition graph* $\mathcal{Q}[\Gamma^r_{M+|r|}(\mathbb{T})]$ *is connected; for* $p \in \mathrm{ssc}(\mathbb{T})$*, every partition graph* $\mathcal{Q}[\Gamma^p_{2M-1}(\mathbb{T})]$ *is connected; for* $p \in \mathrm{ssc}(\mathbb{T})$*, every graph* $\Gamma^p_n(\mathbb{T})$ *with* $|p| + 2 \leq n \leq M - 1$ *is connected.*

b) For $r \in \mathrm{lsc}(\mathbb{T})$*, graph* $\Gamma^r_{M+|r|}(\mathbb{T})$ *is connected; for* $p \in \mathrm{ssc}(\mathbb{T})$*, graph* $\Gamma^p_{2M-1}(\mathbb{T})$ *is connected; for* $p \in \mathrm{ssc}(\mathbb{T})$ *and* $|p| + 2 \leq n \leq M - 1$*, every graph* $\Gamma^p_n(\mathbb{T})$ *is connected.*

c) For every taboo-free string s and $n \geq 0$, graph $\Gamma^s_{|s|+n}(\mathbb{T})$ is connected.

Proof. Prop. 2.2 states that the connectivity of a graph is equivalent to the connectivity of each of its quotient graphs. Hence $b) \Rightarrow a)$ follows, because if graphs $\Gamma^r_{M+|r|}(\mathbb{T})$ and $\Gamma^p_{2M-1}(\mathbb{T})$ are connected, then also partition graphs $\mathcal{Q}[\Gamma^r_{M+|r|}(\mathbb{T})]$ and $\mathcal{Q}[\Gamma^p_{2M-1}(\mathbb{T})]$ are connected. Since the implication $c) \Rightarrow b)$ is obvious, it only remains to prove $a) \Rightarrow c)$.

Theorem 2.16 states that, when $\mathbb{T}$ is left proper, every nonempty graph of the form $\Gamma^s_{n+|s|}(\mathbb{T})$ is isomorphic to graph $\Gamma^w_{n+|w|}(\mathbb{T})$, where $w = s[1, k_s] \in \mathrm{suf}(\mathbb{T}) \bigcap \mathrm{suf}(V_M(\mathbb{T}))$. By construction, strings in $\mathrm{suf}(\mathbb{T}) \bigcap \mathrm{suf}(V_M(\mathbb{T}))$ either belong to $\mathrm{lsc}(\mathbb{T})$ or $\mathrm{ssc}(\mathbb{T})$. Therefore, statement $c)$ is equivalent to the connectivity, for every $n \geq 0$, of every $\Gamma^r_{n+|r|}(\mathbb{T})$, where $r \in \mathrm{lsc}(\mathbb{T})$, and of every $\Gamma^p_{n+|p|}(\mathbb{T})$, where $p \in \mathrm{ssc}(\mathbb{T})$.

Assuming statement $a)$, since every partition graph $\mathcal{Q}[\Gamma^r_{M+|r|}(\mathbb{T})]$ is connected for $r \in \mathrm{lsc}(\mathbb{T})$, Prop. 2.20 implies that every graph $\Gamma^w_{M+n}(\mathbb{T})$ is connected, where $w \in V_M(\mathbb{T})$ and $n \geq 0$. For any $r \in \mathrm{lsc}(\mathbb{T})$, there exists by construction a $w \in V_M(\mathbb{T})$ such that $r = w[1, k_w]$. Since $\Gamma^w_{M+n}(\mathbb{T}) \simeq \Gamma^r_{|r|+n}(\mathbb{T})$ due to Prop. 2.15, it follows that $a)$ implies that every $\Gamma^r_{|r|+n}(\mathbb{T})$ is connected, where $r \in \mathrm{lsc}(\mathbb{T})$ and $n \geq 0$.

It remains to prove that $a)$ implies that every $\Gamma^p_{|p|+n}(\mathbb{T})$ is connected, where $p \in \mathrm{ssc}(\mathbb{T})$ and $n \geq 0$. Since every partition graph $\mathcal{Q}[\Gamma^p_{2M-1}(\mathbb{T})]$ is connected, Prop. 2.21 with $k = M - |p|$ implies that $\Gamma^p_{M+n}(\mathbb{T})$ is connected for $n \geq 0$. The connectivity of graphs $\Gamma^p_{|p|+2}(\mathbb{T}), \cdots, \Gamma^p_{M-1}(\mathbb{T})$ is part of the assumptions of $a)$, and graphs $\Gamma^p_{|p|+1}(\mathbb{T})$ and $\Gamma^p_{|p|}(\mathbb{T})$ are trivially connected, finishing the proof. ■ □

In general, if $\mathbb{T}$ has just a few taboos, proving connectivity becomes easier since most of strings are left k-synchronized. In Prop. 2.23 only previous results are used, while in Prop. 2.24 we study this case more exhaustively in a self-contained manner. Note that, when taboo-set $\mathbb{T}$ is minimal, the assumptions of Prop. 2.24 are much easier to check.

Proposition 2.23. *Given a left proper $\mathbb{T}$ such that every pair of strings $w_1, w_2 \in V_M(\mathbb{T})$ with $d(w_1, w_2) = 1$ is left 1-synchronized, it holds that:*

a) For any $r \in \mathrm{lsc}(\mathbb{T})$ and $n \in \mathbb{N}_0$, $\Gamma^r_{n+|r|}(\mathbb{T})$ is connected.

b) For any $p \in \mathrm{ssc}(\mathbb{T})$ with connected $\Gamma^p_M(\mathbb{T})$, $\Gamma^p_n(\mathbb{T})$ is connected for $n \geq M$.

Proof. Prop. 2.10 implies that every pair $w_1, w_2 \in V_M(\mathbb{T})$ is left k-synchronized for any $k \in \mathbb{N}$. We know from Lemma 2.17 that left k-synchronization of two strings with Hamming distance 1 indexing a partition as suffixes is equivalent to

those two strings being adjacent in the partition graph. Therefore any quotient graph $\mathcal{Q}[\Gamma_n^w(\mathbb{T})] = (L^1(w), E_{L^1(w)})$ induced by partition $V_n^w(\mathbb{T}) = \bigsqcup_{a\in L^1(w)} V_n^{aw}(\mathbb{T})$ is fully connected (that is, every two vertices are adjacent). In particular, every $\mathcal{Q}[\Gamma_n^w(\mathbb{T})]$ is connected, and thus Prop. 2.20 implies *a*). Similarly with partition $V_n^p(\mathbb{T}) = \bigsqcup_{w\in L^{M-|p|}(p)} V_n^{wp}(\mathbb{T})$, since $\mathcal{Q}[\Gamma_n^p(\mathbb{T})] \simeq \Gamma_M^p(\mathbb{T})$ for $n \geq M$, Prop. 2.21 implies *b*). ■

□

Example 2.9. *For $\Sigma = \{A, C, G, T\}$ and $\mathbb{T} = \{AA,\ CCC\}$, the strings Tw_1 and Tw_2 are taboo-free for $w_1, w_2 \in V_3(\mathbb{T})$, hence they are left 1-synchronized. Since $\mathrm{lsc}(\mathbb{T}) = \mathrm{suf}(\mathbb{T})$, for any taboo-free string s and $n \geq |s|$, $\Gamma_n^s(\mathbb{T})$ is connected.*

Proposition 2.24. *Given taboo-set $\mathbb{T}$ and set $\Psi(\mathbb{T}) := \bigcup_{t\in\mathbb{T}} t[2, |t|]$, if every pair of taboo-free strings $w_1, w_2 \in \Psi(\mathbb{T})$ with $|w_1| \geq |w_2|$ and $d\big(\, w_1[1, |w_2|] \,,\, w_2 \,\big) \leq 1$ is left 1-synchronized, then it holds that:*

a) Every taboo-free string is 1-prefixable. In particular, $\mathbb{T}$ is left proper.

b) Every two taboo-free strings s_1, s_2 with $d(s_1, s_2) = 1$ are left 1-synchronized.

c) Graph $\Gamma_n^s(\mathbb{T})$ is connected for every taboo-free string s and $n \geq |s|$.

Proof.

a) Consider any taboo-free string s. Assume that, for each $a \in \Sigma$, as is not taboo-free, that is, that for some integer $c_a \geq 2$, $(as)[1, c_a] \in \mathbb{T}$. WLOG assume $c_{a_1} \leq \cdots \leq c_{a_m}$ and consider $s[1, c_{a_m} - 1]$, which satisfies $s[1, c_{a_m} - 1] \in \Psi(\mathbb{T})$ since $(a_m s)[1, c_{a_m}] \in \mathbb{T}$. By construction, for any $a \in \Sigma$, string $as[1, c_{a_m} - 1]$ is not taboo-free. On the other hand, the Hamming distance between $s[1, c_{a_m} - 1] \in \Psi(\mathbb{T})$ and itself is 0, and thus the assumption of the statement implies that $s[1, c_{a_m} - 1]$ is left 1-synchronized with $s[1, c_{a_m} - 1]$. In other words, a symbol $a \in \Sigma$ exists such that $as[1, c_{a_m} - 1]$ is taboo-free, which is a contradiction. All in all, s must be 1-prefixable. Taking $s \in V_M(\mathbb{T})$ we see that $\mathbb{T}$ is left proper.

b) Given taboo-free strings s_1, s_2 such that $d(s_1, s_2) = 1$, assume that they are not 1-synchronized. Then for every $a \in \Sigma$, either $(as_1)[1, c_a] \in \mathbb{T}$ or $(as_2)[1, c_a] \in \mathbb{T}$ for some $c_a \geq 2$. Denote by $C_1 \subseteq \bigcup_{a\in\Sigma}\{c_a\}$ those c_a such that $(as_1)[1, c_a] \in \mathbb{T}$, and analogously with C_2. If C_1 were empty, then s_2 would not be 1-prefixable, contradicting *a*). Thus, both C_1 and C_2 must be nonempty. Consider $d_1 := \max\{c : c \in C_1\}$ and $d_2 := \max\{c : c \in C_2\}$. It holds that $s_1[2, d_1] \in \Psi(\mathbb{T})$ and $s_2[2, d_2] \in \Psi(\mathbb{T})$. Moreover, we have that the

pair $s_1[2, d_1]$, $s_2[2, d_2]$ is not left 1-synchronized. Since $d(s_1, s_2) = 1$, that contradicts the assumptions of the statement, hence s_1 and s_2 must be left 1-synchronized, as desired.

c) Clearly $\Gamma^s_{|s|}(\mathbb{T})$ is connected, so let us proceed by induction. Assume $\Gamma^s_n(\mathbb{T})$ is connected for a fixed $n \geq |s|$ and consider $\Gamma^s_{n+1}(\mathbb{T})$. Since $V^s_{n+1}(\mathbb{T}) \subseteq \Sigma \circ V^s_n(\mathbb{T})$, if $|V^s_n(\mathbb{T})| = 1$, then $\Gamma^s_{n+1}(\mathbb{T})$ is connected. Otherwise we take different $s_1, s_2 \in V^s_{n+1}(\mathbb{T})$; we will prove that they are connected. We know that $s_1, s_2 \in \Sigma \circ V^s_n(\mathbb{T})$, hence let us write $s_1 = c_1 w_1$ and $s_2 = c_2 w_2$ for $c_i \in \Sigma$ and $w_i \in V^s_n(\mathbb{T})$. If $w_1 = w_2$, the result is obvious, so assume $w_1 \neq w_2$.

By hypothesis, $\Gamma^s_n(\mathbb{T})$ is connected, and thus there exists a path of vertices of $V^s_n(\mathbb{T})$, namely $y_1, \cdots, y_D$, such that $d(y_i, y_{i+1}) = 1$, $y_1 = w_1$ and $y_D = w_2$. For every $j \in [1, D-1]$, the pair y_j, y_{j+1} is left 1-synchronized, and thus there exists $b_j \in \Sigma$ such that $b_j y_j$ and $b_j y_{j+1}$ are taboo-free. Since $d(b_j y_j, b_j y_{j+1}) = 1$, $b_j y_j$ and $b_j y_{j+1}$ are adjacent in $\Gamma^s_{n+1}(\mathbb{T})$. Moreover every pair of taboo-free strings contained in $\Sigma \circ y_i$ is adjacent for $i \in [1, D-1]$. Since the relation "being connected" is transitive, vertices $s_1 \in \Sigma \circ y_1$ and $s_2 \in \Sigma \circ y_D$ are connected, as desired. ■

□

Example 2.10. *If* $\Sigma = \{A, C, G, T\}$ *and* $\mathbb{T} = \{AA,\ CC,\ GG,\ TT\}$, *then* $\Psi(\mathbb{T}) = \{A, C, G, T\}$. *Every pair of strings in* $\Psi(\mathbb{T})$ *is left* 1*-synchronized, hence for every taboo-free* s *and* $n \geq |s|$, $\Gamma^s_n(\mathbb{T})$ *is connected.*

Now we aim to find an upper bound for the number of taboos needed to guarantee connectivity of the graphs $\Gamma^s_n(\mathbb{T})$. The following Corollary of Prop. 2.24 holds.

Corollary 2.25. *Consider an alphabet* Σ *and a taboo-set* $\mathbb{T}$. *The following holds:*

a) If $|\mathbb{T}[1,1]| < |\Sigma|$, *then for any taboo-free string* s *and* $n \geq |s|$, $\Gamma^s_n(\mathbb{T})$ *is connected.*

b) If $|\mathbb{T}| < |\Sigma|$, *then for any taboo-free string* s *and* $n \geq |s|$, $\Gamma^s_n(\mathbb{T})$ *is connected.*

Proof.

a) Assume that taboo-free strings s_1, s_2 satisfy $L^1(s_1) \bigcap L^1(s_2) = \emptyset$. That is, for each $a \in \Sigma$, either as_1 or as_2 has a taboo as prefix, contradicting $|\mathbb{T}[1,1]| < |\Sigma|$. Therefore every two taboo-free strings are left 1-synchronized, so we can apply Prop. 2.24.c, implying a).

b) If $|\mathbb{T}| < |\Sigma|$, then $|\mathbb{T}[1,1]| < |\Sigma|$. Thus, statement *a)* yields the result. ■

□

Corollary 2.25.b implies that, if $|\mathbb{T}| < |\Sigma|$, then every $\Gamma_n^s(\mathbb{T})$ is connected. In Examples 2.11 and 2.12, we give examples of taboo-sets over an alphabet with $|\Sigma| = 2$ and $|\Sigma| > 2$ symbols respectively, such that $|\mathbb{T}| = |\Sigma|$ and at least one suffix graph is disconnected. In this sense, the upper bound $|\mathbb{T}| < |\Sigma|$ that guarantees connectivity for every suffix graph cannot be improved.

Example 2.11. *If* $\Sigma = \{0,1\}$ *and* $\mathbb{T} = \{10,\ 01\}$*, then* $\mathbb{T}$ *is left proper and* $|\mathbb{T}[1,1]| = |\mathbb{T}[2,2]| = 2 = |\Sigma|$*. For* $n \geq 2$*,* $V_n(\mathbb{T}) = \{0\cdots 0,\ 1\cdots 1\}$*, which makes* $\Gamma_n(\mathbb{T})$ *disconnected. The trivial graphs* $\Gamma_n^0(\mathbb{T})$ *and* $\Gamma_n^1(\mathbb{T})$ *are both connected.*

Example 2.12. *For* $m \geq 3$*,* $\Sigma = \{a_1, \cdots, a_m\}$ *and the left proper taboo-set*

$$\mathbb{T} = \{a_3a_1,\ a_4a_1,\ a_5a_1,\ \cdots,\ a_ma_1\} \bigsqcup \{a_1a_2,\ a_2a_2\},$$

we claim that $\Gamma_n^{a_1}(\mathbb{T})$ *is disconnected for* $n \geq 3$*. Indeed,*

$$\begin{aligned} V_n^{a_1}(\mathbb{T}) &= V_n^{a_1a_1}(\mathbb{T}) \bigsqcup V_n^{a_2a_1}(\mathbb{T}) = \\ &= \Big(V_n^{a_2a_1a_1}(\mathbb{T}) \bigsqcup V_n^{a_1a_1a_1}(\mathbb{T})\Big) \bigsqcup \Big(\bigsqcup_{i\in[3,m]} V_n^{a_ia_2a_1}(\mathbb{T})\Big), \end{aligned}$$

so take $s \in V_n^{a_2a_1a_1}(\mathbb{T}) \bigsqcup V_n^{a_1a_1a_1}(\mathbb{T})$ *and* $r \in \bigsqcup_{i\in[3,m]} V_n^{a_ia_2a_1}(\mathbb{T})$*. It holds that* $d(s,r) \geq 2$*, hence we found two disconnected components in graph* $\Gamma_n^{a_1}(\mathbb{T})$*. This is coherent with* $|\mathbb{T}[1,1]| = |\Sigma| = m$*.*
To generalize this example, for $i \in \mathbb{N}_0$*, denote by* $s_i := a_1\overset{i)}{\cdots}a_1$ *the concatenation of* i a_1*'s. The taboo-set*

$$\mathbb{T}_i = \{a_3s_i, a_4s_i, \cdots, a_ms_i\} \bigsqcup \{a_1a_2s_{i-1}, a_2a_2s_{i-1}\}$$

satisfies that graph $\Gamma_n^{s_i}(\mathbb{T}_i)$ *is disconnected for* $n \geq i+2$*.*

In this section, we have stated various results regarding the connectivity of every suffix Hamming graph given a left proper taboo-set $\mathbb{T}$. Up to Theorem 2.16, our aim was to characterize the connectivity of every suffix Hamming graph. Then we found sufficient conditions in Prop. 2.24 and Corollary 2.25 that are easier to apply. When studying this connectivity problem, the practitioner should firstly try to apply the results requiring easy-to-check assumptions, and increasingly use the more complicated ones. Given a taboo-set $\mathbb{T}$, a possible workflow would be the following:

1) We check if $|\mathbb{T}[1,1]| < |\Sigma|$. If it holds, we can apply Corollary 2.25.a. Otherwise go to step 2)

2) In order to apply Prop. 2.24, we check if every pair of taboo-free strings $w_1, w_2 \in \Psi(\mathbb{T})$ with $|w_1| \geq |w_2|$ and $d\big(\, w_1[1,|w_2|]\,,\, w_2\,\big) \leq 1$ is left 1-synchronized. If it does not hold, go to step 3)

3) We check whether $\mathbb{T}$ is left proper (this holds in all the biological examples that we considered so far). Otherwise redefine an equivalent left proper taboo-set and apply the characterization of Theorem 2.22. Two possibilities can arise: Either every suffix Hamming graph is connected, and thus evolution can explore all the space of taboo-free strings; or some taboo-free strings belonging to $\mathrm{lsc}(\mathbb{T})$ or $\mathrm{ssc}(\mathbb{T})$ induce disconnected suffix graphs $\Gamma^s_{n_0}(\mathbb{T})$ for some $n_0 \geq |s|+M$, implying that $\Gamma^s_n(\mathbb{T})$ stays disconnected for $n \geq n_0$.

2.9 Examples of plausible bacterial taboo-sets

Taboo-sets as generated by the avoidance of restriction sites can assume various levels of complexities. In this section, we discuss some examples from REBASE [Roberts et al., 2014] using the theory developed in this work. Note that many restriction enzymes of REBASE database have an unknown recognition site, hence our taboo-sets may underestimate the actual amount of taboos. Before describing the examples, we will briefly review essential nomenclature for DNA sequences.

DNA is double-stranded, where A pairs with T and G pairs with C, hence it suffices to discuss only one of the strands. We adopt the convention that, given any of the strands, the DNA sequence is always represented from the 5' end to the 3' end (which is chemically determined). As a consequence, given a DNA sequence, **its complementary DNA sequence**, the one lying on the opposite strand, is obtained by inverting the order of the symbols and carrying through substitutions $A \leftrightarrow T$ and $C \leftrightarrow G$. If a DNA sequence s is identical to its complementary DNA sequence, we say that s is an **inverted repeat** [Ussery et al., 2008]. For example, sequence $CCGG$ is an inverted repeat.

The fact that DNA is double-stranded implies that each recognition site induces taboos in pairs, namely itself and its complementary DNA sequence. For example, if $AGGGC$ is a recognition site, then also the complementary strand $GCCCT$ is a taboo. If, however, the recognition site is an inverted repeat such as $TGCA$, then this pair is actually one single recognition site. Recognition sites of type II R-M systems are nearly always an inverted repeat [Rusinov et al., 2015, Gelfand and

Koonin, 1997], and therefore one recognition site induces one single taboo. This is specially interesting because, according to Rusinov et al. [2015, 2018a], only type II R-M systems induce taboos.

A permutation of the symbols of alphabet Σ does not alter any of the results that we proved along this work. Moreover, by reversing the order of the symbols, any statement regarding e.g. left-properness and suffixes has an analogous one in which right-properness and suffixes are involved. On the other hand, taboo-sets induced by restriction enzymes remain invariant when we interchange every recognition site by its complementary sequence. Therefore, note that, for a bacterial taboo-set $\mathbb{T}$, if we prove that every graph $\Gamma_n^s(\mathbb{T})$ is connected, then also every graph ${}^s\Gamma_n(\mathbb{T})$ is connected.

2.9.1 A frequent case: *Turneriella parva*

The *Turneriella parva* (REBASE organism number 8970) strain produces a restriction enzyme with recognition site $GATC$, an inverted repeat. Similarly, another of its enzymes has recognition sites $GGACC$ and $GGTCC$. Thus, these restriction enzymes generate the taboo-set

$$\mathbb{T}_{T.pa} = \{GATC\} \bigcup \{GGACC, GGTCC\}. \tag{2.4}$$

Since $|\mathbb{T}_{T.pa}[1,1]| < 4$, Corollary 2.25.a implies that every graph $\Gamma_n^s(\mathbb{T}_{T.pa})$ is connected. Therefore the evolution of the DNA sequences can potentially reach any other taboo-free DNA sequence, no matter which suffix was conserved along this process.

Among the 3623 bacteria in REBASE [2020b], only 465 have more than three type II restriction enzymes. Assuming that only type II restriction enzymes induce taboos, as stated by Rusinov et al. [2015, 2018a], Corollary 2.25.b implies that at least 87% (3158/3623) of bacterial taboo-sets in REBASE [2020b] yield connected taboo-free Hamming graphs. Similarly, at least 90% (139/153) of archea in REBASE [2020a] induce connected taboo-free Hamming graphs, because they have less than four type II restriction enzymes. The following example describes a more complex collection of restriction enzymes.

2.9.2 *Helicobacter pylori*

In *H. pylori* 21-A-EK1, studied by Ailloud et al. [2019], many restriction enzymes have been identified. For the sake of clarity, let us write $\mathbb{T}_{H.py} = {}^A\mathbb{T} \bigcup {}^G\mathbb{T} \bigcup {}^C\mathbb{T} \bigcup {}^T\mathbb{T}$, where ${}^a\mathbb{T}$ denotes those taboos in $\mathbb{T}_{H.py}$ whose **first** symbol is $a \in \Sigma$. Then we have

$$
\begin{aligned}
{}^{A}\mathbb{T} =& \{AC \circ \Sigma \circ GT\}, \\
{}^{G}\mathbb{T} =& (GT \circ \Sigma^2 \circ AC) \bigcup \{GTCAC, GTGAC\} \bigcup \\
& \bigcup \{GTAC, GAGG\} \\
{}^{C}\mathbb{T} =& \{CCGG, CCTC, CATG\}, \\
{}^{T}\mathbb{T} =& \{TGCA\},
\end{aligned}
\tag{2.5}
$$

where $GT \circ \Sigma^2 \circ AC$ represents taboos of the type $GTabAC$ with $a, b \in \Sigma$, and so on for analogous notations.

We want to apply Prop. 2.24. Take any $r_1, r_2 \in \Psi(\mathbb{T}_{H.py})$ and assume that they are **not** left 1-synchronized. In particular WLOG we can assume that $T \notin L^1(r_1)$, implying $r_1 = GCA$. If $C \notin L^1(r_1)$, then $r_1 \in \{CGG, CTC, ATG\}$, which contradicts $r_1 = GCA$. Therefore it must be $C \notin L^1(r_2)$, yielding $r_2 \in \{CGG, CTC, ATG\}$. In any case, $d(r_1, r_2) \geq 2$. Thus, for any $w_1, w_2 \in \Psi(\mathbb{T})$ with $d\big(\, w_1[1, |w_2|] \,,\, w_2 \,\big) \leq 1$, it holds that w_1 and w_2 are left 1-synchronized, so Prop. 2.24 can be applied: Every graph $\Gamma_n^s(\mathbb{T}_{H.py})$ is connected and, in particular, $\Gamma_n(\mathbb{T}_{H.py})$ is connected.

2.9.3 *An imaginary bacterium*

The taboo-set can significantly influence evolution in the cases where some $\Gamma_n^s(\mathbb{T})$ is disconnected. To explain this, we will create a plausible, nonexistent example. Suppose that a strain of *Bacterium imaginara* has taboo-set

$$
\mathbb{T}_{B.im} = \{ACCC, TCCC, CGCC, GGCC\} \bigcup \{GGGT, GGGA, GGCG\},
$$

where the second set contains the complementary DNA sequences of the first set, except that of $GGCC$, which is an inverted repeat. Thus, taboo-set $\mathbb{T}_{B.im}$ is induced by 4 restriction enzymes. At first glance, taboo-set $\mathbb{T}_{B.im}$ seems less restrictive than $\mathbb{T}_{H.py}$, which has 6 taboos of length four and 22 taboos of length five or more.

Proposition 2.24 cannot be applied because CCC and GCC are not left 1-synchronized, and actually we can find a disconnected suffix graph. Let us take $V_n^{CCC}(\mathbb{T}_{B.im})$, which satisfies

$$
\begin{aligned}
& V_n^{CCC}(\mathbb{T}_{B.im}) = V_n^{GCCC}(\mathbb{T}_{B.im}) \bigcup V_n^{CCCC}(\mathbb{T}_{B.im}) = \\
& \Big(V_n^{AGCCC}(\mathbb{T}_{B.im}) \bigcup V_n^{TGCCC}(\mathbb{T}_{B.im})\Big) \bigcup \Big(V_n^{GCCCC}(\mathbb{T}_{B.im}) \bigcup V_n^{CCCCC}(\mathbb{T}_{B.im})\Big),
\end{aligned}
$$

implying that, for any strings $s_1 \in V_n^{GCCC}(\mathbb{T}_{B.im})$ and $s_2 \in V_n^{CCCC}(\mathbb{T}_{B.im})$, it holds that $d(s_1, s_2) \geq 2$. Thus, we found two disconnected components in $\Gamma_n^{CCC}(\mathbb{T}_{B.im})$, namely $\Gamma_n^{GCCC}(\mathbb{T}_{B.im})$ and $\Gamma_n^{CCCC}(\mathbb{T}_{B.im})$. All in all, the graph $\Gamma_n^{CCC}(\mathbb{T}_{B.im})$ is disconnected for $n \geq 5$.

This produces the following evolutionary implications: Assume that we have two correctly aligned DNA fragments f_α and f_β of the genome of *Bacterium imaginara*. Assume moreover that we can write $f_\alpha = r_\alpha GCCC$ and $f_\beta = r_\beta CCCC$ for some strings r_α and r_β, as also that the suffix CCC is invariable due to functional constrains. Then f_α cannot have evolved from f_β by simple point mutations, because at some point in evolution a taboo string is produced that is lethal for the carrier. Thus, the standard models of sequence evolution [Strimmer and von Haeseler, 2009] do not apply.

2.10 Concluding remarks

Using the results proven in this work, it is possible to decide whether every Hamming graph $\Gamma_n^s(\mathbb{T})$ is connected. The connectivity of the taboo-free Hamming graphs induced by the restriction enzymes of the bacteria listed in REBASE could be quickly analysed with our tools. Unfortunately, for many organisms listed in REBASE, the recognition sites of restriction enzymes are not available.

Based on the current version of REBASE [2020b], we conclude using Corollary 2.25 that taboo-sets of at least 87% (3158/3623) of bacteria in REBASE induce connected taboo-free Hamming graphs, because they have less than four type II restriction enzymes. For larger taboo-sets, Prop. 2.24 can be used, as we did in Subsection 2.9.2, or one can directly use the characterization of Theorem 2.22. Thus, restriction enzymes in bacteria generally do not lead to any disconnected taboo-free Hamming graph, and our models of sequence evolution are by and large applicable. However, the influence of some missing sequences in the Hamming graph on the estimation of evolutionary parameters deserves further investigations. We also would like to emphasize that still many recognition sites have to be identified, and thus it may be well possible that we find disconnected taboo-free Hamming graphs in the next future.

We consider the formal framework developed in this paper as a first and necessary step to understand the effect of restriction enzymes (and possibly other taboo sequences) on the DNA composition of bacteria and viruses, or more generally on the sequence space modelled as a Hamming graph. Consider, for example, the phylogenetic studies by Ailloud et al. [2019], where the *H. pylori* taboo-set $\mathbb{T}_{H.py}$ of

Subsection 2.9.2 was taken from. The following natural questions arise: How are inferred evolutionary times between the two *H. pylori* populations affected by $\mathbb{T}_{H.py}$? Has their *GC* content varied due to the taboos of restriction enzymes?

To answer such questions, we need to develop models of sequence evolution that take taboos into account. Taboo avoidance induces complex dependencies along a DNA sequence, which can be measured using Markov Chain Monte Carlo (MCMC) simulations. If all taboo-free Hamming graphs $\Gamma_n^s(\mathbb{T})$ are connected, then MCMC methods are easy to apply [Manuel et al., unpublished]. A disconnected taboo-free Hamming graph, however, leads to a reducible Markov chain, which complicates simulation of taboo-free evolution.

Another application of our framework is the construction of combinations of restriction enzymes that lead to a disconnected Hamming graph, and thus limit evolutionary freedom. This may help to efficiently treat viral infections. Some progress has been made in the usage of restriction enzymes for the treatment of viral infections [Weber et al., 2014]. Since one or just a few SNPs can significantly alter the symptoms or even the mortality associated to a pathogen [Collery et al., 2017, Yuan et al., 2017], our characterization of the connectivity of taboo-free Hamming graphs could help to delete SNPs from the viral genome that are detrimental to humans. Although the treatment of an infection using restriction enzymes is mostly unexplored, this work could be a first theoretical guide to a successful treatment.

References in this chapter

F. Ailloud, X. Didelot, S. Woltemate, G. Pfaffinger, J. Overmann, R. C. Bader, C. Schulz, P. Malfertheiner, and S. Suerbaum. Within-host evolution of Helicobacter pylori shaped by niche-specific adaptation, intragastric migrations and selective sweeps. *Nature Communications*, 10(1):2273, 2019. ISSN 2041-1723. doi: 10.1038/s41467-019-10050-1.

B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. Watson. *Molecular Biology of the Cell*, chapter 8, pages 532–534. Garland, 5th edition, 2004.

A. Asinowski, A. Bacher, C. Banderier, and B. Gittenberger. *Analytic Combinatorics of Lattice Paths with Forbidden Patterns: Enumerative Aspects*, pages 195–206. ., 01 2018. doi: 10.1007/978-3-319-77313-1-15.

A. Asinowski, A. Bacher, C. Banderier, and B. Gittenberger. Analytic combinatorics of lattice paths with forbidden patterns, the vectorial kernel method, and generating functions for pushdown automata. *Algorithmica*, pages 1–43, 10 2019. doi: 10.1007/s00453-019-00623-3.

M. M. Collery, S. A. Kuehne, S. M. McBride, M. L. Kelly, M. Monot, A. Cockayne, B. Dupuy, and N. P. Minton. What's a SNP between friends: The influence of single nucleotide polymorphisms on virulence and phenotypes of Clostridium difficile strain 630 and derivatives. *Virulence*, 8(6):767–781, Aug 2017. ISSN 2150-5608. doi: 10.1080/21505594.2016.1237333.

W. M. Fitch and E. Margoliash. A method for estimating the number of invariant amino acid coding positions in a gene using cytochrome c as a model case. *Biochemical Genetics*, 1(1):65–71, Jun 1967. ISSN 1573-4927. doi: 10.1007/BF00487738.

M. Gelfand and E. Koonin. Avoidance of palindromic words in bacterial and archaeal genomes: A close connection with restriction enzymes. *Nucleic acids research*, 25: 2430–9, 07 1997. doi: 10.1093/nar/25.24.0.

W. J. Hsu and M. J. Chung. Generalized Fibonacci Cubes. *1993 International Conference on Parallel Processing - ICPP'93*, 1:299–302, Aug 1993. doi: 10.1109/ICPP.1993.95.

A. Ilić, S. Klavžar, and Y. Rho. Generalized Fibonacci cubes. *Discrete Mathematics*, 312:2 – 11, 2012.

S. Klavžar. Structure of Fibonacci cubes: a survey. *Journal of Combinatorial Optimization*, 25:505–522, 2013. doi: 10.1007/s10878-011-9433-z.

V. Kommireddy and V. Nagaraja. Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiology and molecular biology reviews : MMBR*, 77:53–72, 03 2013. doi: 10.1128/MMBR.00044-12.

C. Manuel and A. von Haeseler. Structure of the space of taboo-free sequences. *Journal of Mathematical Biology*, 81(4):1029–1057, 2020.

C. Manuel, S. Pfannerer, and A. von Haeseler. Etaboo: Modelling and measuring taboo-free evolution. Unpublished, unpublished.

REBASE. The Restriction Enzyme Database: Archea. http://rebase.neb.com/rebase/arcbaclistA.html, 2020a. Accessed: 2020-06-17.

REBASE. The Restriction Enzyme Database: Bacteria. http://rebase.neb.com/rebase/arcbaclistB.html, 2020b. Accessed: 2020-06-17.

R. J. Roberts, T. Vincze, J. Posfai, and D. Macelis. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Research*, 43(D1):D298–D299, 11 2014. ISSN 0305-1048. doi: 10.1093/nar/gku1046.

E. Rocha, A. Danchin, and A. Viari. Evolutionary role of restriction/modification systems as revealed by comparative genome analysis. *Genome Research*, 11, 06 2001. doi: 10.1101/gr.GR-1531RR.

I. Rusinov, A. Ershova, A. Karyagina, S. Spirin, and A. Alexeevski. Lifespan of restriction-modification systems critically affects avoidance of their recognition sites in host genomes. *BMC Genomics*, 16(1):1084, 2015. ISSN 1471-2164. doi: 10.1186/s12864-015-2288-4.

I. S. Rusinov, A. S. Ershova, A. S. Karyagina, S. A. Spirin, and A. V. Alexeevski. Avoidance of recognition sites of restriction-modification systems is a widespread

but not universal anti-restriction strategy of prokaryotic viruses. *BMC Genomics*, 19(1):885, 2018a. ISSN 1471-2164. doi: 10.1186/s12864-018-5324-3.

I. S. Rusinov, A. S. Ershova, A. S. Karyagina, S. A. Spirin, and A. V. Alexeevski. Comparison of methods of detection of exceptional sequences in prokaryotic genomes. *Biochemistry (Moscow)*, 83(2):129–139, 2018b. ISSN 1608-3040. doi: 10.1134/S0006297918020050.

P. Sanders and C. Schulz. High quality graph partitioning. *Proceedings of the 10th DIMACS implementation challenge workshop*, 03 2013. doi: 10.1090/conm/588.

J. S. Shoemaker and W. M. Fitch. Evidence from nuclear sequences that invariable sites should be considered when sequence divergence is calculated. *Molecular Biology and Evolution*, 6(3):270–289, 05 1989. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a040550.

K. Strimmer and A. von Haeseler. Genetic distances and nucleotide substitution models. In P. Lemey, M. Salemi, and V. Anne-Mieke, editors, *The Phylogenetic Handbook: a Practical Approach to Phylogenetic Analysis and Hypothesis Testing*, pages 111–141. Cambridge University Press, Cambridge, 2 edition, 2009. doi: 10.1017/CBO9780511819049.006.

D. W. Ussery, T. M. Wassenaar, and S. Borini. *Computing for Comparative Microbial Genome: Bioinformatics for Microbiologists*. Springer., 1st edition, 2008. ISBN 978-1-84800-254-8.

N. D. Weber, M. Aubert, C. H. Dang, D. Stone, and K. R. Jerome. DNA cleavage enzymes for treatment of persistent viral infections: recent advances and the pathway forward. *Virology*, 454-455:353–361, Apr 2014. doi: 10.1016/j.virol.2013.12.037.

R. J. Wilson. *Introduction to Graph Theory*. John Wiley & Sons, Inc., New York, NY, USA, 1986. ISBN 0-470-20616-0.

L. Yuan, X.-Y. Huang, z.-y. Liu, F. Zhang, X. L. Zhu, J.-Y. Yu, X. Ji, Y. Xu, G. Li, C. Li, H.-J. Wang, Y.-Q. Deng, M. Wu, M.-L. Cheng, Q. Ye, D.-Y. Xie, X.-F. Li, X. Wang, W. Shi, and C.-F. Qin. A single mutation in the prM protein of Zika virus contributes to fetal microcephaly. *Science*, 358, 09 2017. doi: 10.1126/science.aam7120.

Chapter 3

New Measures of Phylogenetic Information Allow to Test for Saturation

Publication history and status

Manuscript in preparation. Authors: Cassius Manuel and Arndt von Haeseler.

Supplementary material: We include 8 Mathematica notebooks and two text files respectively containing a two-sequence and a five-sequence SIV alignment, available at `https://ucloud.univie.ac.at/index.php/s/U9b7ynPA0eEGNtc`.

Abstract

We introduce two new measures of information in a phylogenetic tree: The coherence of a branch, which quantifies the dependence between the two clades split by the branch, and the memory of a clade, which quantifies the identification of the parent node of a clade. The interplay between these measures is described assuming a stationary and reversible evolutionary process.

To apply these measures, we present the problem of substitution saturation in a phylogeny. We define branch saturation as a statistical hypothesis and propose the asymptotic test, whose statistic is based on the coherence. As a practical example, we show that long branch repulsion can be resolved using the asymptotic test.

3.1 Introduction

After the work of Shannon [1948], the entropy or "amount of choice" was accepted as a useful measure of the information carried by a message. As a familiar example, a receiver of a message may quantify the information received in bits, which is a unit of entropy. In phylogenetics, each observed species in a rooted phylogeny indeed receives a message (its orthologous nucleotide sequence), although our purpose is to reconstruct what we *cannot* observe, namely the ancestral sequences and the evolutionary process. This fundamental difference with respect to conventional communication suggests that we can build a more specific theory of information to assess phylogenetic reconstruction.

Since ancestral sequences cannot be observed, here we use the likelihood of each nucleotide as a proxy for the identity of the site. Then, instead of the entropic "amount of choice", the basis of our description of phylogenetic information is the "amount of identification" of an ancestral sequence, which we call **memory**. Moreover, given two adjacent nodes A and B of the phylogenetic tree, we define **the coherence** of branch AB as the "amount of dependence" between the two ancestral sequences at nodes A and B.

The memory and the coherence are not a direct function of the given multiple sequence alignment, because they are model-dependent. We assume that sites mutated independently as determined by a given reversible rate matrix Q and a phylogenetic tree. This framework corresponds to the typical workflow of current software implementations of phylogenetic reconstruction using a Maximum Likelihood Estimate (MLE) [Encyclopedia of Mathematics, 2021b], such as IQ-TREE [Minh et al., 2020] and RaxML [Stamatakis, 2014].

To show the usefulness of these measures, we employ them to construct a convenient test for phylogenetic saturation. Intuitively, saturation is the occurrence of too many mutations in an alignment as to provide information about its evolutionary history (cfr. Strimmer and von Haeseler [2009], Salemi [2009]). Starting from Archie [1989], the first saturation tests were based on parsimonious reconstruction, which we do not consider here.

Later on, Xia et al. [2003] proposed an entropy-based index of substitution saturation that is applied to an alignment using the observed state distribution at each site, recently used by Duchêne et al. [2021]. This index does not assume any rate matrix or phylogenetic tree, being in this aspect less parameterized than our asymptotic test. However, lacking a tree-like structure, the saturation index cannot judge whether some subset of the aligned species is informative.

By construction, the sample coherence of a branch is normally distributed and its expected value decreases exponentially as the branch length increases. This simple behaviour offers a useful tool to test for saturation, which can be viewed as an statistical description of mixing. Probabilistic descriptions of the mixing of Markov chains can be found e.g. in Levin and Peres [2017].

This work is structured as follows: We provide some introductory examples for practitioners in Section 3.3. In Section 3.4, we formally introduce the elements of an evolutionary process. Then, in Section 3.5, we reexplain the well-known computation of likelihoods in a phylogenetic tree. In Section 3.6, we define our new measures of phylogenetic information, namely the memory of a clade and the coherence of a branch. These measures are eigendecomposed in Section 3.7.

Our core results describing the expectation and variance of the coherence are stated in Section 3.8. The coherence emerges as a useful tool for asymptotic analysis of the log-likelihood in Section 3.9. The asymptotic test for saturation is described in Section 3.10 and employed in Section 3.11 to resolve long branch repulsion (LBR).

3.2 Results

Here we describe the memory of a clade and the coherence of a branch assuming stationarity and reversibility. We then employ these measures to construct the so-called asymptotic test for saturation. Our asymptotic test decides whether a branch is saturated, that is, if too many mutations occurred along the tested branch as to provide information.

Notably, the critical value of the asymptotic test can be computed analytically and is easy to estimate (Eq. 3.68). Moreover, the power of the asymptotic test is optimal for phylogenies close to saturation (Subsection 3.10.3).

As an application, in Section 3.11 we show that the asymptotic test can detect Long Branch Repulsion (LBR), a problematic phenomenon where a tree is wrongly reconstructed due to saturation. In particular, for each group of simulations with sequence length $n = 5000$, the amount of wrong reconstructions after applying the asymptotic test always stayed below 4%.

Moreover, we obtained a first moment estimate of the branch length between two nodes on a tree (Eq. 3.45), so far lacking in the literature. This estimate, as also its faster approximation of Eq. 3.46, can be used by practitioners aiming to accelerate or even avoid the iterative Newton method [Encyclopedia of Mathematics, 2021a] to compute the MLE of the branch length.

Along the way to expose these results, we made other interesting observations for phylogenetists: The likelihood of a branch length can have multiple maxima, and thus the Newton method can output a suboptimal maximum (Appendix 3.E); The divergence of the Newton method can be predicted using the dominant sample coherence (Section 3.9, Cor. 3.5); A finite MLE is uninformative about the finiteness of the true branch length (Section 3.D).

3.3 Examples of the Asymptotic Test

In Section 3.10, we formally introduce the concept of saturation and propose a test for branch saturation. Although the theory leading to the asymptotic test may be too technical for practitioners, its usage is relatively simple, as the examples presented in this section will show.

The numerical analysis of Subsection 3.3.1 is included Supp. Notebook 7, while those of Subsections 3.3.2 and 3.3.3 can be found in Supp. Notebook 8.

3.3.1 Saturation between two sequences

We are given an alignment of two DNA sequences $\mathbf{y}$ and $\mathbf{z}$ of length $n = 100$, which we suspect that mutated from a common ancestor a very long time ago. We arrange the alignment of sequences $\mathbf{y}$ and $\mathbf{z}$ as a 4×4 matrix $N = (n_{ij})$, where n_{ij} is the number of sites where we observe nucleotide i in sequence $\mathbf{y}$ and nucleotide j in sequence $\mathbf{z}$. Following the order A, C, G, T, say that we obtain

$$N = \begin{pmatrix} 5 & 4 & 7 & 5 \\ 4 & 11 & 7 & 7 \\ 9 & 9 & 7 & 8 \\ 8 & 4 & 1 & 4 \end{pmatrix}$$

Somehow we infer that the model used to generate the alignment was the K80 model with transition-transversion ratio $k = 2$ [Kimura, 1980]. Thus the rate matrix has the form

$$Q = \frac{1}{4} \begin{pmatrix} * & 1 & 2 & 1 \\ 1 & * & 1 & 2 \\ 2 & 1 & * & 1 \\ 1 & 2 & 1 & * \end{pmatrix},$$

where we divide by 4 to make the largest nonzero eigenvalue of Q be $\lambda_1 = -1$, which has multiplicity one and right eigenvector $\boldsymbol{v_1} = (-1, 1, -1, 1)^T$.

We ignore the evolutionary time t^* between $\mathbf{y}$ and $\mathbf{z}$. In particular, we even ignore if $t^* \to \infty$, or equivalently, if sequences $\mathbf{y}$ and $\mathbf{z}$ were sampled at random independently and have no evolutionary history in common. If we cannot reject the null hypothesis $t^* \to \infty$, we say that branch $\mathbf{yz}$ is saturated. With this setup, is branch $\mathbf{yz}$ saturated?

First of all, we need to choose a significance level, say $\alpha = 0.05$. That means that we are willing to wrongly reject $t^* \to \infty$ in 5% of cases. In a standard normal distribution, the centile covering a 95% probability from $-\infty$ to z_α is $z_\alpha \approx 1.6$. Now we compute the so-called dominant sample coherence of branch $\mathbf{yz}$ (see Eq. 3.63 and Prop. 3.6.a) as

$$\hat{\delta} = \hat{C}_1(\mathbf{y}; \mathbf{z}) = \boldsymbol{v_1}^T N \boldsymbol{v_1}/n = 8/100 = 0.08.$$

The critical value of the asymptotic test is $c_S \approx z_\alpha/\sqrt{n} \approx 0.16$ (Eq. 3.75). That is, if $\hat{\delta} > c_S$, then we reject $t^* \to \infty$. Since $0.08 \not> 0.16$, we cannot reject $t^* \to \infty$ and conclude that branch $\mathbf{yz}$ is saturated with significance $\alpha = 0.05$.

The main consequence of the saturation of branch $\mathbf{yz}$ is that we should not try to compute an estimate $\hat{t}$ of t^*. Any estimate $\hat{t}$ is unreliable, since the given data cannot reject that sequences $\mathbf{y}$ and $\mathbf{z}$ were sampled at random.

The matrix N used here was generated from an actual simulation where we set true time $t^* = 3$. A natural question is which minimum true time t_S (called saturation time) leads to an expected alignment where we remain in hypothesis $t^* \to \infty$. Since $\lambda_1 = -1$, we know that $\mathbb{E}[\hat{\delta}] = e^{-\lambda_1 t^*} = e^{-t^*}$ (Prop. 3.6.b). Thus the saturation time is the solution of equation $e^{-t_S} = c_S \approx 0.16$, giving $t_S \approx 1.8$. We plot t_S as a function of the sequence length n in Figure 3.1.

3.3.2 Saturation of an external branch

In Supp. File `SIV_5_species.txt` we have a DNA alignment of length 3265 of the ENV gene of 5 SIV samples, obtained from LANL [2020]. Due to their high rate of mutation, phylogenies of SIV tend to vary depending on the data used for the analysis [Salemi, 2009]. Using the asymptotic test, we can test if an alignment region is not supporting a particular branch of the alignment. We will first focus on the longest reconstructed branch.

We use IQ-TREE [Minh et al., 2020] with a GTR model to reconstruct the

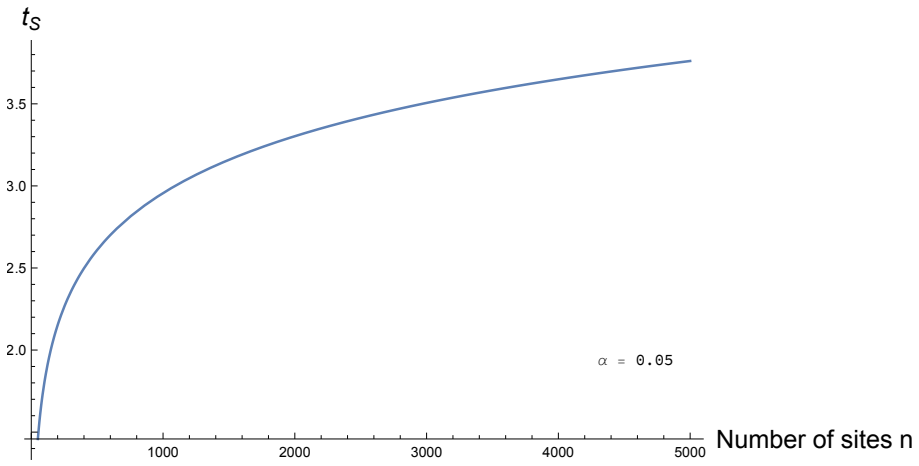

Figure 3.1: Plot of the saturation time t_S as a function of the number of sites n when $\lambda_1 = -1$ has multiplicity one and $\alpha = 0.05$. Since $t_S = -\log(z_\alpha/\sqrt{n})$, the saturation time grows linearly on the logarithm of the sequence length.

phylogeny of these SIV species, assuming moreover a Gamma model with 4 rates [Gu et al., 1995, Yang, 1994]. This implies that the evolutionary process starts by assigning to each site one of 4 possible average rates of mutation. For each site, the branch lengths of the original evolutionary tree are multiplied by their assigned rate.

Attending to the MLE, the largest of this rates is $f_4 = 2.3$. The reconstructed phylogenetic tree multiplied by rate f_4 is shown in Figure 3.2. Then we use maximum likelihood to estimate which of the 3265 sites had rate f_4, giving an alignment region of length $n = 1008$. A necessary condition for this region to support the branch between node A and SIV5 is a significant rejection of the hypothesis that subsequence SIV5 was sampled independently from the rest of the alignment region. Equivalently, we must reject hypothesis $t^* \to \infty$, where t^* is the true time length t^* between node A and sequence SIV5.

We choose significance level $\alpha = 0.05$, giving the centile $z_\alpha \approx 1.6$. Assume that the largest nonzero eigenvalue of Q has multiplicity one with right eigenvector $\boldsymbol{v_1} = (v_1^A, v_1^C, v_1^G, v_1^T)$ and left eigenvector $\boldsymbol{h_1}$.

When considered independently, each site of the given alignment determines a pattern ∂, which are strings of length m, with one nucleotide per aligned sequence (see e.g. Table 3.1). As explained in Section 3.5, branch $A\mathbf{y}$ splits each pattern ∂ into two subpatterns: the observed nucleotide i at sequence $\mathbf{y}$, and the rest of the pattern, say ∂A. Given ∂A, we can compute the normalized likelihood vector $\tilde{\boldsymbol{\alpha}}_{\partial A}$ as exemplified in Figure 3.3. If pattern $\partial = (\partial A, i)$ is observed n_∂ times, we compute

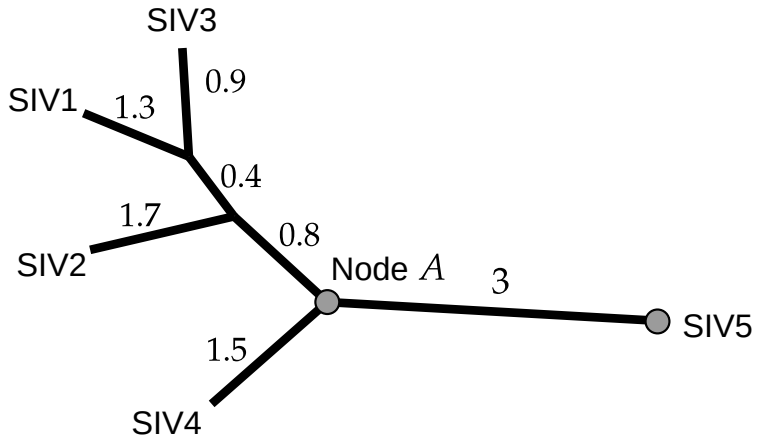

Figure 3.2: Reconstructed phylogenetic tree of the 5 SIV species for the region with estimated rate category f_4.

the dominant sample coherence of branch A$\mathbf{y}$ (see Equations 3.63 and 3.37) as

$$\hat{\delta} = \hat{C}_1(A; \mathbf{y}) := \sum_{\partial} \frac{n_\partial}{n} (\tilde{\boldsymbol{\alpha}}_{\partial A} \cdot \boldsymbol{h_1}) v_1^i. \tag{3.1}$$

In general, we also need to compute the 11-projection of the sample memory at node A (Eq. 3.43), which is $\hat{M}_{11}(A) \geq 0$. Assuming a large n and using Eq. 3.73, the critical value of the asymptotic test can be approximated as

$$c_S \approx \frac{z_\alpha}{\sqrt{n}} \sqrt{\hat{M}_{11}(A)} \approx 0.05 \sqrt{\hat{M}_{11}(A)}. \tag{3.2}$$

That is, if $\hat{\delta} > c_S$, then we reject $t^* \to \infty$. We compute $\hat{\delta} \approx -0.05$, while $c_S > 0$ by construction. Since $\hat{\delta} < c_S$, we cannot reject $t^* \to \infty$ and conclude that, for the region with reconstructed rate f_4, the branch connecting node A and SIV5 is saturated with significance $\alpha = 0.05$. In informal terms, the region with reconstructed rate f_4 has not collaborated in the reconstruction of this branch. In this particular case, since $\hat{\delta} < 0$, any level of significance α leads to the same conclusion.

3.3.3 Saturation of an internal branch

In the previous subsection we could see that the asymptotic test does not depend on the reconstructed length of the tested branch, in that case $\hat{t} = 3$. Consequently, although it may seem intuitive that only long branches lead to branch saturation, this is actually a wrong intuition, as clarified by the formalism of the asymptotic

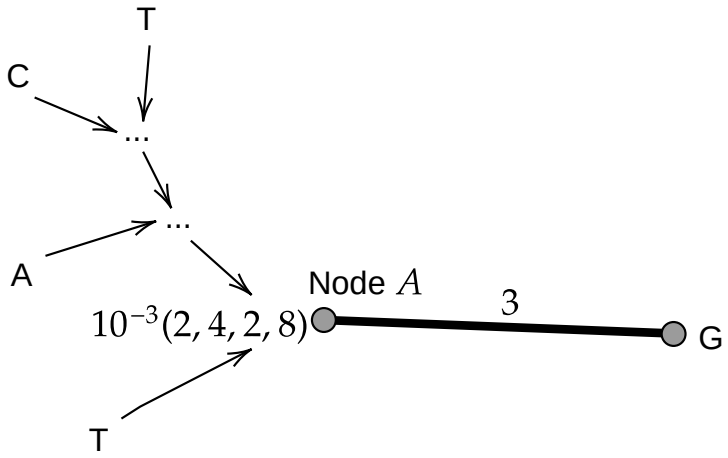

Figure 3.3: Diagram of the computation of the likelihood vector $\boldsymbol{\alpha_{\partial A}}$ at node A given pattern $\partial = CATTG$, inducing subpattern $\partial A = CATT$. For a given site, we recursively compute the likelihood vector at deeper nodes using Eq. 3.13. Finally we obtain $\boldsymbol{\alpha_{\partial A}} = 10^{-3}(2,4,2,7)$. The equilibrium distribution of the rate matrix is $\boldsymbol{\pi} \approx (0.33, 0.19, 0.23, 0.24)$, and the normalized likelihood vector $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} = \boldsymbol{\alpha_{\partial A}}/(\boldsymbol{\alpha_{\partial A}} \cdot \boldsymbol{\pi}) \approx (0.6, 1, 0.5, 2)$. See Section 3.5 for more details.

test. To see an example, consider the shortest branch of the SIV phylogeny of Figure 3.2, which has length $\hat{t} = 0.4$ and induces the split $13|245$ between the SIV species. From now on, we refer to the parent node of clade 13 as node A, while the parent node of clade 245 will be node B.

Again we choose significance level $\alpha = 0.05$, giving the centile $z_\alpha \approx 1.6$, and consider the same rate matrix Q with left eigenvector $\boldsymbol{h_1}$. Branch AB splits each pattern ∂ into subpatterns ∂A and ∂B, as described in Figure 3.4. Given ∂A, we can compute the normalized likelihood vector $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}$ as exemplified in Figure 3.4, and proceed analogously with pattern ∂B. If pattern $\partial = (\partial A, \partial B)$ is observed n_∂ times, we compute the dominant sample coherence of branch AB (see Equations 3.63 and 3.37) as

$$\hat{\delta} = \hat{C}_1(A;B) := \sum_{\partial} \frac{n_\partial}{n} (\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} \cdot \boldsymbol{h_1})(\tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} \cdot \boldsymbol{h_1}). \tag{3.3}$$

In general, we need to compute the 11-projections of the sample memories $\hat{M}_{11}(A)$ and $\hat{M}_{11}(B)$ (Eq. 3.43), and the saturation coherence is

$$c_S \approx \frac{z_\alpha}{\sqrt{n}} \sqrt{\hat{M}_{11}(A)\hat{M}_{11}(B)} \approx 0.05 \sqrt{\hat{M}_{11}(A)\hat{M}_{11}(B)}. \tag{3.4}$$

In this particular case, however, we compute $\hat{\delta} = -0.03$. Since $c_S > 0$ by con-

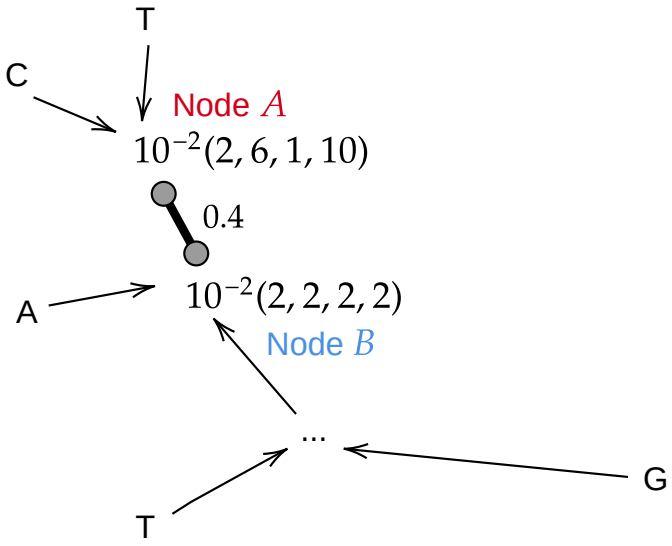

Figure 3.4: Diagram of the computation of the likelihood vectors $\boldsymbol{\alpha}_{\partial A}$ at node A and $\boldsymbol{\beta}_{\partial B}$ at node B given pattern $\partial = CATTG$. Branch AB induces the split $13|245$ and thus subpatterns $\partial A = CT$ and $\partial B = ATG$. For a given site, we recursively compute the likelihood vector at deeper nodes using Eq. 3.13. After computing $\boldsymbol{\alpha}_{\partial A}$ and $\boldsymbol{\beta}_{\partial B}$, the normalized likelihood vectors are $\tilde{\boldsymbol{\alpha}}_{\partial A} = \boldsymbol{\alpha}_{\partial A}/(\boldsymbol{\alpha}_{\partial A} \cdot \boldsymbol{\pi}) \approx (0.4, 1.3, 0.3, 2)$ and $\tilde{\boldsymbol{\beta}}_{\partial B} = \boldsymbol{\beta}_{\partial B}/(\boldsymbol{\beta}_{\partial B} \cdot \boldsymbol{\pi}) \approx (1.1, 1, 0.9, 1)$. Recall that the equilibrium distribution is $\boldsymbol{\pi} \approx (0.33, 0.19, 0.23, 0.24)$. See Section 3.5 for more theoretical details.

struction, it follows that $\hat{\delta} < c_S$, and thus branch AB is saturated with significance $\alpha = 0.05$. As we can see, the fact that branch AB is short did not play any role in its saturation status. The emergence of short internal branches in saturated alignments is clarified in Subsection 3.9.1.

Actually, performing the asymptotic test for all branches, we see that all branches of the region with reconstructed rate f_4 are saturated, because they have a negative dominant sample coherence. The intuition emerges that the IQ-TREE reconstruction method has grouped under rate f_4 all uninformative patterns. Consequently, this region can be ignored without significantly affecting the phylogeny, or equivalently we could just set $f_4 \to \infty$. All in all, the asymptotic test for saturation has recognized, in a systematic way, a region not providing significant information for the reconstructed phylogeny.

3.4 The Evolutionary Process and its Reconstruction

In a broad sense, an **evolutionary process** is a random variable that outputs sequences (strings over an alphabet $\mathbb{A}$ of possible states) using a model of evolution.

Here we only consider continuous Markov models of evolution, where each site of the ancestral sequence mutates independently as determined by a rate matrix Q and a tree with some branch lengths. Only substitutions are allowed under this model.

We will focus on processes assuming two properties: **Stationarity**, meaning that the prior state distribution is the unique equilibrium distribution $\boldsymbol{\pi} = (\pi_i)$ of matrix Q, defined by equation $\boldsymbol{\pi}^T Q = Q$ (see Prop. 3.1); and **reversibility**, meaning that the detailed balance equations $\Psi Q = Q^T \Psi$ hold, where $\Psi = \text{Diag}\,(\boldsymbol{\pi})$. In this work, italic bold letters as $\boldsymbol{\pi}$ always denote column vectors, while indices $i \in \mathbb{A}$ and $j \in \mathbb{A}$ always refer to states of the alphabet $\mathbb{A}$.

As an easy and important example, Figure 3.5 shows the evolutionary process E_{root} that outputs sequences $\mathbf{y}$ and $\mathbf{z}$ of length n as follows:

1) For each site $s \in [n] := \{1, \cdots, n\}$, sample a state $r_s \in \mathbb{A}$ independently, as determined by the equilibrium distribution $\boldsymbol{\pi}$ of Q. This yields the ancestral sequence $\mathbf{r} = r_1 \cdots r_n$.

2) Each state r_s mutates for t_1^* time units as determined by Q, meaning that the probability of state $i \in \mathbb{A}$ mutating to $j \in \mathbb{A}$ is $p_{ij}(t_1^*)$, where the transition matrix is the exponential matrix $(p_{ij}(t_1^*)) = e^{Qt_1^*}$ (see Prop. 3.1). This yields state y_s, forming sequence $\mathbf{y} = y_1 \cdots y_n$.

3) Analogously, generate each state z_s by mutating r_s for t_2^* time units as determined by Q. This yields sequence $\mathbf{z} = z_1 \cdots z_n$.

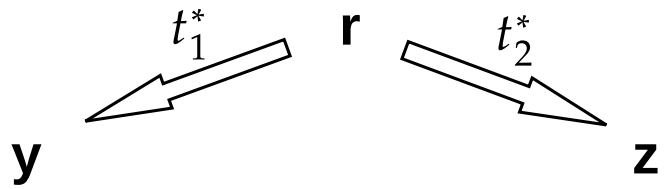

Figure 3.5: Diagram of the stationary and reversible evolutionary process E_{root}. Sequences $\mathbf{y}$ and $\mathbf{z}$ are the outcome of E_{root}, which were obtained by mutating the root $\mathbf{r}$ as determined by rate matrix Q.

The more general evolutionary process E on a tree with m leaves is constructed analogously [Felsenstein, 2004]: At the root node R, we sample a sequence $\mathbf{r}$ as determined by distribution $\boldsymbol{\pi}$. Sequence $\mathbf{r}$ evolves independently towards each child node of R, and we repeat this process until each of the m leaves receives a mutated sequence $\boldsymbol{l^c} = l_1^c \cdots l_n^c$, where $c \in [m]$.

A **realization of** E is the alignment output by the process, that is, the ordered set of sequences $(\boldsymbol{l^1}, \cdots, \boldsymbol{l^m})$ at the leaves. For example, a realization of E_{root} for $n = 3$ using the nucleotide alphabet $\mathbb{A} = \{A, C, G, T\}$ could be an alignment $(\mathbf{y}, \mathbf{z})$ as $\mathbf{y} = ACG$ and $\mathbf{z} = ACC$. The s'th site of the alignment $(\boldsymbol{l^1}, \cdots, \boldsymbol{l^m})$ is the string $l_s^1 \cdots l_s^m$, which belongs to set $\mathbb{A}^m$, that is, the set of strings of length m over alphabet $\mathbb{A}$. Each string $\partial \in \mathbb{A}^m$ is called a **pattern**.

If the process E is known, the probability that the alignment has pattern ∂ at site $s \in [n]$ is designated as $\Pr(\partial)$. Conversely, given pattern ∂ at site $s \in [n]$, the **likelihood** of process E is $\Pr(\partial \mid E)$. Considering the whole given alignment, if each pattern ∂ is observed n_∂ times, then the log-likelihood of process E is

$$L(E) = \sum_{\partial} n_\partial \log \Pr(\partial \mid E). \tag{3.5}$$

Normally the true process E^* that generated the data is (at least partially) unknown. To estimate E^*, we use a **maximum likelihood estimate** (MLE). The MLE is a process $\hat{E}$ with the highest log-likelihood among all processes, given the observed alignment.

Notably, if the rate matrix Q is reversible and the process is stationary, then the root R of the tree cannot be identified, in the sense that, for any process E, any replacement of the root R on the tree gives the same probability $\Pr(\partial \mid E)$ of observing pattern ∂. This property is called the Pulley principle, stated by Felsenstein [1981].

The Pulley principle implies the equality $\Pr(\partial \mid E_{root}) = \Pr(\partial \mid E_{seq})$, where process E_{seq}, shown in Figure 3.6, has sequence $\mathbf{y}$ as root and $t^* = t_1^* + t_2^*$. For future examples, we will focus on process E_{seq}, which is simpler than E_{root}.

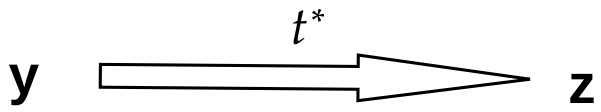

Figure 3.6: Diagram of the stationary and reversible evolutionary process E_{seq}. Sequences $\mathbf{y}$ and $\mathbf{z}$ are the outcome of E_{seq}. The root sequence $\mathbf{y}$ is sampled using distribution $\boldsymbol{\pi}$, while $\mathbf{z}$ is generated by mutating the root $\mathbf{y}$ as determined by rate matrix Q.

As an example, if the rate matrix Q is fixed, the MLE of the true distance t^* of process E_{seq} is obtained as follows. Let $ij \in \mathbb{A}^2$ denote the pattern where $\mathbf{y}$ has state i and $\mathbf{z}$ has state j. It holds that $\Pr(ij \mid t) = \pi_i p_{ij}(t)$, where $\boldsymbol{\pi} = (\pi_i)$

and $e^{Qt} = (p_{ij}(t))$. Given an alignment where pattern ij is observed n_{ij} times, the log-likelihood of distance t between sequences $\mathbf{y}$ and $\mathbf{z}$ is

$$L(t) = \sum_{i,j} n_{ij} \log(\pi_i p_{ij}(t)). \tag{3.6}$$

The MLE of t^* is a time $\hat{t} \in [0, \infty]$ where $L(t)$ reaches its absolute maximum.

3.5 Likelihood Vectors in a Phylogenetic Tree

In this section, we describe the likelihood vector, introduced by Felsenstein [1981] as a tool to compute the likelihood $\Pr(\partial \mid E)$.

Consider an evolutionary process E on a tree rooted at R. **The likelihood vector at R given a pattern** ∂ is a vector $\boldsymbol{\rho_\partial}$ of probabilities assuming each possible root identity; more formally, it is defined as

$$\boldsymbol{\rho_\partial} := (\rho^i_\partial) := (\Pr(\partial \mid i \text{ at node } R)), \tag{3.7}$$

where $i \in \mathbb{A}$ and we omitted the assumption of process E for simplicity. Since $\sum_\partial \Pr(\partial \mid i \text{ at node } R) = 1$ for all $i \in \mathbb{A}$, it follows that

$$\sum_\partial \boldsymbol{\rho_\partial} = \mathbf{1}, \tag{3.8}$$

where $\mathbf{1}$ is the column vector of 1's. Moreover, using the law of total probability, the likelihood of a process E given pattern ∂ is

$$\Pr(\partial \mid E) = \sum_i \Pr(i \text{ at node } R) \Pr(\partial \mid i \text{ at node } R) = \boldsymbol{\pi} \cdot \boldsymbol{\rho_\partial}, \tag{3.9}$$

where "$\cdot$" denotes the Euclidean dot product. Vector $\boldsymbol{\rho_\partial}$ allows to compute the likelihood $\Pr(\partial \mid E)$ more easily, because $\boldsymbol{\rho_\partial}$ can be expressed in terms of likelihood vectors induced by smaller subtrees. This method, described by Felsenstein [1981], works as follows.

Consider a branch AB where the root R is lying, as shown in Figure 3.7. We consider a split determined by branch AB, that is, we consider two subtrees rooted at nodes A and B, called **clades**, namely a clade A with k leaves and a clade B with the other $m - k$ leaves, as Figure 3.7 shows. Note that clades are named after their parent node. A pattern ∂ induces subpatterns ∂A and ∂B at the leaves of clades A and B, as exemplified in Table 3.1. Conversely, the subpatterns ∂A and ∂B

determine pattern ∂, because by construction all leaves are partitioned into those of clade A and those of clade B.

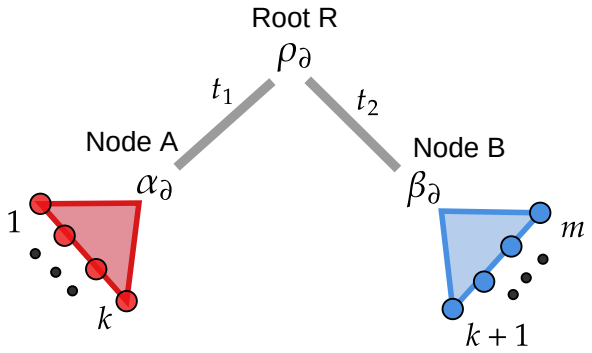

Figure 3.7: Diagram of the likelihood vectors at branch AB given ∂. Branch AB splits the phylogeny into clades A and B, with respectively k and $m-k$ leaves. Each of these leaves has a nucleotide identity determined by pattern ∂. Likelihood vectors $\boldsymbol{\alpha_\partial}$ and $\boldsymbol{\beta_\partial}$ depend only on their respective clades A and B. The likelihood vector $\boldsymbol{\rho_\partial}$ is computed using vectors $\boldsymbol{\alpha_\partial}$, $\boldsymbol{\beta_\partial}$ and branch lengths $t_1 \geq 0$ and $t_2 \geq 0$.

	Site 1	**Site 2**	
1	C	T	∂A
2	C	T	
3	G	T	∂B
4	G	T	
5	G	T	

Table 3.1: Alignment of 5 sequences with 2 sites. We assume that branch AB splits the phylogeny into a clade A with sequences 1 and 2 and a clade B with sequences 3, 4 and 5. At site 1, we observe pattern $\partial = CCGGG$, inducing subpatterns $\partial A = CC$ and $\partial B = GGG$. At site 2, we observe pattern $\partial = TTTTT$, inducing subpatterns $\partial A = TT$ and $\partial B = TTT$.

Now imagine that we can observe only pattern ∂A and the process consists only on clade A. We define the likelihood vector at node A given ∂A as

$$\boldsymbol{\alpha_{\partial A}} := (\alpha^i_{\partial A}) := \Pr(\partial A \mid \ i \text{ at node } A), \tag{3.10}$$

where we omit the assumed evolutionary process. Analogously, the likelihood vector at node B given ∂B is

$$\boldsymbol{\beta_{\partial B}} := (\beta^i_{\partial B}) := \Pr(\partial B \mid \ \ i \text{ at node } B). \tag{3.11}$$

Recall that the probability of state $i \in \mathbb{A}$ mutating to $j \in \mathbb{A}$ in time t is $p_{ij}(t)$,

where $e^{Qt} := (p_{ij}(t))$. Using the law of total probability, the likelihood vector at node R given only clade A is $e^{Qt_1}\boldsymbol{\alpha}_{\partial A}$, or more explicitly,

$$e^{Qt_1}\boldsymbol{\alpha}_{\partial A} = \Pr(\partial A \mid \ i \text{ at node } R). \tag{3.12}$$

Analogously, the likelihood vector at node R given only clade B is $e^{Qt_2}\boldsymbol{\beta}_{\partial B}$. We define the entrywise product "$\circ$" between two vectors $\boldsymbol{v} = (v_i)$ and $\boldsymbol{w} = (w_i)$ as $\boldsymbol{v} \circ \boldsymbol{w} := (v_i w_i)$. Since subpatterns ∂A and ∂B are a partition of ∂ and were obtained independently from the sequence at R, we have

$$\boldsymbol{\rho}_{\boldsymbol{\partial}} = (e^{Qt_1}\boldsymbol{\alpha}_{\boldsymbol{\partial A}}) \circ (e^{Qt_2}\boldsymbol{\beta}_{\boldsymbol{\partial B}}). \tag{3.13}$$

This equation can be used to recursively compute the likelihood $\Pr(\partial \mid E)$. However, in this work the likelihood vectors $\boldsymbol{\alpha}_{\boldsymbol{\partial A}}$ and $\boldsymbol{\beta}_{\boldsymbol{\partial B}}$ are not only a tool to compute $\Pr(\partial \mid E)$, but also of fundamental importance to understand the phylogenetic relation between clades A and B. In particular, let us compute the likelihood $\Pr(\partial \mid t)$ of the total branch length $t := t_1 + t_2$ between nodes A and B, represented in Figure 3.8. For simplicity, in likelihood $\Pr(\partial \mid t)$ we omit the assumption of the rest of process E.

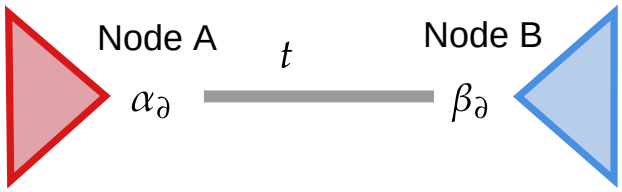

Figure 3.8: Diagram of the branch AB of length t, whose log-likelihood $\Pr(\partial \mid t)$ we want to compute.

Assuming a stationary and reversible rate matrix Q, the likelihood of the alignment is independent of the placement of the root, as implied by the Pulley Principle [Felsenstein, 1981]. Thus we place the root R on node A, and the likelihood vector $\boldsymbol{\rho}_{\boldsymbol{\partial}}$ can be inferred from Eq. 3.13 by setting $t_1 = 0$ and $t_2 = t$. This gives

$$\boldsymbol{\rho}_{\boldsymbol{\partial}} = \boldsymbol{\alpha}_{\boldsymbol{\partial A}} \circ (e^{Qt}\boldsymbol{\beta}_{\boldsymbol{\partial B}}). \tag{3.14}$$

Therefore, given pattern ∂, the likelihood of branch AB having length t equals

$$\Pr(\partial \mid t) = \boldsymbol{\pi} \cdot \boldsymbol{\rho}_{\boldsymbol{\partial}} = \boldsymbol{\pi} \cdot \left(\boldsymbol{\alpha}_{\boldsymbol{\partial A}} \circ (e^{Qt}\boldsymbol{\beta}_{\boldsymbol{\partial B}})\right) = \boldsymbol{\alpha}_{\boldsymbol{\partial A}}{}^{T} \Psi e^{Qt}\boldsymbol{\beta}_{\boldsymbol{\partial B}}, \tag{3.15}$$

where we set $\Psi := \mathrm{Diag}(\boldsymbol{\pi})$.

3.6 Measures of phylogenetic information

In this section, we present all definitions required to construct our measures of phylogenetic information, which are the coherence of a branch and the memory of a clade. The memory vector and its module as a measure of information were already introduced by Manuel [2022]. Basic bounds involving these measures can be found in Appendix 3.B.

3.6.1 The memory vector

All multiples $C\boldsymbol{\rho}_{\boldsymbol{\partial}}$ of the likelihood vector can be used to have an MLE of the ancestor identity at the root. To have a unique representative of each ray of likelihood vectors, we define **the normalized likelihood vector at** R **given** ∂ as

$$\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} := \frac{\boldsymbol{\rho}_{\boldsymbol{\partial}}}{\boldsymbol{\pi} \cdot \boldsymbol{\rho}_{\boldsymbol{\partial}}} = \frac{\boldsymbol{\rho}_{\boldsymbol{\partial}}}{\Pr(\partial)}, \tag{3.16}$$

where recall that $\boldsymbol{\pi} \cdot \boldsymbol{\rho}_{\boldsymbol{\partial}} = \Pr(\partial)$ assuming stationarity. This normalization is very convenient, because matrix e^{Qt} is an endomorphism of normalized likelihood vectors. To see this more clearly, consider an evolutionary process E with a tree rooted at node R with likelihood vector $\boldsymbol{\rho}_{\boldsymbol{\partial}}$ given pattern ∂. We can modify process E by making the ancestral sequence at R mutate for t additional time units, as shown in Figure 3.9. We say that the root of this modified process is $e^{Qt}R$, which has likelihood vector $e^{Qt}\boldsymbol{\rho}_{\boldsymbol{\partial}}$ given ∂. Our normalization is convenient because, when rooted at node $e^{Qt}R$, we have

$$\Pr(\partial \mid \text{Root at } e^{Qt}R) = \boldsymbol{\pi}^T e^{Qt}\boldsymbol{\rho}_{\boldsymbol{\partial}} = \boldsymbol{\pi}^T\boldsymbol{\rho}_{\boldsymbol{\partial}} = \boldsymbol{\pi} \cdot \boldsymbol{\rho}_{\boldsymbol{\partial}}, \tag{3.17}$$

and consequently the normalized likelihood vector at node $e^{Qt}R$ is $e^{Qt}\boldsymbol{\rho}_{\boldsymbol{\partial}}/(\boldsymbol{\pi} \cdot \boldsymbol{\rho}_{\boldsymbol{\partial}}) = e^{Qt}\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$. Therefore there is no need to renormalize after the action of e^{Qt}.

The normalized likelihood vector is closely related to the posterior probability after observing pattern ∂. Define the posterior distribution $\boldsymbol{r}_{\boldsymbol{\partial}} = (r^i_{\partial})$ as $r^i_{\partial} := \Pr(i \text{ at node } R \mid \partial)$, where $i \in \mathbb{A}$. Then, using Bayes' theorem and the stationary prior $\boldsymbol{\pi}$, we get $\boldsymbol{r}_{\boldsymbol{\partial}} = \boldsymbol{\pi} \circ \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$.

It is useful to know whether $\boldsymbol{\rho}_{\boldsymbol{\partial}}$ has nearly uniform entries, since uniformity implies that it is difficult to estimate the ancestor identity at R. Moreover, $\boldsymbol{\rho}_{\boldsymbol{\partial}}$ is

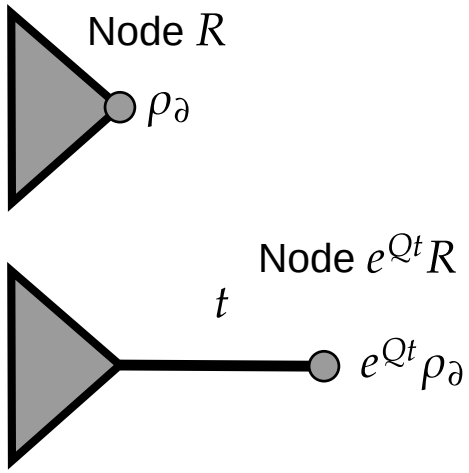

Figure 3.9: Diagram of a tree or clade whose parent node evolves t additional time units.

uniform iff $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = \mathbf{1}$, where $\mathbf{1}$ is the column vector of 1's.

Motivated by this observation, we define **the memory vector at** R **given** ∂ as $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}$. We can state analogous definitions for the clades of a tree. Given a branch AB inducing subpattern ∂A, if clade A has likelihood vector $\boldsymbol{\alpha}_{\boldsymbol{\partial A}}$, then $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} = \boldsymbol{\alpha}_{\boldsymbol{\partial A}} / \Pr(\partial A)$, where $\Pr(\partial A) = \boldsymbol{\pi} \cdot \boldsymbol{\alpha}_{\boldsymbol{\partial A}}$ assuming stationarity, and the memory vector at node A given ∂A is $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} - \mathbf{1}$. Similarly, the memory vector at node B given ∂B is $\tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} - \mathbf{1}$, where $\Pr(\partial B) = \boldsymbol{\pi} \cdot \boldsymbol{\beta}_{\boldsymbol{\partial B}}$.

3.6.2 The coherence of a branch

Given two vectors $\boldsymbol{v} = (v_i)$ and $\boldsymbol{w} = (w_i)$, we define their π-inner product as

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle_\pi := \boldsymbol{\pi} \cdot (\boldsymbol{v} \circ \boldsymbol{w}) = \boldsymbol{v}^T \Psi \boldsymbol{w} = \sum_i \pi_i v_i w_i. \tag{3.18}$$

Consider two adjacent nodes A and B on the tree, determining branch AB. We define **the coherence of branch** AB **given** ∂, denoted as $C^\partial(A;B)$, as the π-inner product between the memory vectors at nodes A and B, namely

$$C^\partial(A;B) := \langle \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} - \mathbf{1}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} - \mathbf{1} \rangle_\pi \tag{3.19}$$

Developing the inner product, we obtain the alternative form

$$C^\partial(A;B) = \langle \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} \rangle_\pi - 1, \tag{3.20}$$

where we used the fact that $\langle \tilde{\boldsymbol{\alpha}}_{\partial A}, \mathbf{1} \rangle_\pi = \langle \tilde{\boldsymbol{\beta}}_{\partial B}, \mathbf{1} \rangle_\pi = 1$. The coherence quantifies the dependence between the observation of patterns ∂A and ∂B, as formally stated in Eq. 3.31. For statistical applications, the coherence given ∂ is of little use, since ∂ is a single observation among a potentially huge set $\mathbb{A}^m$. Therefore we define **the population coherence of branch** AB as

$$C(A;B) := \mathbb{E}[C^\partial(A;B)] = \sum_\partial \Pr(\partial) C^\partial(A;B). \tag{3.21}$$

The population coherence quantifies the dependence between clades A and B, because it tends to zero as the true length of branch AB grows (see Prop. 3.2.c).

We will estimate the population coherence using the observed alignment as follows. Given an alignment of n sites where pattern ∂ is observed n_∂ times, we define **the sample coherence of branch** AB as

$$\hat{C}(A;B) := \sum_\partial \frac{n_\partial}{n} C^\partial(A;B). \tag{3.22}$$

3.6.3 The memory of a clade

The $L_2(\pi)$-norm is the norm induced by the π-inner product. More explicitly, given a vector $\boldsymbol{v} = (v_i)$, we define its $L_2(\pi)$-norm as

$$\|\boldsymbol{v}\|_\pi = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle_\pi} = \sqrt{\sum_i \pi_i v_i^2}. \tag{3.23}$$

Consider any tree or subtree with root R and pattern ∂. We define **the memory** $M^\partial(R)$ **of clade** R **given pattern** ∂ as the squared $L_2(\pi)$-norm of the memory vector at R, namely

$$M^\partial(R) := \|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi^2. \tag{3.24}$$

Abusing of the notation, we can write $C^\partial(R;R) := M^\partial(R)$ in a strictly algebraic sense, ignoring the fact that branch RR does not exist in our tree. As we did with the coherence, we define **the population memory of clade** R as

$$M(R) := \mathbb{E}[M^\partial(R)] = \sum_\partial \Pr(\partial) M^\partial(R). \tag{3.25}$$

The population memory has a direct interpretation, since it is an average for all patterns of the norm of $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}$, which is small when the likelihood of the ancestral states at the root is uniform. Thus $M(R)$ quantifies the identification of the root, that is, how confidently we expect to reconstruct the root ancestral state. In par-

ticular, $M(R) = 0$ iff $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = \mathbf{1}$ for all patterns ∂, meaning that no pattern provides information about the ancestral root state.

To estimate $M(R)$ given an alignment where pattern ∂ is observed n_∂ times, we define **the sample memory of clade** R as

$$\hat{M}(R) = \sum_{\partial} \frac{n_\partial}{n} M^\partial(R). \tag{3.26}$$

3.7 Spectral Decomposition

In this section, the eigendecomposition of a reversible rate matrix is introduced, giving us the eigendecomposition of the objects defined in Section 3.6. Explicit examples of the eigendecomposition of the coherence and the memory are presented in Appendix 3.A.

3.7.1 Decomposition of the rate matrix

Consider a rate matrix $Q = (q_{ij})$ with $q_{ii} = -\sum_{j \neq i} q_{ij}$, that is, such that the sum of each **row** of Q is 0. Along this work, **we always assume that Q is irreducible**, meaning that we have a positive probability $p_{ij}(t) > 0$ of mutating from i to j for all times $t > 0$. In the following proposition, we state the well known eigenvector decomposition of a reversible rate matrix, similarly explained by Levin and Peres [2017] (Chapter 1, Section 12.1).

Proposition 3.1. *For an irreducible rate matrix $Q = (q_{ij})$ over an alphabet $\mathbb{A}$ of $K+1$ states, the following holds:*

a) Matrix Q has eigenvalue $\lambda_0 = 0$ with algebraic multiplicity 1. A right eigenvector of λ_0 is $\mathbf{1}$, whose left eigenvector is the unique equilibrium distribution $\boldsymbol{\pi}^T > 0$, defined by equation $\boldsymbol{\pi}^T Q = \boldsymbol{\pi}^T$. The rest of eigenvalues of Q are complex number with strictly negative real part.

Finally, the exponential matrix e^{Qt} satisfies

$$e^{Qt} \to \mathbf{1}\boldsymbol{\pi}^T \text{ as } t \to \infty.$$

b) If matrix Q is reversible, then it has real eigenvalues $0 > \lambda_1 \geq \cdots \geq \lambda_K$. Moreover, matrix Q has an orthogonal basis of right eigenvectors $\boldsymbol{v_k}$ and left eigenvectors $\boldsymbol{h_k^T}$ such that $\boldsymbol{h_k} = \boldsymbol{\pi} \circ \boldsymbol{v_k} = \Psi \boldsymbol{v_k}$ for $k \in [0, K]$, where $\boldsymbol{v_0} = \mathbf{1}$, $\boldsymbol{h_0} = \boldsymbol{\pi}$

and

$$Q = \mathbf{1}\boldsymbol{\pi}^T + \boldsymbol{v_1}\boldsymbol{h_1}^T\lambda_1 + \cdots + \boldsymbol{v_K}\boldsymbol{h_K}^T\lambda_K.$$

Finally, the exponential matrix e^{Qt} can be computed as

$$e^{Qt} = \mathbf{1}\boldsymbol{\pi}^T + \boldsymbol{v_1}\boldsymbol{h_1}^T e^{\lambda_1 t} + \cdots + \boldsymbol{v_K}\boldsymbol{h_K}^T e^{\lambda_K t}.$$

3.7.2 Decomposition of the likelihood of a branch

If Q is reversible, then we can decompose the likelihood of a branch length of Eq. 3.15, which assumes stationarity and the root placed at node A. Since $\Psi e^{Qt} = \boldsymbol{\pi}\boldsymbol{\pi}^T + \sum_{k\in[K]} \boldsymbol{h_k}\boldsymbol{h_k}^T e^{\lambda_k t}$, we have

$$\Pr(\partial \mid t) = \boldsymbol{\alpha_{\partial A}}^T \Psi e^{Qt} \boldsymbol{\beta_{\partial B}} = \Pr(\partial A)\Pr(\partial B) + \sum_{k\in[K]} (\boldsymbol{\alpha_{\partial A}}\cdot\boldsymbol{h_k})(\boldsymbol{\beta_{\partial B}}\cdot\boldsymbol{h_k})e^{\lambda_k t}. \quad (3.27)$$

Hoang et al. [2017] explain that, during likelihood computation assuming a reversible process, instead of storing the likelihood vectors $\boldsymbol{\alpha_\partial}$ and $\boldsymbol{\beta_\partial}$, IQ-TREE stores the quantities $\boldsymbol{h_k}\cdot\boldsymbol{\alpha_\partial}$ and $\boldsymbol{h_k}\cdot\boldsymbol{\beta_\partial}$ for $k \in [0, K]$, allowing a faster computation of the likelihood $\Pr(\partial \mid t)$.

If subpatterns ∂A and ∂B were independent events, we would have $\Pr(\partial \mid t) = \Pr(\partial A)\Pr(\partial B)$. Consequently, to measure the dependence between events ∂A and ∂B, we define **the dependence factor** $D(\partial \mid t)$ as

$$D(\partial \mid t) := \frac{\Pr(\partial \mid t)}{\Pr(\partial A)\Pr(\partial B)} = \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}^T \Psi e^{Qt} \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}, \quad (3.28)$$

which equals 1 if ∂A and ∂B are independent. The quantity $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}^T \Psi$ can be interpreted as the posterior state distribution given ∂A. In any case, using the eigendecompostion of Eq. 3.27, we can write

$$D(\partial \mid t) = 1 + \sum_{k\in[K]} (\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}\cdot\boldsymbol{h_k})(\tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}\cdot\boldsymbol{h_k})e^{\lambda_k t}, \quad (3.29)$$

showing an important property, namely the limit $D(\partial \mid t) \to 1$ as $t \to \infty$. In other words, the subpatterns ∂A and ∂B become independent as the assumed branch length t grows. This property also holds assuming only stationarity: If the tree is rooted at A, substitute $e^{Qt} \to \mathbf{1}\boldsymbol{\pi}^T$ (Prop. 3.1.a) in Eq. 3.28, giving

$$D(\partial \mid t) \to \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}^T \boldsymbol{\pi}\boldsymbol{\pi}^T \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} = 1 \times 1 = 1. \quad (3.30)$$

Note that $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}\cdot\boldsymbol{h_k} = \boldsymbol{a}_{\boldsymbol{\partial A}}\cdot\boldsymbol{v_k}$, where $\boldsymbol{a}_{\boldsymbol{\partial A}}$ is the posterior distribution given

∂A with prior $\boldsymbol{\pi}$. This alternative expression is sometimes easier to handle. As an example, if clade A is a single sequence and $\partial A = i$, then $\boldsymbol{a_{\partial A}}$ has a single nonzero entry 1 at position i, and $\boldsymbol{a_{\partial A}} \cdot \boldsymbol{v_k} = v_k^i$, where $\boldsymbol{v_k} = (v_k^0, \cdots, v_k^K)$.

3.7.3 Decomposition of the coherence

The dependence factor is nearly a reformulation of the coherence, because the coherence between nodes A and $e^{Qt}B$ is

$$C^{\partial}(A; e^{Qt}B) = \langle \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}, e^{Qt}\tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} \rangle_{\pi} - 1 = \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}^{T} \Psi e^{Qt} \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} - 1 = D(\partial \mid t) - 1. \quad (3.31)$$

This gives the intuition that the coherence of branch AB vanishes as subpatterns ∂A and ∂B are close to independent. Since $C^{\partial}(A; e^{Qt}B) = C^{\partial}(A; B)$ for $t = 0$, the reversible decomposition of the dependence factor of Eq. 3.29 with $t = 0$ gives

$$C^{\partial}(A; B) = \sum_{k \in [K]} (\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} \cdot \boldsymbol{h_k})(\tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} \cdot \boldsymbol{h_k}). \quad (3.32)$$

Motivated by this decomposition, we define **the k-projection of the coherence of branch AB given $\partial = (\partial A, \partial B)$** as

$$C_k^{\partial}(A; B) := (\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} \cdot \boldsymbol{h_k})(\tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} \cdot \boldsymbol{h_k}), \quad (3.33)$$

and Eq. 3.32 can be rewritten as

$$C^{\partial}(A; B) = \sum_{k \in [K]} C_k^{\partial}(A; B). \quad (3.34)$$

The eigendecomposition of Eq. 3.29 further implies that the action of matrix e^{Qt} gives

$$C^{\partial}(A; e^{Qt}B) = \sum_{k \in [K]} e^{\lambda_k t} C_k^{\partial}(A; B), \quad (3.35)$$

indicating that each k-projection decreases as an exponential decay of rate λ_k as the assumed length of branch AB grows. Analogously, we define **the k-projection of the population coherence of branch AB** as

$$C_k(A; B) := \mathbb{E}[C_k^{\partial}(A; B)], \quad (3.36)$$

and **the k-projection of the sample coherence of branch** AB as

$$\hat{C}_k(A;B) := \sum_{\partial} \frac{n_\partial}{n} C_k^\partial(A;B) = \sum_{\partial} \frac{n_\partial}{n} (\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} \cdot \boldsymbol{h_k})(\tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} \cdot \boldsymbol{h_k}). \tag{3.37}$$

3.7.4 Decomposition of the memory

Since for a clade R, we have $M^\partial(R) = C^\partial(R,R)$ algebraically, Eq. 3.32 with $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} = \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} = \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$ leads to the reversible decomposition

$$M^\partial(R) = \sum_{k\in[K]} (\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} \cdot \boldsymbol{h_k})^2. \tag{3.38}$$

We define **the kl-projection of the memory of clade R given ∂** as

$$M_{kl}^\partial(R) := (\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} \cdot \boldsymbol{h_k})(\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} \cdot \boldsymbol{h_l}), \tag{3.39}$$

and thus Eq. 3.32 is rewritten as

$$M^\partial(R) = \sum_{k\in[K]} M_{kk}^\partial(R). \tag{3.40}$$

When the transition matrix e^{Qt} acts on node R, Eq. 3.35 implies that

$$M^\partial(e^{Qt}R) = \sum_{k\in[K]} e^{2\lambda_k t} M_{kk}^\partial(R). \tag{3.41}$$

We further define **the kl-projection of the population memory of clade** R as

$$M_{kl}(R) := \mathbb{E}[M_{kl}^\partial(R)], \tag{3.42}$$

and **the kl-projection of the sample memory of clade** R as

$$\hat{M}_{kl}(R) = \sum_{\partial} \frac{n_\partial}{n} M_{kl}^\partial(R) = \sum_{\partial} \frac{n_\partial}{n} (\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} \cdot \boldsymbol{h_k})(\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} \cdot \boldsymbol{h_l}). \tag{3.43}$$

A very useful kl-projection of the population memory is obtained when the tree is composed by a single sequence $R = \mathbf{y}$. Then we have $M_{kl}(\mathbf{y}) = \delta_{kl}$ (Prop. 3.6.f), where $\delta_{kl} = 1$ if $k = l$ and $\delta_{kl} = 0$ if $k \neq l$.

3.8 The Relation between Coherence and Memory

In this section, we describe the expectation and the variance of the coherence of branch AB, with focus on the assumption that branch AB has infinite length. This assumption is the basis of the tests for saturation described in Section 3.10.

In the following proposition, we compute the expectation of the projections of the coherence of a branch. The non-reversible analogous result is stated in Prop. 3.10.

Proposition 3.2. *Consider a reversible and stationary evolutionary process on a tree. If branch AB has true length t^*, then the following holds.*

a) If $t^* \to \infty$, $C_k(A;B) := \mathbb{E}[C_k^{\partial}(A;B) \mid t^* \to \infty] = 0$ for all $k \in [K]$.

b) For any branch length $t^* \geq 0$, $C_k(A;B) = \sum_{l\in[K]} e^{\lambda_l t^*} M_{kl}(A) M_{kl}(B)$.

Proof.

a) By making $t^* \to \infty$ in Eq. 3.27, it follows that ∂A and ∂B are independent events. Thus the k-projection of the population coherence can be computed as

$$\mathbb{E}[(\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}})(\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}) \mid t^* \to \infty] = \mathbb{E}[\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}]\mathbb{E}[\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}]. \tag{3.44}$$

Using linearity, we compute e.g. $\mathbb{E}[\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}] = \sum_{\partial A} \boldsymbol{\alpha}_{\boldsymbol{\partial A}} = \mathbf{1}$, and $\boldsymbol{h_k} \cdot \mathbf{1} = 0$.

b) Substituting $\Pr(\partial)$ in the definition of $C_k(A;B)$ as implied by Eq. 3.27, we get

$$\begin{aligned}
&C_k(A;B) = C_k(A;B) - \mathbb{E}[C_k^{\partial}(A;B) \mid t^* \to \infty] = \\
&= \sum_{\partial} (\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}})(\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}) \Big(\Pr(\partial A) \Pr(\partial B) \sum_{l\in[K]} e^{\lambda_l t^*} (\boldsymbol{h_l} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}})(\boldsymbol{h_l} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}) \Big) = \\
&= \sum_{\partial} \Pr(\partial A) \Pr(\partial B) (\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}})(\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}) \sum_{l\in[K]} e^{\lambda_l t^*} (\boldsymbol{h_l} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}})(\boldsymbol{h_l} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}) = \\
&= \sum_{l\in[K]} e^{\lambda_l t^*} \sum_{\partial} \Pr(\partial A) \Pr(\partial B) (\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}})(\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}})(\boldsymbol{h_l} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}})(\boldsymbol{h_l} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}) = \\
&= \sum_{l\in[K]} e^{\lambda_l t^*} \sum_{\partial A, \partial B} \Pr(\partial A)(\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}})(\boldsymbol{h_l} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}) \Pr(\partial B)(\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}})(\boldsymbol{h_l} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}) = \\
&= \sum_{l\in[K]} e^{\lambda_l t^*} \sum_{\partial A} \Pr(\partial A)(\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}})(\boldsymbol{h_l} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}) \sum_{\partial B} \Pr(\partial B)(\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}})(\boldsymbol{h_l} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}) = \\
&= \sum_{l\in[K]} e^{\lambda_l t^*} M_{kl}(A) M_{kl}(B).
\end{aligned}$$

□

Notably, Prop. 3.2.b gives a first moment estimate (FME) $\hat{t}$ of the true length t^* of branch AB. Indeed, for any $k \in [K]$, estimate $\hat{t}$ is the solution(s) (if any) of equation

$$\hat{C}_k(A;B) = \sum_{l\in[K]} e^{\lambda_l \hat{t}} \hat{M}_{kl}(A)\hat{M}_{kl}(B), \tag{3.45}$$

which can be computed numerically. In Supp. Notebook 5 we see that, for $k \neq l$, $\mathbb{E}[M_{kl}(A)]$ is around two orders of magnitude smaller than $\mathbb{E}[M_{kk}(A)]$. Moreover, we have $e^{\lambda_1 t^*} >> e^{\lambda_l t^*}$ if $\lambda_1 > \lambda_l$. Thus for any k such that $\boldsymbol{h_k}$ has eigenvalue λ_1, a rough estimate of t^* can be obtained assuming $M_{kl}(A) \approx 0$ for $k \neq l$, giving

$$\hat{t} \approx \frac{1}{\lambda_k} \log\Big(\frac{\hat{C}_k(A;B)}{\hat{M}_{kk}(A)\hat{M}_{kk}(B)}\Big). \tag{3.46}$$

Given the likelihood vectors $\boldsymbol{\alpha_{\partial A}}$ and $\boldsymbol{\beta_{\partial B}}$, this FME approximation is $\Theta(K)$ times faster than one single iteration of the Newton method, computed as stated by Hoang et al. [2017]. We further analyze this approximation in Subsec. 3.9.1.

The FME approximation of Eq. 3.46 is unnecessary when branch AB is external. Indeed, if e.g. clade A is a single sequence $\mathbf{y}$, then $M_{kl}(\mathbf{y}) = \delta_{kl}$ (Prop. 3.6.f) and Prop. 3.2.b gives the FME

$$\hat{t} = \frac{1}{\lambda_k} \log\Big(\frac{\hat{C}_k(\mathbf{y};B)}{\hat{M}_{kk}(B)}\Big). \tag{3.47}$$

The following proposition describes the variance of the coherence assuming that the true length of branch AB is infinity. Its non-reversible counterpart is stated in Prop. 3.11.

Proposition 3.3. *Consider a reversible and stationary process on a tree over an alphabet with $K+1$ states. If branch AB has true length t^*, then the following holds.*

a) For any subset $S \subseteq [K]$ we have

$$\mathrm{Var}[\sum_{k\in S} C_k^{\partial}(A;B) \mid t^* \to \infty] = \sum_{k,l\in S} M_{kl}(A)M_{kl}(B) \leq K^2.$$

b) Assume that $\sum_{k\in S} C_k^{\partial}(A;B)$ is not constant in ∂, as also that Q has eigenvalues $0 > \lambda_1 \geq \cdots \geq \lambda_K$ such that λ_1 has multiplicity $D \leq K$. Then, for

$t^* < \infty$ *large enough,*

$$\sum_{k\in[D]} C_k(A;B) > 0.$$

Proof.

a) We know from Eq. 3.27 that $t^* \to \infty$ implies that ∂A and ∂B are independent events. We also proved that $\mathbb{E}[C_k^{\partial}(A;B) \mid t \to \infty] = 0$ in Prop. 3.2.a. Therefore

$$\mathrm{Var}[\sum_{k\in S} C_k^{\partial}(A;B) \mid t^* \to \infty] = \mathbb{E}[(\sum_{k\in S} C_k^{\partial}(A;B))^2 \mid t^* \to \infty]. \tag{3.48}$$

After expanding the squared sum $(\sum_{k\in S} C_k^{\partial}(A;B))^2$, note that

$$C_k^{\partial}(A;B)C_l^{\partial}(A;B) = M_{kl}^{\partial A}(A)M_{kl}^{\partial B}(B). \tag{3.49}$$

Thus we have

$$\begin{aligned} &\mathbb{E}[C_k^{\partial}(A;B)C_l^{\partial}(A;B) \mid t^* \to \infty] = \\ =&\mathbb{E}[M_{kl}^{\partial A}(A)M_{kl}^{\partial B}(B) \mid t^* \to \infty] = \\ =&\ \mathbb{E}[M_{kl}^{\partial A}(A)]\mathbb{E}[M_{kl}^{\partial B}(B)] = M_{kl}(A)M_{kl}(B). \end{aligned} \tag{3.50}$$

Summing for all $k, l \in [S]$ we get the desired equality. Using Prop. 3.8.c with $n_{\partial}/n = \Pr(\partial)$, we obtain the inequality

$$\sum_{k,l\in S} M_{kl}(A)M_{kl}(B) \leq (\sum_{k\in S} M_{kk}(A))(\sum_{k\in S} M_{kk}(B)). \tag{3.51}$$

In Prop.3.7.c, we proved that $M(A) \leq K$, and thus $\sum_{k\in S} M_{kk}(A) \leq M(A) \leq K$. Analogously we have $\sum_{k\in S} M_{kk}(B) \leq K$, yielding the bound.

b) Using Prop. 3.2.b, we know that

$$e^{-\lambda_1 t^*} \sum_{k\in[D]} C_k(A;B) = \sum_{k\in[D],l\in[K]} e^{(\lambda_l-\lambda_1)t^*} M_{kl}(A)M_{kl}(B) = \tag{3.52}$$

$$= \sum_{k,l\in[D]} M_{kl}(A)M_{kl}(B) + o(1), \tag{3.53}$$

where we used the fact that $e^{(\lambda_l-\lambda_1)t^*} \to 0$ as $t^* \to \infty$ for any $\lambda_l < \lambda_1$. Since the variance of a nonconstant rv is positive, using item a) with $S = [D]$, we know that $0 < \sum_{k,l\in[D]} M_{kl}(A)M_{kl}(B)$. Thus for t^* large enough, Eq. 3.53 is positive, giving the result.

□

3.9 Asymptotics of the Log-Likelihood of a Branch

The log-likelihood of a process E was introduced in Eq. 3.5. Considering the length of branch AB as the only variable of a stationary process rooted at A, the log-likelihood of branch length t equals

$$L(t) = \sum_{\partial} n_{\partial} \log \Pr(\partial \mid t) = \sum_{\partial} n_{\partial} \log \Big(\boldsymbol{\alpha_{\partial A}}^T \Psi e^{Qt} \boldsymbol{\beta_{\partial B}} \Big) = \tag{3.54}$$

$$= \sum_{\partial} n_{\partial} \log \Big(\langle \boldsymbol{\alpha_{\partial A}}, e^{Qt} \boldsymbol{\beta_{\partial B}} \rangle_{\pi} \Big), \tag{3.55}$$

where we substituted $\Pr(\partial \mid t)$ using Eq. 3.15 and the definition of the π-scalar product of Eq. 3.18. We define the log-likelihood of independence as

$$L_{\infty} := \sum_{\partial} n_{\partial} \Big(\log \Pr(\partial A) + \log \Pr(\partial B) \Big), \tag{3.56}$$

which together with $C^{\partial}(A; e^{Qt}B) = \langle \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}, e^{Qt} \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} \rangle_{\pi} - 1$ (Eq. 3.31) allows to rewrite Eq. 3.55 as

$$L(t) - L_{\infty} = \sum_{\partial} n_{\partial} \log \Big(1 + C^{\partial}(A; e^{Qt}B) \Big). \tag{3.57}$$

The well-known first order approximation $\log(1 + x) \approx x$ for small x implies that, for small $C^{\partial}(A; e^{Qt}B)$, we have

$$L(t) - L_{\infty} \approx n \hat{C}(A; e^{Qt}B), \tag{3.58}$$

and thus the sample coherence emerges as a useful tool to describe the log-likelihood for large t. Going further, if we assume reversibility, then the decomposition of the coherence of Eq. 3.35 applied to Eq. 3.57 gives

$$L(t) - L_{\infty} = \sum_{\partial} n_{\partial} \log \Big(1 + \sum_{k \in [K]} e^{\lambda_k t} C_k^{\partial}(A; B) \Big) = \tag{3.59}$$

$$= \sum_{\partial} n_{\partial} \log \Big(1 + \sum_{k \in [K]} e^{\lambda_k t} (\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}) (\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}) \Big). \tag{3.60}$$

In the following proposition, we describe $L(t)$ asymptotically.

Proposition 3.4. *Given an alignment where pattern ∂ is observed n_∂ times, assume that the alignment is the realization of a stationary and reversible process on a tree with rate matrix Q. If matrix Q has eigenvalues $0 > \lambda_1 \geq \cdots \geq \lambda_K$ such that λ_1 has multiplicity $D \leq K$, then the log-likelihood $L(t)$ of branch AB having length t satisfies the following.*

a) $L(t) \sim L_\infty$, that is, the quotient $L(t)/L_\infty \to 1$ as $t \to \infty$.

b) Define the dominant sample coherence as

$$\hat{\delta} := \sum_{k\in[D]} \hat{C}_k(A;B) = \sum_{k\in[D]} \sum_{\partial} \frac{n_\partial}{n} (\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}})(\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}).$$

If $\delta \neq 0$, then $L(t) - L_\infty \sim \hat{\delta} n e^{\lambda_1 t}$.

Proof.

Note that, when either $\boldsymbol{\alpha_\partial} = \mathbf{1}$ or $\boldsymbol{\beta_\partial} = \mathbf{1}$, the log-likelihood $\Pr(\partial \mid t)$ is constant on t. Consequently, to avoid degenerated cases, assume that, for at least one observed pattern, at least one entry of $\boldsymbol{\alpha_\partial}$ or $\boldsymbol{\beta_\partial}$ is strictly smaller than 1, implying $(\boldsymbol{\alpha}_{\boldsymbol{\partial A}} \cdot \boldsymbol{\pi})(\boldsymbol{\beta}_{\boldsymbol{\partial B}} \cdot \boldsymbol{\pi}) < 1$.

a) Using Eq. 3.56, $L(t) - L_\infty$ tends to zero as $t \to \infty$, because $\log(1) = 0$. To prove $L(t) \sim L_\infty$, it is enough to show that $L_\infty < 0$. By assumption, for some ∂, it holds that $\log(\boldsymbol{\alpha}_{\boldsymbol{\partial A}} \cdot \boldsymbol{\pi}) + \log(\boldsymbol{\beta}_{\boldsymbol{\partial B}} \cdot \boldsymbol{\pi}) < 0$, and thus L_∞ is strictly negative.

b) Using L'Hôpital's rule, it is enough to show that the derivative $L'(t)$ satisfies $L'(t) \sim \hat{\delta} n \lambda_1 e^{\lambda_1 t}$. In the expression

$$[\log\Big(1 + \sum_{k\in[K]} e^{\lambda_k t} C_k^\partial(A;B)\Big)]' = \frac{\sum_{k\in[K]} \lambda_k e^{\lambda_k t} C_k^\partial(A;B)}{1 + \sum_{k\in[K]} e^{\lambda_k t} C_k^\partial(A;B)},$$

the denominator is asymptotically equivalent to 1. On the other hand, if $\sum_{k\in[D]} C_k^\partial(A;B) \neq 0$, then the numerator is equivalent to $\lambda_1 e^{\lambda_1 t} \sum_{k\in[D]} C_k^\partial(A;B)$, and $o(e^{\lambda_1 t})$ otherwise. The equivalences can be summed up due to the assumption $\hat{\delta} \neq 0$.

□

Figure 3.10 shows the limit of Prop 3.4.b for two sequences. Notably, Prop. 3.4.b implies that whether $L(t)$ increases or decreases for large t uniquely depends on the sign of the dominant sample coherence $\hat{\delta}$, as stated in the following corollary.

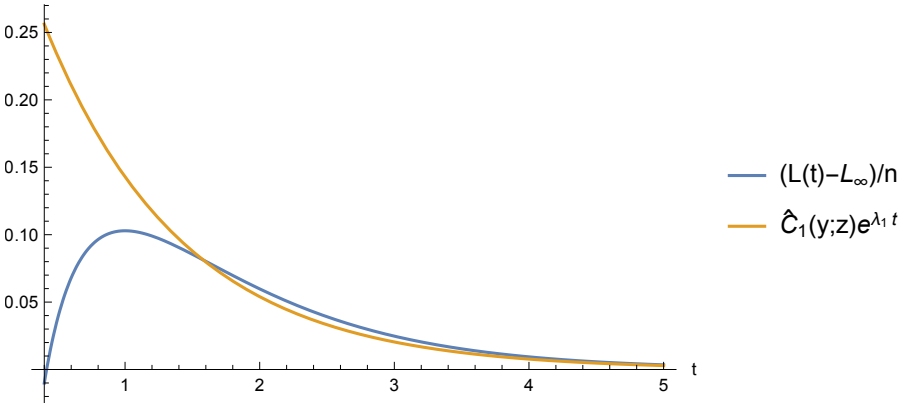

Figure 3.10: Plot of the difference $(L(t)-L_\infty)/n$ and the dominant exponential decay $\hat{C}_1(\mathbf{y},\mathbf{z})e^{\lambda_1 t}$ for the expected alignment of two sequences generated by a random reversible rate matrix Q where λ_1 has multiplicity $D=1$. See Supp. Notebook 1.

Corollary 3.5. *Given an alignment and assuming a stationary and reversible process on a tree, the log-likelihood $L(t)$ of branch AB having length t satisfies the following.*

a) If $\hat{\delta} > 0$, then $L(t)$ has local minimum L_∞ when $t \to \infty$.

b) If $\hat{\delta} < 0$, then $L(t)$ has local maximum L_∞ when $t \to \infty$.

If the log-likelihood $L(t)$ has a unique maximum for $t \in [0,\infty]$, Corollary 3.5 characterizes the convergence of the Newton method in the search for the absolute maximum of $L(t)$. It is remarkable, however, that the log-likelihood can have multiple maxima (even if infrequently), as explained in Appendix 3.E. For practitioners, we recommend using Corollary 3.5 to predict the possible numerical divergence of the Newton method.

3.9.1 The MLE and its asymptotic approximation

Interestingly, Corollary 3.5 implies that the MLE and the FME approximation of Eq. 3.46 behave similarly in extreme cases. Assume that the multiplicity of eigenvalue λ_1 is one. When $C_1(A;B) < 0$, Corollary 3.5 implies that the MLE may output $\hat{t} \to \infty$. On the other hand, when $C_1(A;B) \leq 0$, in Eq. 3.46 it makes sense to define $\log(C_1(A;B)) := -\infty$, giving $\hat{t} := \infty$.

At the opposite extreme, when $\tilde{\boldsymbol{\alpha}}_{\partial A} = \tilde{\boldsymbol{\beta}}_{\partial B}$ for all ∂, then

$$\hat{C}_1(A;B) = \hat{M}_{11}(A) = \hat{M}_{11}(B), \tag{3.61}$$

implying that $\hat{t} \leq 0$ in Eq. 3.46. Since only nonnegative lengths are allowed, we define $\hat{t} := 0$. Regarding the MLE, if we substitute $\tilde{\boldsymbol{\alpha}}_{\partial A} = \tilde{\boldsymbol{\beta}}_{\partial B}$ in the log-likelihood

of Eq. 3.60, we obtain

$$L(t) - L_\infty = \sum_{\partial} \log\left(1 + \sum_{k\in[3]} e^{\lambda_k t}(\boldsymbol{h_k}\cdot\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}})^2\right), \tag{3.62}$$

which decreases in t. Due to restriction $\hat{t} \geq 0$, we conclude that the MLE is $\hat{t} = 0$. Thus again the FME approximation of Eq. 3.46 and the MLE coincide.

It is remarkable for practitioners that Eq. 3.62 reaches its absolute maximum in $[0, \infty]$ at $t = 0$, no matter how mixed (that is, close to $\mathbf{1}$) the normalized likelihood vector $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}$ is. This explains why short internal branches may be estimated when reconstructing saturated phylogenies, because saturation may imply $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} \approx \mathbf{1} \approx \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}$ for nearly all patterns ∂, leading to the MLE $\hat{t} \approx 0$. It follows that a short MLE reconstructed internal branch can also be saturated according to the asymptotic test described in Section 3.10.

3.10 A Test for Branch Saturation

In this section, we will show that the sample coherence of a branch provides an adequate tool to test for saturation. Notably, a raw MLE cannot detect saturation (Appendix 3.D), implying that a test for saturation must be included in any reconstruction protocol.

We define saturation formally in Subsection 3.10.1. Then, in Subsection 3.10.2, we state the asymptotic test in its most general form, which is the most powerful α-level test for long branches, as shown in Subsection 3.10.3. Simple versions of the asymptotic test are stated in Subsection 3.10.4. The consequences of branch saturation are described in Subsection 3.10.5.

3.10.1 Statistical definition of saturation

Given an alignment, we define saturation as **the lack of *significance* to reject the null hypothesis that the alignment was generated from an infinite evolutionary process**.

More concretely, given an alignment, we assume as usual that it is the realization of a stationary and reversible process on a tree. We consider a branch AB on this tree with true length t^*. We say that branch AB is saturated if we cannot reject the null hypothesis that the true length t^* of AB is infinite, that is, the null hypothesis $t^* \to \infty$.

We know from Eq. 3.27 that $\Pr(\partial \mid t) \to \Pr(\partial A)\Pr(\partial B)$ as $t \to \infty$. Consequently, under the null hypothesis $t^* \to \infty$, subpatterns ∂A and ∂B are independent. This means that each subpattern is the realization of an independent process on clades A and B, respectively.

3.10.2 The asymptotic test

In Section 3.8, we computed the expectations and variances of any sum $\sum_{k\in S} C_k^{\partial}(A;B)$ under hypothesis $t^* \to \infty$, where S is a subset of $[K]$. With these results, many tests can be constructed using linear combinations of the projected coherences $C_k^{\partial}(A;B)$. However, we are specially interested in constructing a powerful test able to distinguish a large t^* from $t^* \to \infty$.

To that end, assume that the rate matrix Q has eigenvalues $0 \geq \lambda_1 \geq \cdots \geq \lambda_K$, where λ_1 has multiplicity $D \leq K$, and consider the estimator

$$\hat{\delta} := \sum_{k\in[D]} \hat{C}_k(A;B) \tag{3.63}$$

as defined in Prop. 3.4.b. We know that $\mathbb{E}[\hat{\delta} \mid t^* \to \infty] = 0$ (Prop. 3.2.a) and $\mathbb{E}[\hat{\delta}] > 0$ for large $t^* < \infty$ (Prop. 3.3.b). Therefore, given a level of significance α, we define **the asymptotic test** as:

$$\text{"Reject } t^* \to \infty \text{ if } \hat{\delta} > c_S\text{"}, \tag{3.64}$$

where the saturation coherence $c_S \in [0,1]$ is chosen so that

$$\Pr(\hat{\delta} > c_S \mid t^* \to \infty) = \alpha. \tag{3.65}$$

When $\hat{\delta} < c_S$, we say that branch AB is **saturated** (with significance α), because we do not reject $t^* \to \infty$. Otherwise we say that branch AB is **informative** (with significance α).

We can obtain an explicit expression for c_S as follows. Recall that, if $\mathrm{Var}[X] = \sigma^2$ and $X_s \sim X$ are i.i.d., then the variance of $1/n \sum_{s\in[n]} X_s$ is σ^2/n. This fact, combined with Prop. 3.3.a, gives

$$\mathrm{Var}[\hat{\delta} \mid t^* \to \infty] = \frac{1}{n} \sum_{k,l\in[D]} M_{kl}(A) M_{kl}(B), \tag{3.66}$$

while $\mathbb{E}[\hat{\delta} \mid t^* \to \infty] = 0$ using Prop. 3.2.a. Thus assuming $t^* \to \infty$, we have the

normal approximation

$$\hat{\delta} \sim N\Big(0, \frac{1}{n} \sum_{k,l \in [D]} M_{kl}(A) M_{kl}(B)\Big), \tag{3.67}$$

where $N(\mu, \sigma^2)$ is the normal distribution. We infer the approximation of c_S for large n,

$$c_S \approx z_\alpha \sqrt{\frac{1}{n} \sum_{k,l \in [D]} M_{kl}(A) M_{kl}(B)}, \tag{3.68}$$

where we define z_α by $\Pr(Z > z_\alpha) = \alpha$ for $Z \sim N(0,1)$. Computing $M_{kl}(A)M_{kl}(B)$ is often unfeasible, since the set of possible patterns $\mathbb{A}^m$ grows exponentially with m. Although the sample statistics $\hat{M}_{kl}(A)\hat{M}_{kl}(B)$ can be used as proxies, more formal alternatives are described in Appendix 3.F.

3.10.3 Comparison to the likelihood-ratio test

Let us show that the asymptotic test has optimal power for large t^*. We start by considering the likelihood-ratio test, which is the most powerful α-level test, as proved by Neyman and Pearson [1933]. If the MLE is $\hat{t} < \infty$, we compare the alternative hypothesis $t^* = \hat{t}$ versus the null hypothesis $t^* \to \infty$ using statistic $L(\hat{t}) - L_\infty$, or more explicitly

$$\text{"Reject } t^* \to \infty \text{ if } L(\hat{t}) - L_\infty > c\text{"}, \tag{3.69}$$

where $c \in \mathbb{R}_{>0}$ is chosen so that $\Pr(L(\hat{t}) - L_\infty > c \mid t^* \to \infty) = \alpha$.

Assuming that $\hat{t}$ is far enough from the singularity of $L(t)$, we can use the approximation $L(\hat{t}) \approx L_\infty + \hat{\delta} n e^{\lambda_1 \hat{t}}$ of Prop. 3.4.b, where $\hat{\delta} := \sum_{k \in [D]} \hat{C}_k(A;B)$. This gives the test

$$\text{"Reject } t^* \to \infty \text{ if } \hat{\delta} > e^{-\lambda_1 \hat{t}} c/n\text{"}, \tag{3.70}$$

where $c \in \mathbb{R}_{>0}$ is chosen so that $\Pr(\hat{\delta} > e^{-\lambda_1 \hat{t}} c/n \mid t^* \to \infty) = \alpha$. Setting $c_S := e^{-\lambda_1 \hat{t}} c/n$, we get the asymptotic test.

Notably, the asymptotic test does not depend on the MLE $\hat{t}$, unlike the likelihood-ratio test. This is an important advantage, since the numerical search for the absolute maximum $\hat{t}$ of $L(t)$ may be impeded by the presence of local maxima, as exemplified in Appendix 3.E.

3.10.4 Simple asymptotic tests

The asymptotic test becomes much simpler if $D = 1$. Then $\hat{\delta} = \hat{C}_1(A; B)$ and Eq. 3.66 becomes

$$\mathrm{Var}[\hat{C}_1(A; B) \mid t^* \to \infty] = \frac{1}{n} M_{11}(A) M_{11}(B), \tag{3.71}$$

implying that the saturation coherence of Eq. 3.68 is

$$c_S \approx z_\alpha \sqrt{\frac{M_{11}(A) M_{11}(B)}{n}}, \tag{3.72}$$

and thus we reject hypothesis $t^* \to \infty$ if $\hat{C}_1(A; B) > c_S$, or the upper bound of some confidence interval around c_S.

The asymptotic test is also simplified when applied to an external branch, meaning that clade A or clade B is a single sequence. Assume for example that $B = \mathbf{z}$, and recall that $M_{kl}(\mathbf{z}) = \delta_{kl}$ (Prop. 3.6.f). Then, Eq. 3.66 gives

$$\mathrm{Var}[\hat{\delta} \mid t^* \to \infty] = \frac{1}{n} \sum_{k \in [D]} M_{kk}(A), \tag{3.73}$$

meaning that, for $B = \mathbf{z}$, the variance of the sum of projected memories is the sum of variances. In this case, we reject hypothesis $t^* \to \infty$ if $\hat{\delta} > c_S$, where

$$c_S \approx z_\alpha \sqrt{\frac{1}{n} \sum_{k \in [D]} M_{kk}(A)}. \tag{3.74}$$

If also clade A is a single sequence, say $A = \mathbf{y}$, then the evolutionary process is E_{seq} (Figure 3.6) and the saturation coherence is

$$c_S \approx z_\alpha \sqrt{\frac{D}{n}}. \tag{3.75}$$

Hypothesis $t^* \to \infty$ is rejected if $\hat{\delta} > c_S$, or equivalently if $\sum_{k \in [D]} \boldsymbol{v_k}^T N \boldsymbol{v_k}/n > c_S$ (Prop. 3.6.a).

The simplest of all tests is obtained by considering process E_{seq} and $D = 1$, where $\hat{C}_1(\mathbf{y}; \mathbf{z}) = \boldsymbol{v_1}^T N \boldsymbol{v_1}/n$. If we set for example $n = 10000$ and significance level $\alpha = 0.01$, then since $\Pr(Z > 2.3) \approx 0.01$, the saturation coherence is $c_S \approx 2.3/100$. Thus we reject $t^* \to \infty$ if $\boldsymbol{v_1}^T N \boldsymbol{v_1}/n > 0.023$. Actually, in this very simple case, the asymptotic test can be intuitively expressed in terms of evolutionary time. Consider **the saturation time** $t_S := 1/\lambda_1 \log(c_S)$ and the first moment estimate $\hat{t} := 1/\lambda_1 \log(\boldsymbol{v_1}^T N \boldsymbol{v_1}/n)$ (Eq. 3.47). Then branch **yz** is saturated with signifi-

cance α iff $\hat{t} > t_S$. In Figure 3.11, for various significances α, we represent the scaled saturation time $|\lambda_1|t_S$ as a function of the number of sites n.

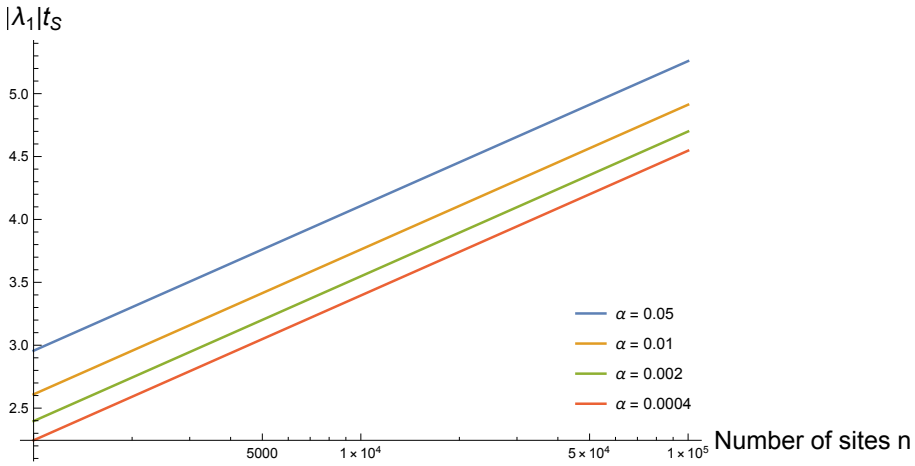

Figure 3.11: Log-linear plots of the scaled saturation time $|\lambda_1|t_S$ as a function of the number of sites n for $\alpha \in 0.25 \times \{5^{-1}, 5^{-2}, 5^{-3}, 5^{-4}\}$. Given α, the saturation time grows as a linear function of the logarithm of the sequence length. See Supp. Notebook 2.

3.10.5 The meaning of a saturated branch

In general, when a reconstructed process is given, we can apply the asymptotic test looking for saturated branches. Rejecting saturation using the asymptotic test is a necessary condition for a reliable reconstructed branch AB. Indeed, if a finite branch AB is saturated, then a realization of the reconstructed process may lead to an infinite estimate for this branch. This lack of reproducibility makes branch AB unreliable, and therefore we remain in the null hypothesis that the true length t^* of AB is infinite.

Recall moreover that $\Pr(\partial \mid t) \to \Pr(\partial A)\Pr(\partial B)$ as $t \to \infty$ (Eq. 3.27), and thus hypothesis $t^* \to \infty$ implies that observations ∂A and ∂B are independent. On the other side, since we assume a reversible process, the root of any phylogeny is unidentifiable. Therefore, if branch AB is saturated, then the reconstructed tree is composed by unrooted clade A (removing parent node A) independent from unrooted clade B (removing parent node B).

If unrooted clades A and B are binary trees and have respectively k and $m-k$ leaves with $m-k \geq k \geq 2$, then they have respectively $2k-3$ and $2(m-k)-3$ branches. There are $(2k-3)(2(m-k)-3)$ possible topologies obtainable by placing new parent nodes A' and B' at a branch of unrooted clades A and B, and then joining $A'B'$. Any true topology obtained by determining branch $A'B'$ agrees with

our reconstructed topology as long as we assume that branch $A'B'$ has infinite length $t_{A'B'} \to \infty$. In the particular case when A has $k = 1$ node, then the reconstructed topology agrees with $2m-5$ possible true topologies. This case is extensively studied in Section 3.11 for $m = 4$ sequences.

Summarizing, in a practical case, a saturated branch AB can be explained in three ways:

1 Too many mutations have happened due to t^* being too large, or $t_{A'B'}$ being too large for some ignored branch $A'B'$ of the true tree. An alignment with more sites may reject saturation, as represented in Figure 3.11.

2 Some sites are wrongly aligned. Under our assumptions, every insertion of sites starts a new evolutionary history. In this case, the consequence of a saturated branch AB is that, indeed, clades A and B have independent evolutionary histories.

3 The assumed evolutionary process is misspecified. Either the tree topology, the branch lengths or the rate matrix are wrong.

3.11 The Asymptotic Test Detects Long Branch Repulsion

In this section we show that the asymptotic test is able to recognize and resolve the phenomenon called **long branch repulsion (LBR)** in an alignment with 4 sequences. In particular, we apply the asymptotic test of Subsection 3.10.2 with significance $\alpha = 0.01$ assuming that the multiplicity of eigenvalue λ_1 is one, that is, $D = 1$. Using Eq. 3.72, the saturation coherence of any branch is

$$c_S \approx z_\alpha \sqrt{\frac{M_{11}(A)M_{11}(B)}{n}}. \tag{3.76}$$

For an external branch, say where $A = \mathbf{y}$, we have $M_{11}(A) = 1$ (Prop. 3.6.f) and saturation coherence

$$c_S \approx z_\alpha \sqrt{\frac{M_{11}(B)}{n}}. \tag{3.77}$$

To upper estimate the projected memories $M_{11}(B)$, we use estimator $\hat{M}_{11}(B) + 2\hat{s}$, where $\hat{s}$ is an upper bound of the standard deviation of any sample projected memory, giving confidence intervals of at least 95% (see Appendix 3.F). Using Prop. 3.7.f, we set $\hat{s} := \sqrt{U \min(K, U/4)/n}$.

3.11.1 Description of LBR

LBR, described by Siddall [1998], is an artifact that can occur when an evolutionary process is reconstructed using an MLE. It consists on the wrong reconstruction of two sister sequences due to their long external branches, as exemplified in Figure 3.12. Parameters p and q determine the true length of two and three branches, respectively.

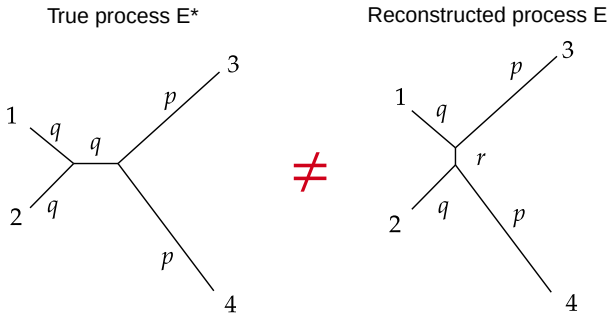

Figure 3.12: Graphic representation of LBR. In the true process E^*, the interior branch induces split 12|34, while in the reconstructed tree the interior branch induces split 13|24. By symmetry, LBR also occurs when the split 14|23 is induced by the interior branch. When reconstructing process E, we only estimate the interior branch r using an MLE (differing from Siddall [1998], who estimates more branches simultaneously). For the wrong 13- and 14-topologies, if p is large and q is small, then r is frequently close to 0.

In Eq. 3.103 we prove that deciding whether $t^* \to \infty$ on a branch AB is infeasible using just a finite MLE. Recall moreover that $t^* \to \infty$ implies that clades A and B are independent. Considering these results, the existence of LBR is not surprising: If p is very large, then observations 3 and 4 are in practice independent of subtree 12. When reconstructing process E, sequences 3 and 4 are placed nearly at random in the tree. Thus, under our reconstruction protocol, mainly the constrains over the distance between sequences 1 and 2 determine the estimated topology.

3.11.2 How to apply the asymptotic test

In Figure 3.13, we have a typical case of the asymptotic test applied **simultaneously** to all branches of the well reconstructed 12-tree and the wrongly reconstructed 13-tree. For both topologies, branches 3 and 4 are saturated, and therefore the asymptotic test applied to the reconstructed $12-$ and $13-$topologies implies the following: Sequences 1 and 2 are distant around $2q$ time units, while sequences 3 and 4 were sampled independently.

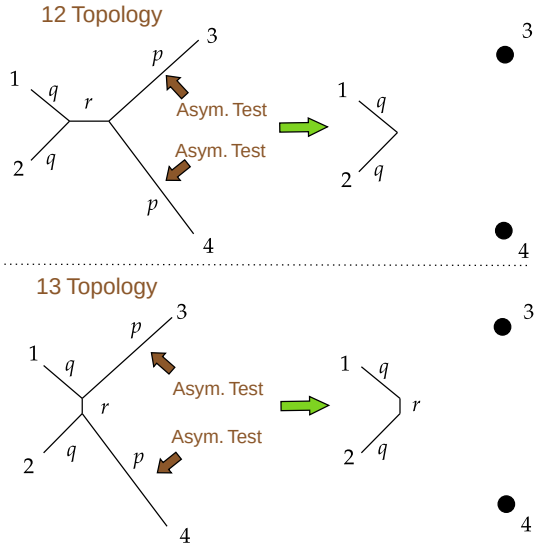

Figure 3.13: Graphic representation of the asymptotic test applied simultaneously to the long branches of the 12− and 13-topology. We do not apply the test sequentially to preserve the symmetry between sequences 3 and 4.

Intuitively, in Figure 3.13, parameter r of the 13-topology is close to 0 because the resulting subtree 12 is a reliable estimate the true subtree 12. As explained in Subsection 3.10.5, the tested 13-topology of Figure 3.13 does not contradict the true 12-topology, because the tested 13-topology is actually agnostic regarding the placement of sequences 3 and 4 in the tree. Note that this reasoning also applies when only one of branches 3 and 4 is saturated, since both subtrees 123 and 124 are subtrees of the true 12-topology.

The same way as branches are often optimized sequentially, also the asymptotic test can be applied sequentially until all remaining branches reject saturation. An example of the asymptotic test applied sequentially to the 12-topology is represented in Figure 3.14. Again, the final tree states that sequences 1 and 2 are around $2q$ time units apart, while sequences 3 and 4 were sampled independently from the rest of the tree.

3.11.3 Simulations

In Supp. Notebook 6, using a fixed random reversible rate matrix Q and for four pairs of parameters (p, q), we simulated 4×100 realizations of length $n = 5000$ of the true process E^* with a 12-topology.

Then we computed an MLE of the interior branch length and the true tree. After this, we simultaneously applied the asymptotic test to branches 3 and 4 to obtain their saturation status, abbreviated using the initials "i" for informative and "s" for

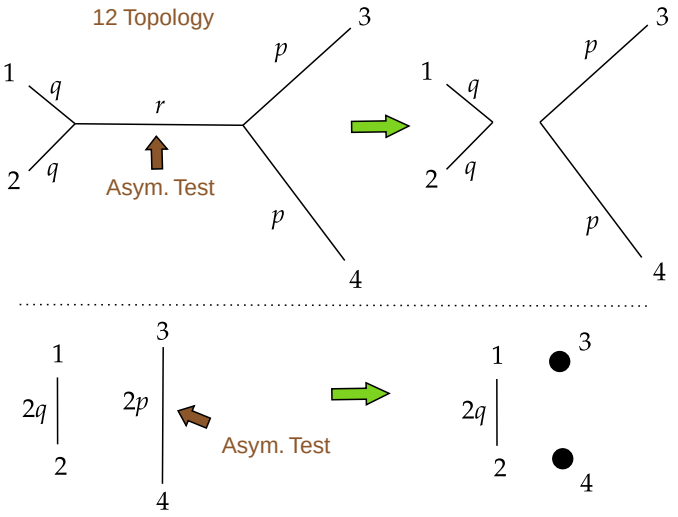

Figure 3.14: Graphic representation of the asymptotic test applied sequentially to the long branches of the 12-topology until all branches reject saturation.

saturated. For example, abbreviation is denotes that branch 3 is informative, while branch 4 is saturated. A summary of these simulations is presented in Tables 3.2 and 3.3.

Table 3.2 shows that the saturation status of branches 3 and 4 is the same one for all candidate topologies assuming that the true parameters are far enough from the saturation threshold (in our case, somewhere around $p = 2$). Thus for most of true parameters, the asymptotic test seems to remain invariant under a mild topological misspecification as a leaf rearrangement. Moreover, as expected, branches 3 and 4 are more often saturated as p increases.

In Table 3.3 we see that the frequency of wrongly estimated topologies increases with p, in agreement with Siddall [1998]. However, when the wrong 13- and 14-topologies are the estimate and $p = 3$, nearly always (32/33) at least one of branches 3 and 4 is saturated. Since the subtrees 123 and 134 are subtrees of the true 12-topology, those tested wrong estimates actually agree with the true 12-topology. For $p = 1.5$, LBR is absent, and accordingly branches 3 and 4 are always informative. When $p = 2$, closer to the saturation threshold, the asymptotic test has its poorest relative performance, providing a wrong tested topology with status ii in half (4/8) of the wrongly reconstructed topologies.

In absolute terms, however, Table 3.3 shows the usefulness of the asymptotic test: For any choice of parameters, in at least 96% of repeats we either estimated the correct 12-topology or found saturated branches in the 13- or 14-topology (thus agreeing with the 12-topology).

(p, q)	All ii	All si/ All is	All ss	Conflict
(1.5, 0.3)	100	0	0	0
(2, 0.3)	47	25	9	19
(2.5, 0.3)	6	20	68	6
(3, 0.3)	1	6	93	0

Table 3.2: Summary of our simulations using a fixed rate matrix Q and the true 12-topology to generate alignments of length $n = 5000$. For each (p, q) and each of the 100 repeats, we count in how many instances **All** candidate topologies have the same saturation status for branches 3 and 4, using significance $\alpha = 0.01$. If not all candidate topologies have the same saturation status, we count it as a **Conflict**. See Supp. Notebook 6.

(p, q)	Estimated Topology	Times	ii	is/si	ss
(1.5, 0.3)	12	100	100	0	0
	13 or 14	0	0	0	0
(2, 0.3)	12	92	43	28	21
	13 or 14	8	4	4	0
(2.5, 0.3)	12	79	2	16	61
	13 or 14	21	4	4	13
(3, 0.3)	12	67	0	0	67
	13 or 14	33	1	6	26

Table 3.3: Summary of our simulations using a fixed rate matrix Q and the true 12-topology to generate alignments of length $n = 5000$. For each (p, q) and each of the 100 repeats, we find the estimated topology by optimizing the interior branch length of each topology. Then we compute the saturation status of branches 3 and 4 **of the MLE** using significance $\alpha = 0.01$. In red, the wrong estimates after testing. See Supp. Notebook 6.

Our simulations show that, with this setup, LBR is mostly a consequence of saturation and can be detected using the asymptotic test. In particular, the existence of LBR does not imply that the MLE is "prone to failure" in the Farris zone. In general, we cannot expect that any reconstruction method will always recover the true topology no matter how long the true branches are. The virtue of ML, however, is that using its estimated parameters we can test the quality of our reconstruction, as we have exemplified in this section.

3.12 Conclusions and future research

Attending to our theoretical results, the coherence of a branch and the memory of a clade provide well-behaved statistics describing the underlying tree structure of an evolutionary process. Moreover, the asymptotic test is simple and powerful enough to detect long branch repulsion in our simulations.

In order to systematically apply the asymptotic test to phylogenies obtained from real data, future work should focus on its software implementation. Then, highly mutated alignments such as the SIV sequences available at Los Alamos database [Foley et al., 2020] could provide good examples of saturation, as we can see in Subsection 3.3.2.

Once systematic data is available, the sample coherence could be compared to the branch support values of the bootstrap [Felsenstein, 1985, Efron, 1992]. Intuitively, a small sample coherence leads to branch saturation, which implies low branch support values, as exemplified by the wrong reconstructed topologies of Section 3.11. However, we ignore up to which extent a big sample coherence implies a high bootstrap support, even if saturation is rejected. Since the bootstrap is computationally expensive, the sample coherence of a branch could offer an efficient alternative with a solid theoretical basis.

Regarding generalizations, it is remarkable that the memory and the coherence are well-defined also under nonreversible and nonstationary models of evolution, as we have exemplified in Appendix 3.C. With this more general setup, Manuel [2022] used the memory vector to upper-bound information flow on trees. However, the relationship between these generalized measures and MLE reconstruction is more complicated.

As an example, it seems feasible to construct a non-reversibility asymptotic test for a branch AB rooted at A (see Appendix 3.C). However, testing other branches not adjacent to the root requires recomputing the probability of observing each

pattern under the null hypothesis. In contrast, other useful generalizations, such as the inclusion of varying evolutionary rates in the alignment, are straightforward, as exemplified in Subsection 3.3.2.

Appendix 3.A Explicit Computation with Two Sequences

To provide some examples, the process E_{seq} represented in Figure 3.6 has a simple tree structure that allows to explicitly compute the coherence and the memory, as we do in Prop. 3.6. We assume that clade A is composed uniquely by sequence $\mathbf{y}$, and thus write $A = \mathbf{y}$, and similarly we set $B = \mathbf{z}$. Given the alignment $(\mathbf{y}, \mathbf{z})$, if pattern $ij \in \mathbb{A}^2$ is observed n_{ij} times, then we define **the diversity matrix** as $N := (n_{ij})$. We define moreover the vector $\boldsymbol{n_y} = (n_i)$, where $n_i := \sum_j n_{ij}$, that is, n_i counts how many times state i was observed in sequence $\mathbf{y}$. Recall also that the Kronecker delta δ_{kl} equals 1 if $k = l$ and 0 otherwise.

Proposition 3.6. *Consider process E_{seq} with true transition matrix $e^{Qt^*} = (p_{ij}(t^*))$, where the reversible rate matrix Q over $K+1$ states with equilibrium frequency $\boldsymbol{\pi} > 0$ is assumed to be known. Then the following holds.*

a) $\hat{C}_k(\boldsymbol{y}, \boldsymbol{z}) = \boldsymbol{v_k}^T N \boldsymbol{v_k}/n$.

b) $C_k(\boldsymbol{y}, \boldsymbol{z}) = e^{\lambda_k t^*}$.

c) $\hat{C}(\boldsymbol{y}, \boldsymbol{z}) = -1 + \sum_i \frac{n_{ii}}{n} \frac{1}{\pi_i}$.

d) $C(\boldsymbol{y}, \boldsymbol{z}) = -1 + \sum_i p_{ii}(t^*) = \sum_{k \in [K]} e^{\lambda_k t^*}$.

e) $\hat{M}_{kl}(\boldsymbol{y}) = (\boldsymbol{v_k} \circ \boldsymbol{v_l}) \cdot \boldsymbol{n_y}/n$.

f) $M_{kl}(\boldsymbol{y}) = \delta_{kl}$.

g) $\hat{M}(\boldsymbol{y}) = -1 + \sum_i \frac{n_i}{n} \frac{1}{\pi_i}$.

h) $M(\boldsymbol{y}) = K$.

Proof.

a) We write $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} = (\tilde{\alpha}_i)$ and $\tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} = (\tilde{\beta}_i)$. If $\partial = (\partial A, \partial B) = ij$, then $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} = \boldsymbol{e_i}/\pi_i$ and $\tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} = \boldsymbol{e_j}/\pi_j$, where $\boldsymbol{e_i}$ is the vector with 1 at the ith entry and 0 everywhere else. Recall that $\boldsymbol{h_k} = \boldsymbol{v_k} \circ \boldsymbol{\pi}$, and thus if we write $\boldsymbol{v_k} = (v_k^i)$, then

$C_k^{\partial}(\mathbf{y};\mathbf{z}) = v_k^i v_k^j$. Therefore

$$\hat{C}_k(\mathbf{y};\mathbf{z}) := \sum_{ij} \frac{n_{ij}}{n} C_k^{\partial}(\mathbf{y};\mathbf{z}) = \sum_{ij} \frac{n_{ij}}{n} v_k^i v_k^j = \boldsymbol{v_k}^T N \boldsymbol{v_k}/n.$$

b) Since $\Pr(ij) = \pi_i p_{ij}(t^*)$, it holds that $\mathbb{E}[N/n] = \Psi e^{Qt^*}$, where $\Psi = \mathrm{Diag}(\boldsymbol{\pi})$. Taking the expected value in item a), we obtain

$$C_k(\mathbf{y};\mathbf{z}) = \boldsymbol{v_k}^T \Psi e^{Qt^*} \boldsymbol{v_k} = \boldsymbol{h_k}^T e^{Qt^*} \boldsymbol{v_k} = e^{\lambda_k t^*}. \tag{3.78}$$

c) It holds that $C^{ij}(\mathbf{y};\mathbf{z}) = \delta_{ij}/\pi_i - 1$. Therefore we have

$$\hat{C}(\mathbf{y};\mathbf{z}) = \sum_i \frac{n_{ii}}{n}(\frac{1}{\pi_i} - 1) - \sum_{i \neq j} \frac{n_{ij}}{n} = \sum_i \frac{n_{ii}}{n}\frac{1}{\pi_i} - 1, \tag{3.79}$$

where we used the fact that $\sum_{ij} n_{ij} = n$.

d) To obtain the first equality, substitute $\mathbb{E}[n_{ii}/n] = \pi_i p_{ii}(t^*)$ in item c). To obtain the second equality, use item b) and the decomposition $C(\mathbf{y};\mathbf{z}) = \sum_{k\in[K]} C_k(\mathbf{y};\mathbf{z})$. Alternatively, use the fact that the trace of a matrix equals the sum of its eigenvalues.

e) For any pattern $\partial = i$, $M_{kl}^i(\mathbf{y}) = v_k^i v_l^i$, and thus

$$\hat{M}_{kl}(\mathbf{y}) = \sum_i \frac{n_i}{n} v_k^i v_l^i = (\boldsymbol{v_k} \circ \boldsymbol{v_l}) \cdot \boldsymbol{n_y}/n.$$

f) It holds that $\mathbb{E}[\boldsymbol{n_y}/n] = \boldsymbol{\pi}$. Taking the expected value in item e), we have

$$M_{kl}(\mathbf{y}) = (\boldsymbol{v_k} \circ \boldsymbol{v_l}) \cdot \boldsymbol{\pi} = \boldsymbol{h_k} \cdot \boldsymbol{v_l} = \delta_{kl}. \tag{3.80}$$

g) Just assume that $\mathbf{z} = \mathbf{y}$ in the alignment of item c), implying $n_{ii} = n_i$.

h) Using item f), $M(\mathbf{y}) = \sum_{k\in[K]} M_{kk}(\mathbf{y}) = \sum_{k\in[K]} 1 = K$.

□

Appendix 3.B Bounds of the Coherence and the Memory

In this section we state basic bounds involving the memory and the coherence. We start with some bounds of the memory of a clade and its variance.

Proposition 3.7. *Consider a process on a tree with a rate matrix over an alphabet with $K+1$ states and equilibrium frequency $\boldsymbol{\pi} > 0$. Setting $U := 1/\min_i \pi_i - 1$, for any clade R, the following bounds hold assuming stationarity [s] and/or reversibility [r].*

a) $M^{\partial}(R) \leq U$ *for any pattern* ∂.

b) $\hat{M}(R) \leq U$ *for any alignment.*

c) *[s]* $M(R) \leq K$.

d) *[r]* $|\hat{M}_{kl}(R)| \leq \sqrt{\hat{M}_{kk}(R)\hat{M}_{ll}(R)} \leq U$ *for any* $k, l \in [K]$.

e) *[s,r]* $|M_{kl}(R)| \leq \sqrt{M_{kk}(R)M_{ll}(R)} \leq K$ *for any* $k, l \in [K]$.

f) *[s,r]* $\mathrm{Var}[\sum_{k\in S} M^{\partial}_{kk}(R)] \leq U \min(K, U/4)$ *for any subset* $S \subseteq [K]$.

g) *[s,r]* $\mathrm{Var}[M^{\partial}_{kl}(R)] \leq UK$ *for any* $k, l \in [K]$.

Proof.

a) We know that $M^{\partial}(R) = \|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}\|^2_{\pi} - 1$ and $\boldsymbol{\pi} \cdot \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = 1$. If we write $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = (\tilde{\rho}_i)$, then

$$\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}\|^2_{\pi} = \sum_i \tilde{\rho}_i^2 \pi_i \leq \max_i \tilde{\rho}_i \leq 1/\min_i \pi_i. \tag{3.81}$$

b) Since $\hat{M}(R)$ is a weighed mean of memories $M^{\partial}(R)$, it is consequence of item *a*).

c) We know that $M(R) = -1 + \sum_{\partial} \Pr(\partial) \|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}\|^2_{\pi}$. If we write $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = (\tilde{\rho}^{\partial}_i)$, Eq. 3.81 yields

$$\sum_{\partial} \Pr(\partial) \|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}\|^2_{\pi} \leq \sum_{\partial} \Pr(\partial) \max_i \tilde{\rho}^{\partial}_i \leq \sum_{\partial} \Pr(\partial) \sum_i \tilde{\rho}^{\partial}_i. \tag{3.82}$$

Now we write $\boldsymbol{\rho}_{\boldsymbol{\partial}} = (\rho^{\partial}_i)$. Since $\Pr(\partial)\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = \boldsymbol{\rho}_{\boldsymbol{\partial}}$, the last term of Eq. 3.82 equals

$$\sum_{\partial} \Pr(\partial) \sum_i \tilde{\rho}^{\partial}_i = \sum_{\partial} \sum_i \rho^{\partial}_i = \sum_i \sum_{\partial} \rho^{\partial}_i = \sum_i 1 = K + 1, \tag{3.83}$$

where we used the definition $\rho_i^{\partial} = \Pr(\partial \mid i \text{ at node } R)$, and thus $\sum_{\partial} \rho_i^{\partial} = 1$.

d) Apply the Cauchy-Schwartz inequality to sequences $x_{\partial} = \sqrt{n_{\partial}/n}|\boldsymbol{h_k} \cdot \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}|$ and $y_{\partial} = \sqrt{n_{\partial}/n}|\boldsymbol{h_l} \cdot \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}|$. For the second inequality, use the fact that $\hat{M}_{kk}(R) \leq \hat{M}(R) \leq U$.

e) For the first inequality, substitute $n_{\partial}/n = \Pr(\partial)$ in the previous item. For the second inequality, use the fact that $M_{kk}(R) \leq M(R) \leq K$.

f) Inequality $\mathrm{Var}[\sum_{k\in S} M_{kk}^{\partial}(R)] \leq U^2/4$ is a direct consequence of Popoviciu's inequality on variances [Popoviciu, 1935], where we use the lower bound $0 \leq M^{\partial}(R)$ and the upper bound $M^{\partial}(R) \leq U$ from item a). Inequality $\mathrm{Var}[\sum_{k\in S} M_{kk}^{\partial}(R)] \leq UK$ is obtained by doing

$$\mathrm{Var}[\sum_{k\in S} M_{kk}^{\partial}(R)] \leq \mathbb{E}[(\sum_{k\in S} M_{kk}^{\partial}(R))^2] \leq U\mathbb{E}[\sum_{k\in S} M_{kk}^{\partial}(R)] \leq UK, \tag{3.84}$$

where we used the fact that $\sum_{k\in S} M_{kk}(R) \leq M(R) \leq K$ from item c).

g) As in Eq. 3.84, we do

$$\mathrm{Var}[M_{kl}^{\partial}(R)] \leq \mathbb{E}[M_{kl}^{\partial}(R)^2] = \mathbb{E}[M_{kk}^{\partial}(R)M_{ll}^{\partial}(R)] \leq U\mathbb{E}[M_{kk}^{\partial}(R)] \leq UK. \tag{3.85}$$

□

In Prop. 3.8, we state some Cauchy-Schwartz-like bounds of the projected sample coherences of a branch. In particular, we prove that the coherence of a branch cannot be greater than the geometric mean of the memories of the clades induced by this branch. We omit the analogous results for the population coherence, since they are just a particular case where $n_{\partial}/n = \Pr(\partial)$. A similar bound is stated in Prop. 3.9, without assuming reversibility.

Proposition 3.8. *Consider a reversible evolutionary process on a tree. Given branch AB, the following holds for any observed alignment.*

a) $|\hat{C}_k(A;B)| \leq \sqrt{\hat{M}_{kk}(A)\hat{M}_{kk}(B)}$.

b) For any subset $S \subseteq [K]$,

$$\sum_{k\in S}|\hat{C}_k(A;B)| \leq \sqrt{(\sum_{k\in S}\hat{M}_{kk}(A))(\sum_{k\in S}\hat{M}_{kk}(B))}.$$

c) For any subset $S \subseteq [K]$,

$$\sum_{k,l \in S} |\hat{M}_{kl}(A)\hat{M}_{kl}(B)| \leq \sum_{k \in S} \hat{M}_{kk}(A) \sum_{k \in S} \hat{M}_{kk}(B).$$

Proof.

a) Clearly $|C_k^{\partial}(A;B)| = \sqrt{M_{kk}^{\partial A}(A) M_{kk}^{\partial B}(B)}$. Applying the Cauchy-Schwartz inequality to sequences

$$x_{\partial} = \sqrt{M_{kk}^{\partial A}(A) n_{\partial}/n} \ , \ y_{\partial} = \sqrt{M_{kk}^{\partial B}(B) n_{\partial}/n},$$

we obtain

$$|\hat{C}_k(A;B)| \leq \sum_{\partial} \sqrt{M_{kk}^{\partial A}(A) n_{\partial}/n} \sqrt{M_{kk}^{\partial B}(B) n_{\partial}/n} \leq$$
$$\leq \sqrt{\sum_{\partial} M_{kk}^{\partial A}(A) \frac{n_{\partial}}{n} \sum_{\partial} M_{kk}^{\partial B}(B) \frac{n_{\partial}}{n}}.$$

We obtain the result using $\sum_{\partial B} n_{\partial} = n_{\partial A}$, and $\sum_{\partial A} n_{\partial} = n_{\partial B}$.

b) Using item *a*), we can do

$$\sum_{k \in S} |\hat{C}_k(A;B)| \leq \sum_{k \in S} \sqrt{\hat{M}_{kk}(A)} \sqrt{\hat{M}_{kk}(B)} \leq \sqrt{\sum_{k \in S} \hat{M}_{kk}(A)} \sqrt{\sum_{k \in S} \hat{M}_{kk}(B)},$$

where we applied the Cauchy-Schwarz to obtain the last inequality.

c) We know from Prop. 3.7.d that $|\hat{M}_{kl}(A)| \leq \sqrt{\hat{M}_{kk}(A)\hat{M}_{ll}(A)}$ and $|\hat{M}_{kl}(B)| \leq \sqrt{\hat{M}_{kk}(B)\hat{M}_{ll}(B)}$, so it is enough to prove

$$\sum_{k,l \in S} \sqrt{\hat{M}_{kk}(A)\hat{M}_{ll}(A)} \sqrt{\hat{M}_{kk}(B)\hat{M}_{ll}(B)} \leq \sum_{k \in S} \hat{M}_{kk}(A) \sum_{k \in S} \hat{M}_{kk}(B). \quad (3.86)$$

The left hand side of Eq. 3.86 can be rewritten as

$$(\sum_{k \in S} \sqrt{\hat{M}_{kk}(A)\hat{M}_{kk}(B)})^2, \quad (3.87)$$

and thus Eq. 3.86 is consequence of the Cauchy-Schwartz inequality using $x_k = \sqrt{\hat{M}_{kk}(A)}$ and $y_k = \sqrt{\hat{M}_{kk}(B)}$.

□

Appendix 3.C Non-reversible processes

The non-reversible counterparts of three propositions of the article are stated here.

Proposition 3.9. *Consider an evolutionary process over $K+1$ states on a tree rooted at A. Given branch AB, the following holds.*

a) For any observed alignment, $|\hat{C}(A;B)| \leq \sqrt{\hat{M}(A)\hat{M}(B)}$.

b) Assuming stationarity, $|C(A;B)| \leq K$.

Proof.

a) For the π-inner product, the Cauchy-Schwartz inequality reads

$$\langle \tilde{\boldsymbol{\alpha}}_{\partial \boldsymbol{A}} - \mathbf{1}, \tilde{\boldsymbol{\beta}}_{\partial \boldsymbol{B}} - \mathbf{1} \rangle \pi \leq \|\tilde{\boldsymbol{\alpha}}_{\partial \boldsymbol{A}} - \mathbf{1}\|_\pi \|\tilde{\boldsymbol{\beta}}_{\partial \boldsymbol{B}} - \mathbf{1}\|_\pi.$$

This gives

$$|\hat{C}(A;B)| \leq \sum_\partial \|\tilde{\boldsymbol{\alpha}}_{\partial \boldsymbol{A}} - \mathbf{1}\|_\pi \|\tilde{\boldsymbol{\beta}}_{\partial \boldsymbol{B}} - \mathbf{1}\|_\pi n_\partial / n.$$

Applying the Cauchy-Schwartz inequality to sequences

$$x_\partial = \|\tilde{\boldsymbol{\alpha}}_{\partial \boldsymbol{A}} - \mathbf{1}\|_\pi \sqrt{n_\partial / n} \ , \ y_\partial = \|\tilde{\boldsymbol{\beta}}_{\partial \boldsymbol{B}} - \mathbf{1}\|_\pi \sqrt{n_\partial / n},$$

we obtain

$$|\hat{C}(A;B)| \leq \sum_\partial \|\tilde{\boldsymbol{\alpha}}_{\partial \boldsymbol{A}} - \mathbf{1}\|_\pi^2 n_\partial / n \sum_\partial \|\tilde{\boldsymbol{\beta}}_{\partial \boldsymbol{B}} - \mathbf{1}\|_\pi n_\partial / n.$$

Pattern ∂ can be partitioned as $\partial = (\partial A, \partial B)$, and thus $\sum_{\partial B} n_\partial = n_{\partial A}$, and $\sum_{\partial A} n_\partial = n_{\partial B}$, giving the result.

b) Consider an alignment such that $n_\partial / n = \Pr(\partial)$ in item *a*). Then, use the bounds $M(A) \leq K$ and $M(B) \leq K$ of Prop. 3.7.c.

□

Proposition 3.10. *Consider a stationary process on a tree rooted at A. Given a branch AB with true length t^*, it holds that*

$$C(A;B) := \mathbb{E}[C^\partial(A;B) \mid t^* \to \infty] = 0.$$

Proof. We know from Eq. 3.30 that $\Pr(\partial) \to \Pr(\partial A)\Pr(\partial B)$. Recall moreover that Eq. 3.8 implies that $\sum_{\partial A} \boldsymbol{\alpha}_{\partial \boldsymbol{A}} = \sum_{\partial B} \boldsymbol{\beta}_{\partial \boldsymbol{B}} = \mathbf{1}$. Therefore

$$\begin{aligned} C(A;B) &= -1 + \sum_{\partial A, \partial B} \Pr(\partial A)\Pr(\partial B)\langle \tilde{\boldsymbol{\alpha}}_{\partial \boldsymbol{A}}, \tilde{\boldsymbol{\beta}}_{\partial \boldsymbol{B}} \rangle_\pi = \\ &= -1 + \sum_{\partial A, \partial B} \langle \boldsymbol{\alpha}_{\partial \boldsymbol{A}}, \boldsymbol{\beta}_{\partial \boldsymbol{B}} \rangle_\pi = -1 + \sum_{\partial A} \langle \boldsymbol{\alpha}_{\partial \boldsymbol{A}}, \sum_{\partial B} \boldsymbol{\beta}_{\partial \boldsymbol{B}} \rangle_\pi = \\ &= -1 + \langle \sum_{\partial A} \boldsymbol{\alpha}_{\partial \boldsymbol{A}}, \mathbf{1} \rangle_\pi = -1 + \langle \mathbf{1}, \mathbf{1} \rangle_\pi = -1 + 1 = 0. \end{aligned} \tag{3.88}$$

□

Proposition 3.11. *Consider a stationary process on a tree rooted at A. Given a branch AB with true length t^*, the following holds.*

a) The coherence of a branch satisfies

$$\mathrm{Var}[C^\partial(A;B) \mid t^* \to \infty] = \mathbb{E}[C^\partial(A;B) \mid t^* = 0] \leq K$$

b) For any true branch length t^*, it holds that

$$C(A;B) \geq 0.$$

If $C^\partial(A;B)$ is not constant in ∂, then the inequality is strict.

Proof.

a) We know that $\Pr(\partial) \to \Pr(\partial A)\Pr(\partial B)$ (Eq. 3.30), and we proved in Prop. 3.10 that $\mathbb{E}[C^\partial(A;B) \mid t^* \to \infty] = 0$. Thus we can compute

$$\mathrm{Var}[C^\partial(A;B) \mid t^* \to \infty] = \mathbb{E}[(C^\partial(A;B))^2 \mid t^* \to \infty] = \tag{3.89}$$

$$= \sum_{\partial A, \partial B} \Pr(\partial A)\Pr(\partial B)\langle \tilde{\boldsymbol{\alpha}}_{\partial \boldsymbol{A}} - \mathbf{1}, \tilde{\boldsymbol{\beta}}_{\partial \boldsymbol{B}} - \mathbf{1} \rangle_\pi^2 = \tag{3.90}$$

$$= \sum_{\partial A, \partial B} \Pr(\partial A)\Pr(\partial B)\langle \tilde{\boldsymbol{\alpha}}_{\partial \boldsymbol{A}} - \mathbf{1}, \tilde{\boldsymbol{\beta}}_{\partial \boldsymbol{B}} - \mathbf{1} \rangle_\pi (\langle \tilde{\boldsymbol{\alpha}}_{\partial \boldsymbol{A}}, \tilde{\boldsymbol{\beta}}_{\partial \boldsymbol{B}} \rangle_\pi - 1). \tag{3.91}$$

$$\tag{3.92}$$

Notice that $\Pr(\partial A)\Pr(\partial B)\langle \tilde{\boldsymbol{\alpha}}_{\partial \boldsymbol{A}}, \tilde{\boldsymbol{\beta}}_{\partial \boldsymbol{B}} \rangle_\pi = \Pr(\partial \mid t^* = 0)$ from Eq. 3.15. On the other side,

$$\sum_{\partial A, \partial B} \Pr(\partial A)\Pr(\partial B)\langle \tilde{\boldsymbol{\alpha}}_{\partial \boldsymbol{A}} - \mathbf{1}, \tilde{\boldsymbol{\beta}}_{\partial \boldsymbol{B}} - \mathbf{1} \rangle_\pi = \mathbb{E}[C^\partial(A;B) \mid t^* \to \infty] = 0.$$

It follows that Eq. 3.91 can be rewritten as

$$\sum_{\partial A,\partial B} \Pr(\partial \mid t^* = 0)\langle \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} - \mathbf{1}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} - \mathbf{1}\rangle_\pi = \mathbb{E}[C^\partial(A;B) \mid t^* = 0], \tag{3.93}$$

giving the desired equality.

b) We know that $\mathrm{Var}[C^\partial(A;B) \mid t^* \to \infty] \geq 0$, being a strict inequality if $C^\partial(A;B)$ is not constant. Therefore item a) implies that

$$\mathbb{E}[C^\partial(A;B) \mid t^* = 0] \geq 0. \tag{3.94}$$

Since Eq. 3.94 applies for any two nodes A and B as long as the tree is rooted at A, we know that

$$\mathbb{E}[C^\partial(A; e^{Qt}B) \mid t^* = 0] \geq 0,$$

and thus it is enough to prove that, for any $t \geq 0$,

$$\mathbb{E}[C^\partial(A; e^{Qt}B) \mid t^* = 0] = \mathbb{E}[C^\partial(A;B) \mid t^* = t]. \tag{3.95}$$

By developing the first expression, where $\Pr(\partial \mid t^* = 0) = \langle \boldsymbol{\alpha}_{\boldsymbol{\partial A}}, \boldsymbol{\beta}_{\boldsymbol{\partial B}}\rangle_\pi$ due to Eq. 3.15, we obtain

$$\begin{aligned}
&\mathbb{E}[C^\partial(A; e^{Qt}B) \mid t^* = 0] = -1 + \sum_{\partial A,\partial B} \langle \boldsymbol{\alpha}_{\boldsymbol{\partial A}}, \boldsymbol{\beta}_{\boldsymbol{\partial B}}\rangle_\pi \langle \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}, e^{Qt}\tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}\rangle_\pi = \\
&= -1 + \sum_{\partial A,\partial B} \langle \boldsymbol{\alpha}_{\boldsymbol{\partial A}}, e^{Qt}\boldsymbol{\beta}_{\boldsymbol{\partial B}}\rangle_\pi \langle \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}}\rangle_\pi = && (3.96)\\
&= \sum_{\partial A,\partial B} \Pr(\partial \mid t^* = t)\langle \tilde{\boldsymbol{\alpha}}_{\boldsymbol{\partial A}} - \mathbf{1}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{\partial B}} - \mathbf{1}\rangle_\pi = && (3.97)\\
&= \mathbb{E}[C^\partial(A;B) \mid t^* = t], && (3.98)
\end{aligned}$$

as desired.

□

Appendix 3.D The MLE Cannot Detect Saturation

Here we show that, when the MLE of a branch length is finite, we cannot draw any conclusion about the finiteness of the true branch length.

Consider an alignment generated from a stationary and reversible process on a tree. Consider a branch AB on this tree with true length t^* and the log-likelihood

$L(t)$ of branch AB having length t (Eq. 3.60). The absolute maximum $\hat{t}$ of function $L(t)$ is the MLE.

We can describe the MLE $\hat{t}$ under the null hypothesis $t^* \to \infty$ using the dominant sample coherence $\hat{\delta}$ defined in Prop. 3.4. Consider the eigenvalues $0 > \lambda_1 \geq \cdots \geq \lambda_K$ of rate matrix Q, where λ_1 has multiplicity $D \leq K$. Assuming $t^* \to \infty$, the dominant sample coherence $\hat{\delta}$, considered as a rv, satisfies that

$$\mathbb{E}[\hat{\delta} \mid t^* \to \infty] = \sum_{k \in [D]} \mathbb{E}[C_k^{\partial}(A; B) \mid t^* \to \infty] = 0, \tag{3.99}$$

where we used Prop. 3.2.a. We define the variance $\sigma^2 := \mathrm{Var}[\hat{\delta} \mid t^* \to \infty]$.

Since we assume that the sites of an alignment are generated independently, the integers n_{∂} are multinomially distributed with probabilities $\Pr(\partial)$. Therefore, for large n, each observed quantity n_{∂} can be approximated as the outcome a normal distribution, as also the sample dominant coherence, which is a linear combination of the integers n_{∂}. In particular, using Eq. 3.99, it follows that the distribution of statistical $\hat{\delta}$ assuming $t^* \to \infty$ can be approximated as

$$\hat{\delta} \sim N(0, \sigma^2), \tag{3.100}$$

where $N(\mu, \sigma^2)$ denotes the normal distribution. In particular,

$$\Pr(\hat{\delta} \neq 0 \mid t^* \to \infty) \approx 1, \tag{3.101}$$

and thus almost surely $\hat{\delta} n e^{\lambda_1 t}$ is the dominant exponential decay of $L(t)$, as implied by Prop. 3.4.b. Moreover, by symmetry of the normal distribution around its mean,

$$\Pr(\hat{\delta} > 0 \mid t^* \to \infty) \approx 1/2. \tag{3.102}$$

When $\hat{\delta} > 0$, the MLE $\hat{t}$ is finite from Corollary 3.5. Consequently,

$$\Pr(\hat{t} \text{ is finite} \mid t^* \to \infty) \geq \Pr(\hat{\delta} > 0 \mid t^* \to \infty) \approx 1/2, \tag{3.103}$$

no matter how large n is. All in all, under the null hypothesis $t^* \to \infty$, a finite MLE $\hat{t}$ is uninformative about the finiteness of the true time t^*.

Appendix 3.E Examples of Multiple Maxima

Here we construct rate matrices inducing log-likelihoods $L(t)$ with multiple maxima. In Subsection 3.E.1, we describe an example of multiple maxima in one entry of the matrix exponential e^{Qt}, and then in Subsection 3.E.2 we give more realistic examples between two SIV sequences.

These examples show that branch length optimization in a phylogenetic tree poses the additional problem of distinguishing local maxima from global maxima. In particular, in Subsection 3.E.2, multiple maxima were present in at least 0.35% (7/2000) and 1.25% (25/2000) of random rate matrices. Thus the occurrence of multiple maxima when studying alignments close to saturation is not negligible and should be taken care of for the Newton method.

3.E.1 Maxima for one entry of the matrix exponential

The numerical part of this subsection can be found in Supp. Notebook 3. Consider the orthogonal base of right eigenvectors

$$U = \begin{bmatrix} 1 & a & -1 & -c \\ 1 & -a & -1 & c \\ 1 & c & 1 & a \\ 1 & -c & 1 & -a \end{bmatrix}, \tag{3.104}$$

where $c^2 = 2 - a^2$, making the module of each right eigenvector be 4. It holds that $U^{-1} = U^T/4$, in particular $\boldsymbol{\pi} = \mathbf{1}/4$. Our reversible rate matrix will be

$$Q := U(\delta D^*)U^T/4, \tag{3.105}$$

where $D^* := \mathrm{Diag}[0, t_1, t_2, t_3]$ for negative real numbers t_i, and δ is chosen so that $\sum_i q_{ii}\pi_i = \mathrm{Tr}(Q)/4 = -1$, making the expected number of substitutions per unit of time be 1. Thus we can write $D := \delta D^* =: \mathrm{Diag}[0, \lambda_1, \lambda_2, \lambda_3]$, where $\lambda_i = t_i\delta$ and we assume $0 > \lambda_1 > \lambda_2 > \lambda_3$. Note that Q is symmetrical, making it easier to handle.

The eigenvector with eigenvalue λ_1 is $\boldsymbol{v_1}^T := (a, -a, c, -c)^T$. From Prop. 3.1.b, we know that

$$e^{Qt} = (p_{ij}(t)) = \frac{1}{4}\mathbf{1}\mathbf{1}^T + \frac{1}{4}\sum_{k\in[3]} \boldsymbol{v_k}\boldsymbol{v_k}^T e^{\lambda_k t}. \tag{3.106}$$

Assuming $\sqrt{2} > a > 0$, the sign of every entry of $\boldsymbol{v_1}\boldsymbol{v_1}^T/4$ is determined, and thus we know whether the dominant exponential decay $v_1^i v_1^j e^{\lambda_1 t}/4$ of $p_{ij}(t)$ increases or

decreases for large t. In particular, $v_1^1 v_1^2/4 = -a^2/4 < 0$, and thus $p_{12}(t)$ has a local maximum as $t \to \infty$.

It remains to determine $0 > \lambda_1 > \lambda_2 > \lambda_3$ and $\sqrt{2} > a > 0$ such that Q does not have negative off-diagonal entries and moreover $p_{12}(t)$ has a finite local maximum point. A satisfying determination is $(t_1, t_2, t_3) = (-1, -1.7, -6)$ and $a = 0.2$. The existence of a finite local maximum point in entry $p_{12}(t)$ is clear from Figure 3.15.

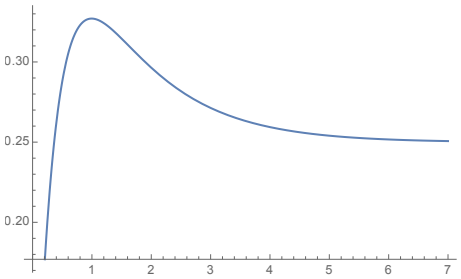

Figure 3.15: Graphic $(t, p_{12}(t))$ for $t < 7$. Entry $p_{12}(t)$ has an absolute maximum at $t \simeq 1.1$. Moreover, by construction $p_{12}(t)$ has a local maximum when $t \to \infty$, not visible in this figure.

3.E.2 Maxima for two SIV sequences

For the numerical analysis of this subsection, see Supp. Notebook 4 and Supp. File `SIVTwoSequences.txt`. We extract two SIV sequences from the ENV gene alignment of year 2018 of an HIV database [LANL, 2020]. If we remove ambiguous sites, SIV1 and SIV2 have a Hamming distance per site of 0.729295, suggesting the possibility of saturation.

We generated random reversible rate matrices and checked whether a finite local maximum point of $L(t)$ existed that was smaller than L_∞, occurring in 0.35% (7/2000) of cases. The rate matrix Q_{AtInf} induces the log-likelihood $L(t)$ shown in Figure 3.16.

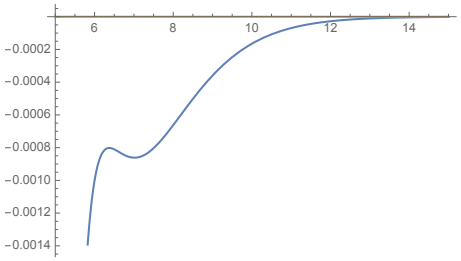

Figure 3.16: Graphic $(t, L(t) - L_\infty)$ for $5 < t < 15$ assuming rate matrix Q_{AtInf} for the alignment of SIV1 and SIV2. Matrix Q_{AtInf} can be found in Supp. Notebook 4.

Moreover, we numerically searched for rate matrices inducing two local maxima, occurring in $12,5\%$ (25/2000) of cases. The rate matrix $Q_{2Maxima}$ induces the log-likelihood $L(t)$ shown in Figure 3.17.

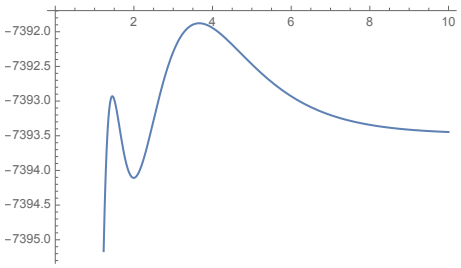

Figure 3.17: Graphic $(t, L(t))$ for $1 < t < 10$ assuming rate matrix $Q_{2Maxima}$ for the alignment of SIV1 and SIV2. The log-likelihood $L(t)$ has maximum points $t_1 \approx 1.45$ and $t_2 \approx 3.66$. Matrix $Q_{2Maxima}$ can be found in Supp. Notebook 4.

Appendix 3.F How to estimate the variance of the coherence

In Eq. 3.68, we stated that the saturation coherence c_S can be approximated as

$$c_S \approx z_\alpha \sqrt{\frac{1}{n} \sum_{k,l \in S} M_{kl}(A) M_{kl}(B)}.$$

If a clade A has a small number of leaves, then it is possible to compute $M_{kl}(A)$ numerically. Notably, if clade A is composed by a single sequence, we know that $M_{kl}(A) = \delta_{kl}$ (Prop. 3.6.f). In more complicated instances, many alternatives are possible. For example, enumerated increasingly in their computing cost and their sharpness, we could do the following.

1. Use the upper bound $\mathrm{Var}[\hat{\delta} \mid t^* \to \infty] \leq K^2/n$ (Prop. 3.3.a). Reject the null hypothesis $t^* \to \infty$ if $\hat{\delta} > K/\sqrt{n}$.
   If $D = [K]$, use the sharper bound $\mathrm{Var}[\hat{C}(A;B) \mid t^* \to \infty] \leq K/n$ (Prop. 3.11.b) and reject $t^* \to \infty$ if $\hat{C}(A;B) > \sqrt{K/n}$.

2. Use the upper bound of Eq. 3.51

$$\mathrm{Var}[\hat{\delta} \mid t^* \to \infty] \leq 1/n(\sum_{k \in S} M_{kk}(A))(\sum_{k \in S} M_{kk}(B)).$$

   Then build confidence intervals of radius ϵ around $\hat{a} := \sum_{k \in S} \hat{M}_{kk}(A)$ using $\mathrm{Var}[\sum_{k \in S} \hat{M}_{kk}(A)] \leq U \min(K, U/4)/n$ (Prop. 3.7.f). Reject $t^* \to \infty$ if

$\hat{\delta} > \sqrt{(\hat{a}+\epsilon)(\hat{b}+\epsilon)/n}$.

3. Build confidence intervals around each sample memory $\hat{M}_{kl}(A)$ and $\hat{M}_{kl}(B)$ using the fact that $\mathrm{Var}[\hat{M}^{\partial}_{kl}(R)] \leq UK/n$ (Prop. 3.7.g). Use these confidence intervals to construct a confidence interval around $\hat{v} := \sum_{k,l\in S} \hat{M}_{kl}(A)\hat{M}_{kl}(B)$ with upper bound $\hat{v}+\epsilon$. Reject $t^* \to \infty$ if $\hat{\delta} > \sqrt{(\hat{v}+\epsilon)/n}$.

All these upper bounds are more accurate as K^2/n decreases. Another alternative, with a potentially much higher variance, follows from the upper bound $\mathrm{Var}[\sum_{k\in S} C^{\partial}_k(A;B)] \leq \mathbb{E}[(\sum_{k\in S} C^{\partial}_k(A;B))^2]$. To estimate this expected value, use the estimator $\hat{e} := \sum_{\partial} n_{\partial}/n(\sum_{k\in S} C^{\partial}_k(A;B))^2$, whose variance is upper bounded by $U^4/(4n)$, as implied by Popoviciu's inequality (see Popoviciu [1935]) with upper bound U^2. If we build the confident interval $(\hat{e}-\epsilon, \hat{e}+\epsilon)$, reject $t^* \to \infty$ if $\sum_{k\in S} \hat{C}_k(A;B) > (\hat{e}+\epsilon)/\sqrt{n}$.

References in this chapter

J. W. Archie. A Randomization Test for Phylogenetic Information in Systematic Data. *Systematic Biology*, 38(3):239–252, 09 1989. ISSN 1063-5157. doi: 10.2307/2992285. URL https://doi.org/10.2307/2992285.

D. A. Duchêne, N. Mather, C. Van Der Wal, and S. Y. W. Ho. Excluding Loci With Substitution Saturation Improves Inferences From Phylogenomic Data. *Systematic Biology*, 71(3):676–689, 09 2021. ISSN 1063-5157. doi: 10.1093/sysbio/syab075. URL https://doi.org/10.1093/sysbio/syab075.

B. Efron. Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer, 1992.

Encyclopedia of Mathematics. Encyclopedia of Mathematics, by Springer Verlag GmbH and European Mathematical Society. URL: https://encyclopediaofmath.org/wiki/Newton_method. Accessed on 2021-09-09, 2021a.

Encyclopedia of Mathematics. Encyclopedia of Mathematics, by Springer Verlag GmbH and European Mathematical Society. URL: https://encyclopediaofmath.org/wiki/Maximum-likelihood_method. Accessed on 2021-09-09, 2021b.

J. Felsenstein. Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, 17(6):368–376, 1981. ISSN 1432-1432. doi: 10.1007/BF01734359.

J. Felsenstein. Confidence limits on phylogenies: an approach using the bootstrap. *evolution*, 39(4):783–791, 1985.

J. Felsenstein. *Inferring phylogenies*, chapter 16. Sinauer Associates, 2004.

B. Foley, T. Leitner, C. Apetrei, B. Hahn, I. Mizrachi, J. Mullins, A. Rambaut, S. Wolinsky, and B. Korber. HIV Sequence Compendium 2018. Theoretical

Biology and Biophysics Group, Los Alamos National Laboratory, NM, LA-UR 18-25673, 2020.

X. Gu, Y. X. Fu, and W. H. Li. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Molecular Biology and Evolution*, 12(4):546–557, 07 1995. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a040235. URL `https://doi.org/10.1093/oxfordjournals.molbev.a040235`.

D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, and L. S. Vinh. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, 35(2):518–522, 10 2017. ISSN 0737-4038. doi: 10.1093/molbev/msx281.

M. Kimura. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution*, 16(2):111–120, 1980.

LANL. Los Alamos National Laboratory. http://www.hiv.lanl.gov/, 2020. Accessed: 2020-12-07.

D. A. Levin and Y. Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

C. Manuel. An upper bound of the information flow from children to parent node on trees. *arXiv preprint arXiv:2204.10618*, 2022.

B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, and R. Lanfear. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, 37 (5):1530–1534, 02 2020. ISSN 0737-4038. doi: 10.1093/molbev/msaa015. URL `https://doi.org/10.1093/molbev/msaa015`.

J. Neyman and E. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London*, IX(231): 289–337, 1933. doi: 10.1098/rsta.1933.0009.

T. Popoviciu. Sur les équations algébriques ayant toutes leurs racines réelles. *Mathematica*, 9(129-145):20, 1935.

M. Salemi. Genetic distances and nucleotide substitution models: practice. In P. Lemey, M. Salemi, and V. Anne-Mieke, editors, *The Phylogenetic Handbook: a Practical Approach to Phylogenetic Analysis and Hypothesis Testing*, pages 125–141. Cambridge University Press, Cambridge, 2 edition, 2009. doi: 10.1017/CBO9780511819049.006.

C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.

M. E. Siddall. Success of Parsimony in the Four-Taxon Case: Long-Branch Repulsion by Likelihood in the Farris Zone. *Cladistics*, 14(3):209–220, 1998. ISSN 0748-3007. doi: https://doi.org/10.1006/clad.1998.0063. URL `https://www.sciencedirect.com/science/article/pii/S0748300798900639`.

A. Stamatakis. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313, 01 2014. ISSN 1367-4803. doi: 10.1093/bioinformatics/btu033. URL `https://doi.org/10.1093/bioinformatics/btu033`.

K. Strimmer and A. von Haeseler. Genetic distances and nucleotide substitution models. In P. Lemey, M. Salemi, and V. Anne-Mieke, editors, *The Phylogenetic Handbook: a Practical Approach to Phylogenetic Analysis and Hypothesis Testing*, pages 111–141. Cambridge University Press, Cambridge, 2 edition, 2009. doi: 10.1017/CBO9780511819049.006.

X. Xia, Z. Xie, M. Salemi, L. Chen, and Y. Wang. An index of substitution saturation and its application. *Molecular phylogenetics and evolution*, 26(1):1–7, 2003.

Z. Yang. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular evolution*, 39(3):306–314, 1994.

Chapter 4

An Upper Bound of the Information Flow from Children to Parent Node on Trees

Publication history and status

Submitted on October 11th 2022 to the Journal of Mathematical Biology. A previous version of this chapter is available in arXiv.

C. Manuel. An upper bound of the information flow from children to parent node on trees. *arXiv preprint arXiv:2204.10618*, 2022

Abstract

We consider an evolutionary process where a state was transmitted from the root of a tree towards its leaves. The states at the leaves are observed, while at deeper nodes we can compute the likelihood of each state given the observation using Felsenstein's Pruning algorithm. In this sense, information flows from child nodes towards the parent node.

Here we find an upper bound of this children-to-parent information flow. To do so, first we introduce the memory vector, whose norm quantifies whether all states have the same likelihood. Then we find conditions such that the norm of the memory vector at the parent node can be linearly bounded by the sum of norms at the child nodes.

We also describe the reconstruction problem of estimating the ancestral state at the root given the observation at the leaves. We infer sufficient conditions under which the original state at the root cannot be confidently reconstructed using the

observed leaves, assuming that the number of levels from the root to the leaves is large.

4.1 Introduction

Clearly we know more about the genome of an observed species than about the genome of its unobserved ancestors. Generalizing this observation, it is natural to ask: Does our capability to identify ancestor genomes decrease as we go deeper into the phylogenetic tree? An answer to this question depends on two opposing tendencies: Intuitively, ancestor identification decreases if the observed species have mutated more with respect to the ancestor, although it increases if more descendant species are available to reconstruct the ancestor identity.

To formalize this intuition, assume a model of evolution where sites mutate independently as determined by some Markov matrices on a phylogenetic tree. Given the observed species, Felsenstein's Pruning algorithm allows to compute the likelihood of each nucleotide at a site of an ancestral sequence. Recall that the Pruning algorithm is a recursion where the likelihoods at each node are computed using the likelihoods at its child nodes [Felsenstein, 1981]. Therefore we can imagine a flow of information (the amount of identification) moving from the observed species towards the deeper nodes of the tree. On every edge, this flow has a bigger leak as the number of mutations grows, although more children will provide a bigger information flow towards the root.

As we can see, the measurement of identification is a fundamental aspect of this description. However, the choice of a natural measure of identification is unclear. For example, imagine that, at a site of an ancestor species, the likelihoods of the DNA nucleotides A, G, T, C are respectively $(0.2, 0.1, 0.1, 0.1)$. We see that the most likely identity of this site is A (adenine), but how worse identified is this site in comparison to a site with likelihoods $(0.2, 0, 0, 0)$?

In this work, we introduce a measure of identification, namely the $L_2(\pi)$-norm of the memory vector. This measure is applied to the Pruning algorithm to upper bound the flow of information on a binary tree assuming a stationary and reversible process. Then we analyze the general d-ary tree assuming that enough mixing has occurred. As a theoretical application of our results, we give conditions for the unsolvability of the reconstruction problem.

Notably, here we consider only the maximum likelihood (ML) reconstruction method, since it has an optimal probability of a correct reconstruction (Theorem 17.2 of Guiasu [1977]). In particular, we give a central role to the likelihood vector,

which vertebrates current software implementations of ML reconstruction [Minh et al., 2020, Stamatakis, 2014]. Different reconstruction methods are analyzed by Mossel and Peres [2003] and Gascuel and Steel [2014]. Similar bounds to ours, directly depending on the entries of the rate matrix instead of the eigenvalues, are stated by Mossel and Steel [2007].

This paper is structured as follows. The evolutionary process is introduced in Section 4.2. Then, in Section 4.3, we describe the Pruning algorithm to compute the likelihood vectors. In Section 4.4 we describe a measure of the information carried by a likelihood vector, namely the $L_2(\pi)$-norm of the memory vector. As an instrumental step, we rewrite the likelihood vector using the memory vector in Section 4.5.

In Section 4.6, we bound the information flow using a common setup, namely a stationary and reversible continuous-time model of evolution on a binary tree. Then we generalize this result for a d-ary tree in Section 4.7. Our core results are Theorems 4.4 and 4.7, where we bound the memory vector norm at a parent node by the memory vector norms at the child nodes. These theorems can be applied to trees with a different transition matrix for every child node. The possible definitions of unsolvability are explained in Section 4.8. Using the bounds of Theorems 4.4 and 4.7, in Section 4.9 we give sufficient conditions on matrix P such that the reconstruction problem is unsolvable.

4.2 The evolutionary process

We consider the following evolutionary process. At the root R of a tree, sample a state from alphabet $\mathbb{A} = \{0, 1, \cdots, K\}$ following distribution $\boldsymbol{\mu} > 0$. The tree is d-ary, meaning that every node has at most d children.

The sampled state at R is then transmitted independently to each child node of R through a noisy channel, described by an aperiodic, irreducible Markov matrix P. In phylogenetics, the noisy channel is normally called *transition matrix*, as we will do in this work. Thus in the transition matrix $P = (p_{ij})$, where $i, j \in \mathbb{A}$, entry p_{ij} is the probability of state i mutating to state j during transmission.

The transmission of states continues analogously until a state is transmitted to every leaf of the tree. The ordered set of states at the leaves is called a *pattern*, represented by symbol ∂. See Figure 4.1 for an example of this process with $\mathbb{A} = \{0, 1, 2, 3\}$. In this example, since the unique path from the root R to each of the leaves has exactly 2 edges, we say that the tree has 2 levels. If the tree has m leaves,

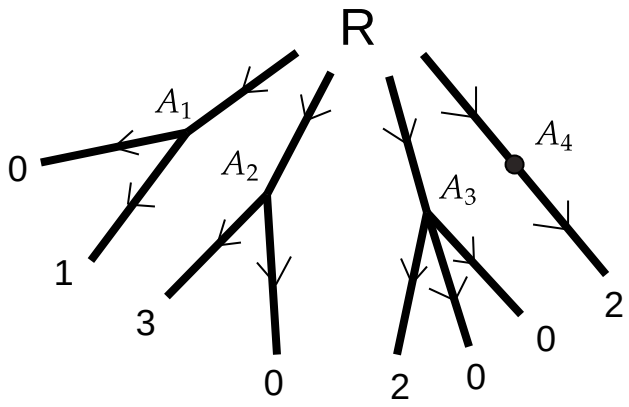

Figure 4.1: Example of an evolutionary process on a 2-level, incomplete 4-ary tree. The ancestral state $i \in \mathbb{A}$ at the root R is sent independently to each node $A_1, \cdots, A_4$. The probability of i mutating to k during this transmission is p_{ik}. Then, the state i_c at each node A_c is sent independently to each of its child leaves, with a probability $p_{i_c k}$ of mutating to k during transmission. The observed pattern is $\partial = 01302002$, while the four subpatterns determined by clades $A_1, \cdots, A_4$ are $\partial_1 = 01$, $\partial_2 = 30$, $\partial_3 = 200$ and $\partial_4 = 2$.

then the set of possible patterns is $\mathbb{A}^m$, that is, the strings of length m over alphabet $\mathbb{A}$. Thus in Figure 4.1, the set of possible patterns is $\mathbb{A}^8$.

4.3 Likelihoods on a d-ary Tree

The Pruning algorithm to recursively compute the likelihood of a model given a pattern was introduced by Felsenstein [1981]. In this section, we rephrase the Pruning algorithm for a d-ary tree with a different transition matrix for each child node.

We want to compute the likelihood of each state at the root R given pattern ∂. To that end, we define the likelihood vector at the root as

$$\boldsymbol{\rho_\partial} := (\Pr(\partial \mid R = 0), \cdots, \Pr(\partial \mid R = K)). \tag{4.1}$$

Note that the likelihood vectors satisfy the equality

$$\sum_{\partial \in \mathbb{A}^m} \boldsymbol{\rho_\partial} = \mathbf{1}, \tag{4.2}$$

where $\mathbf{1}$ is the vector of 1's.

Figure 4.2 shows the root R and its d transition matrices $P_1, \cdots, P_d$ with their respective child nodes $A_1, \cdots, A_d$. Abusing of the notation, we define clade A_c as the subtree rooted at node A_c containing all descendant nodes from A_c. The leaves

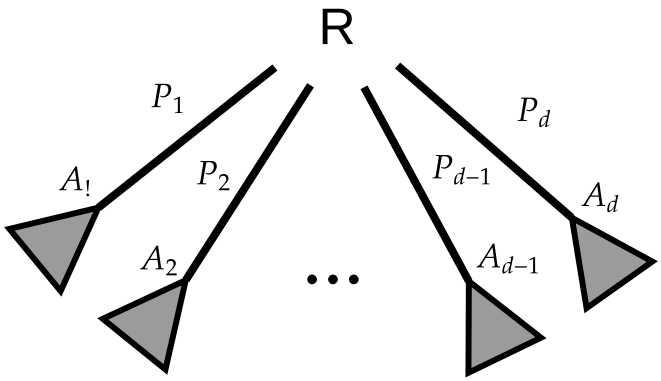

Figure 4.2: Diagram of a d-ary tree. The state at node A_c, for $c \in [d]$, was obtained by transmitting the state at the root R using transition matrix P_c. The likelihood vector $\boldsymbol{\rho_\partial}$ at R satisfies $\boldsymbol{\rho_\partial} = \bigcirc_{c\in[d]} P_c\boldsymbol{\alpha_c}$.

of clade A_c determine subpattern ∂_c. Considered independently, each clade allows to compute a likelihood vector $\boldsymbol{\alpha_c}$ at the root A_c given subpattern ∂_c. More explicitly, we define

$$\boldsymbol{\alpha_c} := (\Pr(\partial_c \mid A_c = 0), \cdots, \Pr(\partial_c \mid A_c = K)). \tag{4.3}$$

The Pruning algorithm consists in expressing $\boldsymbol{\rho_\partial}$ in terms of the likelihood vectors $\boldsymbol{\alpha_1}, \cdots, \boldsymbol{\alpha_d}$. Using the law of total probability, the likelihood vector at the root R considering *only* clade A_c is $P_c\boldsymbol{\alpha_c}$. More explicitly,

$$P_c\boldsymbol{\alpha_c} = (\Pr(\partial_c \mid R = 0), \cdots, \Pr(\partial_c \mid R = K)). \tag{4.4}$$

Now we use the fact that the state at each node A_c was obtained *independently* from the state at R. This implies that $\boldsymbol{\rho_\partial}$ equals the entrywise product of all likelihood vectors $P_c\boldsymbol{\alpha_c}$. To visualize this better, if we focus on state $R = 0$, it holds that

$$\Pr(\partial \mid R = 0) = \prod_{c\in[d]} \Pr(\partial_c \mid R = 0). \tag{4.5}$$

We denote the entrywise product between two vectors $\boldsymbol{v} = (v_i)$ and $\boldsymbol{w} = (w_i)$ as $\boldsymbol{v} \circ \boldsymbol{w} := (v_i w_i)$. Given vectors $\boldsymbol{v_1}, \cdots, \boldsymbol{v_d}$, their entrywise product is denoted as $\bigcirc_{c\in[d]} \boldsymbol{v_c} := \boldsymbol{v_1} \circ \cdots \circ \boldsymbol{v_d}$. With this notation, we can write

$$\boldsymbol{\rho_\partial} = \bigcirc_{c\in[d]} P_c\boldsymbol{\alpha_c}. \tag{4.6}$$

Now we just have to proceed analogously: Every vector $\boldsymbol{\alpha_c}$ can be expressed in

terms of the child nodes of A_c, and so on, until we eventually arrive to the likelihood vectors at the leaves. The likelihood vector at a leaf where state i was observed is the canonical vector $\boldsymbol{e_i}$, composed by 0's except for a 1 at coordinate i. Thus the Pruning algorithm is a recursion based on Eq. 4.6 where the leaves have known likelihood vectors and the computation proceeds towards the root of the tree.

Often one needs to compute $\Pr(\partial)$, that is, the probability of observing pattern ∂. This computation requires the assumption of some prior state distribution $\boldsymbol{\mu} > 0$ at node R. The law of total probability implies that

$$\Pr(\partial) = \boldsymbol{\mu} \cdot \boldsymbol{\rho_\partial}, \tag{4.7}$$

where " $\cdot$ " denotes the Euclidean dot product.

4.4 The memory vector

4.4.1 The equilibrium distribution

The equilibrium distribution of a Markov matrix plays a fundamental role in this work. The necessary preliminaries about the equilibrium distribution can be summarized as follows (see Levin and Peres [2017], Chapter 1 and Section 12.1).

Proposition 4.1. *For an irreducible and aperiodic Markov matrix P, the following holds:*

a) Matrix P has a unique equilibrium distribution $\boldsymbol{\pi} > 0$, defined by equation $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$.

b) Matrix P has eigenvalue $\theta_0 = 1$ with algebraic multiplicity 1, and the other eigenvalues $\{\theta_1, \cdots, \theta_K\} \subset \mathbb{C}$ have a module smaller than 1. We assume $1 > |\theta_1| \geq \cdots \geq |\theta_K|$.

c) If P is reversible, all eigenvalues are real. Moreover, matrix P has an orthogonal basis of right eigenvectors $\boldsymbol{v_k}$ and left eigenvectors $\boldsymbol{h_k^T}$ such that $\boldsymbol{h_k} = \boldsymbol{\pi} \circ \boldsymbol{v_k}$ for $k \in [0, K]$, where $\boldsymbol{v_0} = \mathbf{1}$, $\boldsymbol{h_0} = \boldsymbol{\pi}$ and

$$P = \mathbf{1}\boldsymbol{\pi}^T + \boldsymbol{v_1}\boldsymbol{h_1^T}\theta_1 + \cdots + \boldsymbol{v_K}\boldsymbol{h_K^T}\theta_K.$$

4.4.2 General properties of the norm

For some definitions of this subsection, we follow the notation of Levin and Peres [2017], Section 12.5. Given distribution $\boldsymbol{\pi} > 0$, for any two vectors $\boldsymbol{x} = (x_i)$ and

$\boldsymbol{y} = (y_i)$ where $i \in \mathbb{A}$, we define the π-inner product as

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_\pi := \boldsymbol{\pi} \cdot (\boldsymbol{x} \circ \boldsymbol{y}) = \sum_i \pi_i x_i y_i. \tag{4.8}$$

The π-inner product induces the $L_2(\pi)$-norm, defined as

$$\|\boldsymbol{x}\|_\pi := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle_\pi} = \sqrt{\sum_i \pi_i x_i^2}. \tag{4.9}$$

As in every inner product space, the triangle inequality holds. Explicitly,

$$\|\boldsymbol{x} + \boldsymbol{y}\|_\pi \leq \|\boldsymbol{x}\|_\pi + \|\boldsymbol{y}\|_\pi. \tag{4.10}$$

Denote the Euclidean norm of $\boldsymbol{x}$ as $\|\boldsymbol{x}\| := \sqrt{\sum_i x_i^2}$, and the uniform norm as $\|\boldsymbol{x}\|_\infty := \max_i |x_i|$. A weighted sum is not greater than its biggest summand, and therefore $\|\boldsymbol{x}\|_\pi \leq \|\boldsymbol{x}\|_\infty$. Moreover, since $|x_i| \leq \|\boldsymbol{x}\|$ for all i and $\sqrt{\min_i \pi_i} \|\boldsymbol{x}\| \leq \|\boldsymbol{x}\|_\pi$, it holds that

$$\|\boldsymbol{x}\|_\pi \leq \|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\| \leq \frac{\|\boldsymbol{x}\|_\pi}{\sqrt{\min_i \pi_i}}. \tag{4.11}$$

An important reason to use the $L_2(\pi)$-norm is that a vector cannot increase its $L_2(\pi)$-norm when we centralize its entries with weights $\boldsymbol{\pi}$, as described in the following proposition.

Proposition 4.2 (Centralizing inequality)**.** *Given vector $\boldsymbol{x}$, the $L_2(\pi)$-norm satisfies*

$$\|\boldsymbol{x} - \mathbf{1}\langle \boldsymbol{x}, \mathbf{1} \rangle_\pi\|_\pi^2 = \|\boldsymbol{x}\|_\pi^2 - \langle \boldsymbol{x}, \mathbf{1} \rangle_\pi^2.$$

In particular, it holds that

$$\|\boldsymbol{x} - \mathbf{1}\langle \boldsymbol{x}, \mathbf{1} \rangle_\pi\|_\pi \leq \|\boldsymbol{x}\|_\pi.$$

Proof. Developing the π-inner product, we obtain

$$\Big\langle \boldsymbol{x} - \mathbf{1}\langle \boldsymbol{x}, \mathbf{1} \rangle_\pi, \boldsymbol{x} - \mathbf{1}\langle \boldsymbol{x}, \mathbf{1} \rangle_\pi \Big\rangle_\pi = \langle \boldsymbol{x}, \boldsymbol{x} \rangle_\pi - 2\langle \boldsymbol{x}, \mathbf{1} \rangle_\pi^2 + \langle \boldsymbol{x}, \mathbf{1} \rangle_\pi^2 \langle \mathbf{1}, \mathbf{1} \rangle_\pi. \tag{4.12}$$

Since $\langle \mathbf{1}, \mathbf{1} \rangle_\pi = 1$, it follows that

$$\|\boldsymbol{x} - \mathbf{1}\langle \boldsymbol{x}, \mathbf{1} \rangle_\pi\|_\pi^2 = \|\boldsymbol{x}\|_\pi^2 - \langle \boldsymbol{x}, \mathbf{1} \rangle_\pi^2,$$

as desired. To obtain the inequality, use the fact that $\langle \boldsymbol{x}, \mathbf{1} \rangle_\pi^2 \geq 0$, giving

$$\|\boldsymbol{x} - \mathbf{1}\langle \boldsymbol{x}, \mathbf{1} \rangle_\pi\|_\pi^2 \leq \|\boldsymbol{x}\|_\pi^2.$$

□

4.4.3 Properties of the memory vector

Given a pattern ∂, consider the likelihood vector $\boldsymbol{\rho_\partial}$ at node R. Consider moreover the equilibrium distribution $\boldsymbol{\pi} > 0$ of the Markov matrix P.

Definition 4.1. *The normalized likelihood vector at R given ∂ is defined as*

$$\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} := \frac{\boldsymbol{\rho_\partial}}{\boldsymbol{\rho_\partial} \cdot \boldsymbol{\pi}}.$$

Moreover, we say that the distribution $\boldsymbol{r_\pi} := \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} \circ \boldsymbol{\pi}$ is the posterior distribution at R assuming stationarity.

Recall that, under stationarity, we have $\boldsymbol{\mu} = \boldsymbol{\pi}$, and thus $\Pr(\partial) = \boldsymbol{\rho_\partial} \cdot \boldsymbol{\pi}$ as implied by Eq. 4.7. Thus $\boldsymbol{r_\pi}$ is indeed the posterior distribution given ∂, as implied by Bayes' theorem.

To see the usefulness of the normalized likelihood vector, imagine that, during the process of transmission through transition matrix P, node R received its state from a node called PR, as shown in Fig. 4.3. Then, given ∂, node PR has likelihood vector $P\boldsymbol{\rho_\partial}$. Importantly, since $\boldsymbol{\pi}^T P \boldsymbol{\rho_\partial} = \boldsymbol{\pi}^T \boldsymbol{\rho_\partial} = \boldsymbol{\pi} \cdot \boldsymbol{\rho_\partial}$, it follows that the normalized likelihood vector at node PR is $P\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$. Thus vector $P\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$ is already normalized, that is, matrix P is an endomorphism of normalized likelihood vectors.

We have $\langle \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}, \mathbf{1} \rangle_\pi = \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} \cdot \boldsymbol{\pi} = 1$, and thus if $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} := (\tilde{\rho}_\partial^i)$, then for all $i \in \mathbb{A}$,

$$\tilde{\rho}_\partial^i \leq 1/\pi_i \leq 1/\min_i \pi_i, \tag{4.13}$$

A useful upper bound of the norm of a normalized likelihood vector is

$$\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}\|_\pi^2 = \boldsymbol{r}_{\boldsymbol{\pi}}^T \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} \leq \max_i \tilde{\rho}_\partial^i \leq \frac{1}{\min_i \pi_i}. \tag{4.14}$$

However, we are more interested in the norm of $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}$. Intuitively, as $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$ approaches $\mathbf{1}$, all ancestral states become equally likely, that is, pattern ∂ becomes less informative. This motivates the following definition.

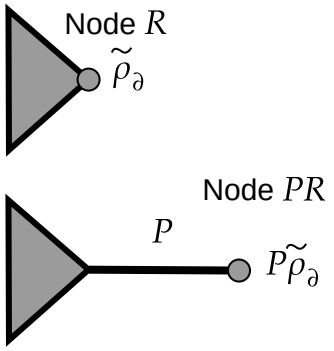

Figure 4.3: Diagram of the action of transition matrix P. We denote by PR the node from which node R received its state through transition matrix P. Given ∂, if node R has likelihood vector $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$, then node PR has likelihood vector $P\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$.

Definition 4.2. *We say that vector $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}$ is the memory vector at node R given ∂.*

When needed, the memory vector will be denoted as $\boldsymbol{m}$, with adequate indices depending on the context. The memory vector satisfies

$$0 \leq \|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi^2 = \langle \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}, \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1} \rangle_\pi = \|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}\|_\pi^2 - 1 \leq \frac{1}{\min_i \pi_i} - 1, \tag{4.15}$$

that is, the $L_2(\pi)$-norm of any memory vector is upper bounded by $\sqrt{1/\min_i \pi_i - 1}$. Assuming stationarity, the expected value of the norm of the memory vector can be bounded using the following result, proved in Appendix 4.B.

Proposition 4.3. *Consider a stationary process on a tree with m leaves. If the alphabet $\mathbb{A}$ has $K+1$ states, then it holds that*

$$\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi] := \sum_{\partial \in \mathbb{A}^m} \Pr(\partial) \|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi \leq \sqrt{K}.$$

4.4.4 The reversible case

If the transition matrix P is reversible, then the eigendecomposition of Prop. 4.1.c applies, yielding simple ways to describe the $L_2(\pi)$-norm and the action of P. Indeed, using decomposition $I = \mathbf{1}\boldsymbol{\pi}^T + \sum_{k \in [K]} \boldsymbol{v_k} \boldsymbol{h_k}^T$ where $\boldsymbol{h_k} = \boldsymbol{\pi} \circ \boldsymbol{v_k}$, we can rewrite the $L_2(\pi)$-norm as

$$\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}\|_\pi^2 = \boldsymbol{r}_{\boldsymbol{\pi}}^T \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = \boldsymbol{r}_{\boldsymbol{\pi}}^T I \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = 1 + \sum_{k \in [K]} (\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} \cdot \boldsymbol{h_k})^2 \geq 1. \tag{4.16}$$

Regarding the action of a reversible Markov matrix P on a normalized likelihood vector $\tilde{\boldsymbol{\alpha}}$, it can be alternatively described as

$$P\tilde{\boldsymbol{\alpha}} = \mathbf{1} + \sum_{k\in[K]} \theta_k(\tilde{\boldsymbol{\alpha}} \cdot \boldsymbol{h_k})\boldsymbol{v_k}, \tag{4.17}$$

where we used the eigendecomposition of Prop. 4.1.c, $P = \mathbf{1}\boldsymbol{\pi}^T + \sum_{k\in[K]} \theta_k \boldsymbol{v_k}\boldsymbol{h_k}^T$. Recall that the eigenvalues θ_k satisfy $1 > |\theta_1| \geq \cdots \geq |\theta_K|$. Using Equations 4.15 and 4.16 with $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = P\tilde{\boldsymbol{\alpha}}$, the memory vector at node PA satisfies

$$\|P\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi^2 = \sum_{k\in[K]} \theta_k^2(\tilde{\boldsymbol{\alpha}} \cdot \boldsymbol{h_k})^2. \tag{4.18}$$

Interestingly, this yields the bound

$$\|P\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi^2 \leq \theta_1^2 \sum_{k\in[K]} (\tilde{\boldsymbol{\alpha}} \cdot \boldsymbol{h_k})^2 = \theta_1^2 \, \|\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi^2, \tag{4.19}$$

or taking the square root,

$$\|P\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi \leq |\theta_1| \, \|\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi. \tag{4.20}$$

Hence the $L_2(\pi)$-norm of the memory vector decreases at least by a factor of $|\theta_1|$ under the action of the reversible matrix P. In Appendix 4.A, we compute a weaker bound if P is not reversible.

4.5 Rewriting the likelihood vectors

In this section, we will rewrite the likelihood vectors described in Section 4.3. We assume that all matrices P_c have the same equilibrium distribution $\boldsymbol{\pi}$. Recall that the normalized likelihood vectors are defined as $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} := \boldsymbol{\rho}_{\boldsymbol{\partial}}/(\boldsymbol{\rho}_{\boldsymbol{\partial}} \cdot \boldsymbol{\pi})$, and analogously $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} := \boldsymbol{\alpha}_{\boldsymbol{c}}/(\boldsymbol{\alpha}_{\boldsymbol{c}} \cdot \boldsymbol{\pi})$. Assuming prior $\boldsymbol{\mu} = \boldsymbol{\pi}$, it holds that $\Pr(\partial) = \boldsymbol{\rho}_{\boldsymbol{\partial}} \cdot \boldsymbol{\pi}$ and $\Pr(\partial_c) = \boldsymbol{\alpha}_{\boldsymbol{c}} \cdot \boldsymbol{\pi}$. Thus we define $\Pr_\pi(\partial) := \boldsymbol{\rho}_{\boldsymbol{\partial}} \cdot \boldsymbol{\pi}$ and $\Pr_\pi(\partial_c) := \boldsymbol{\alpha}_{\boldsymbol{c}} \cdot \boldsymbol{\pi}$

The normalized likelihood vectors are useful to identify the dependence between the subpatterns $\partial_1, \cdots, \partial_d$ composing pattern ∂. Indeed, since $\boldsymbol{\rho}_{\boldsymbol{\partial}} = \bigcirc_{c\in[d]} P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}}$ and $\Pr_\pi(\partial) = \boldsymbol{\rho}_{\boldsymbol{\partial}} \cdot \boldsymbol{\pi}$, it holds that

$$\Pr_\pi(\partial) = \Pr_\pi(\partial_1, \cdots, \partial_d) = \Big(\prod_{c\in[d]} \Pr_\pi(\partial_c)\Big)\Big(\boldsymbol{\pi} \cdot \bigcirc_{c\in[d]} P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}}\Big), \tag{4.21}$$

showing that the dependence factor between observations $\partial_1, \cdots, \partial_d$ is

$\boldsymbol{\pi} \cdot \bigcirc_{c\in[d]} P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}}$. We additionally define the probability of observing ∂ assuming subpattern independence,

$$\Pr_{IND}(\partial) := \prod_{c\in[d]} \Pr_\pi(\partial_c) \tag{4.22}$$

and the dependence factor $D(\partial)$ of the subpatterns of ∂,

$$D(\partial) := \boldsymbol{\pi} \cdot \bigcirc_{c\in[d]} P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}}, \tag{4.23}$$

yielding the relationship

$$\Pr_\pi(\partial) = D(\partial)\Pr_{IND}(\partial). \tag{4.24}$$

By introducing $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} := \boldsymbol{\alpha}_{\boldsymbol{c}}/\Pr_\pi(\partial_c)$ and the memory vectors $\boldsymbol{m}_{\boldsymbol{c}} := P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}$ in equation $\boldsymbol{\rho}_{\boldsymbol{\partial}} = \bigcirc_{c\in[d]} P_c\boldsymbol{\alpha}^{\boldsymbol{c}}_{\boldsymbol{\partial}}$, we obtain

$$\begin{aligned}\frac{\boldsymbol{\rho}_{\boldsymbol{\partial}}}{\Pr_{IND}(\partial)} &= \bigcirc_{c\in[d]} \Big(\mathbf{1} + \boldsymbol{m}_{\boldsymbol{c}}\Big) = \\ &= \mathbf{1} + \sum_{p\in[d]} \sum_{C\in\binom{[d]}{p}} \bigcirc_{c\in C} \boldsymbol{m}_{\boldsymbol{c}}.\end{aligned} \tag{4.25}$$

This equation combined with $\Pr_\pi(\partial) = \boldsymbol{\rho}_{\boldsymbol{\partial}} \cdot \boldsymbol{\pi} = D(\partial)\Pr_{IND}(\partial)$ gives

$$\begin{aligned}D(\partial) = \frac{\Pr_\pi(\partial)}{\Pr_{IND}(\partial)} &= \frac{\boldsymbol{\rho}_{\boldsymbol{\partial}}}{\Pr_{IND}(\partial)} \cdot \boldsymbol{\pi} = \\ &= 1 + \sum_{p\in[d]} \sum_{C\in\binom{[d]}{p}} \boldsymbol{\pi} \cdot \bigcirc_{c\in C} \boldsymbol{m}_{\boldsymbol{c}},\end{aligned} \tag{4.26}$$

where the terms where $p = 1$ vanish, because $\boldsymbol{\pi}^T(P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}) = \mathbf{1} - \mathbf{1} = 0$. The expanded products of Equations 4.25 and 4.26 will be useful to bound information flow in Sections 4.6 and 4.7.

4.6 Information flow on a binary tree

The stationary and reversible process on a binary tree ($d = 2$) using a continuous Markov model of evolution is a common setup for practitioners. With this setup, in this section we introduce the bounds of information flow on a tree with long branches.

Consider a reversible rate matrix Q with eigenvalues $0 > \lambda_1 \geq \cdots \geq \lambda_K$. For

any evolutionary time $t \geq 0$, the reversible Markov matrix e^{Qt} has real eigenvalues $1 > e^{\lambda_1 t} \geq \cdots \geq e^{\lambda_K t}$. Every edge e of the evolutionary tree has a branch length t_e, determining its associated transition matrix $P_e = e^{Qt_e}$.

The Pulley principle, introduced by Felsenstein [1981], states that the root of a reversible process is unidentifiable. Thus the placement of the root with reversible Q is arbitrary, implying that we can compare the expected norm of the memory vector $\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi]$ at each point of a tree assuming that this point were the root. Thus the expected memory vector norm gives a way to distinguish which point of a tree is the least identified.

However, the computation of $\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi]$ on a tree with m aligned sequences is unfeasible for large m, since it requires summing over $(K+1)^m$ patterns. Alternatively, one could estimate $\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi]$ using a large number of observed patterns ∂, although here we do not consider sampling strategies.

Instead, a way to compare the identification of points on a phylogenetic tree is the usage of general upper bounds depending only on the branch lengths of the tree. We introduce the following upper bound, which is only effective when the tree considered has long branches.

Theorem 4.4. *Consider the evolutionary process of Figure 4.2, where we assume* ***stationarity****, $d = 2$ and transition matrices $P_c = e^{Qt_c}$. We denote the equilibrium distribution of Q as $\boldsymbol{\pi} = (\pi_i)$ and define $H := \sqrt{\sum_i 1/\pi_i}$.*
Consider moreover the normalized likelihood vector $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$ at R and memory vectors $\boldsymbol{m_c} := P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}$ at nodes P_cA_c given their respective subpattern ∂_c. The following holds.

a) Set $\mathbb{E}[\|\boldsymbol{m_c}\|_\pi] := \sum_{\partial_c} \Pr(\partial_c)\|\boldsymbol{m_c}\|_\pi$. We have the bound

$$\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi] \leq \mathbb{E}[\|\boldsymbol{m_1}\|_\pi] + \mathbb{E}[\|\boldsymbol{m_2}\|_\pi] + H\mathbb{E}[\|\boldsymbol{m_1}\|_\pi]\mathbb{E}[\|\boldsymbol{m_2}\|_\pi].$$

b) Assuming reversibility, if we define $E(R) := \mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi]$ and $E(A_c) := \mathbb{E}[\|\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}\|_\pi] = \sum_{\partial_c} \Pr(\partial_c)\|\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}\|_\pi$, then

$$E(R) \leq e^{\lambda_1 t_1}E(A_1) + e^{\lambda_1 t_2}E(A_2) + e^{\lambda_1(t_1+t_2)}HE(A_1)E(A_2).$$

Proof.

a) Since we assume stationarity, we know that $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} := \boldsymbol{\rho}_{\boldsymbol{\partial}}/\Pr(\partial)$, while $\Pr(\partial) = D(\partial)\Pr_{IND}(\partial)$ as defined in Eq. 4.24. Therefore

$$D(\partial)\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi = \|\frac{\boldsymbol{\rho}_{\boldsymbol{\partial}}}{\Pr(\partial)}D(\partial) - \mathbf{1}D(\partial)\|_\pi = \\ = \|\frac{\boldsymbol{\rho}_{\boldsymbol{\partial}}}{\Pr_{IND}(\partial)} - \mathbf{1}D(\partial)\|_\pi. \tag{4.27}$$

Eq. 4.25 gives $\boldsymbol{\rho}_{\boldsymbol{\partial}}/\Pr_{IND}(\partial) = \mathbf{1} + \boldsymbol{m_1} + \boldsymbol{m_2} + \boldsymbol{m_1} \circ \boldsymbol{m_2}$, while Eq. 4.26 implies $D(\partial) = 1 + \boldsymbol{\pi} \cdot (\boldsymbol{m_1} \circ \boldsymbol{m_2})$. Using the triangle inequality, we obtain

$$D(\partial)\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi \leq \|\boldsymbol{m_1}\|_\pi + \|\boldsymbol{m_2}\|_\pi + \|\boldsymbol{m_1} \circ \boldsymbol{m_2} - \mathbf{1}(\boldsymbol{\pi} \cdot (\boldsymbol{m_1} \circ \boldsymbol{m_2}))\|_\pi \\ \leq \|\boldsymbol{m_1}\|_\pi + \|\boldsymbol{m_2}\|_\pi + \|\boldsymbol{m_1} \circ \boldsymbol{m_2}\|_\pi, \tag{4.28}$$

where we used the centralizing inequality of Prop. 4.2.

For any vector $\boldsymbol{x}$, we have $\|\boldsymbol{x}\|_\pi = \|\boldsymbol{x} \circ \boldsymbol{\pi}^{1/2}\|$, where the exponentiation occurs entrywise and $\|\cdot\|$ denotes the standard Euclidian norm. Moreover, Lemma 4.12, proved in Appendix 4.B, states that $\|\bigcirc_i \boldsymbol{x_i}\| \leq \prod_i \|\boldsymbol{x_i}\|$. Therefore

$$\|\boldsymbol{m_1} \circ \boldsymbol{m_2}\|_\pi = \|\boldsymbol{\pi}^{1/2} \circ \boldsymbol{m_1} \circ \boldsymbol{m_2}\| = \|\boldsymbol{\pi}^{-1/2} \circ (\boldsymbol{\pi}^{1/2} \circ \boldsymbol{m_1}) \circ (\boldsymbol{\pi}^{1/2} \circ \boldsymbol{m_2})\| \leq \\ \leq \|\boldsymbol{\pi}^{-1/2}\|\|\boldsymbol{\pi}^{1/2} \circ \boldsymbol{m_1}\|\|\boldsymbol{\pi}^{1/2} \circ \boldsymbol{m_2}\| = H\|\boldsymbol{m_1}\|_\pi\|\boldsymbol{m_2}\|_\pi. \tag{4.29}$$

Therefore Equations 4.28 and 4.29 give

$$D(\partial)\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi \leq \|\boldsymbol{m_1}\|_\pi + \|\boldsymbol{m_2}\|_\pi + H\|\boldsymbol{m_1}\|_\pi\|\boldsymbol{m_2}\|_\pi. \tag{4.30}$$

Taking the expected value of Eq. 4.28 under distribution $\Pr_{IND}(\partial) = \Pr(\partial_1)\Pr(\partial_2)$, where $\Pr_{IND}(\partial)D(\partial) = \Pr(\partial)$ due to stationarity, it follows that

$$\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi] \leq \mathbb{E}[\|\boldsymbol{m_1}\|_\pi] + \mathbb{E}[\|\boldsymbol{m_2}\|_\pi] + H\mathbb{E}[\|\boldsymbol{m_1}\|_\pi]\mathbb{E}[\|\boldsymbol{m_2}\|_\pi]. \tag{4.31}$$

b) Due to reversibility, Eq. 4.20 using the Markov matrix $P = e^{Qt_c}$ gives $\|\boldsymbol{m_c}\|_\pi \leq |\theta_1|\ \|\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}\|_\pi$, where $\theta_1 = e^{\lambda_1 t_c}$. This inequality, applied to item a), gives the result.

□

Branch lengths and rate matrices can be inversely rescaled, because $e^{Qt} = e^{(Q/C)(Ct)}$. For simplicity, we will choose the rescaling constant $C = |\lambda_1|$,

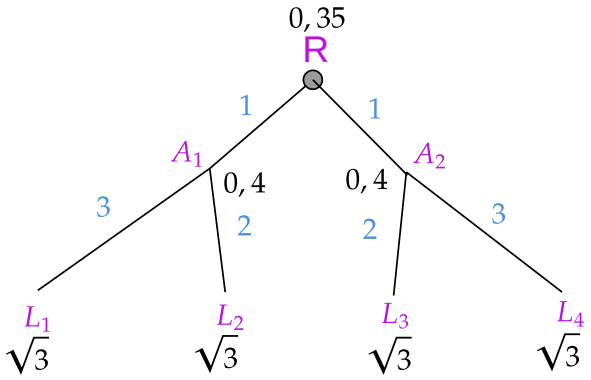

Figure 4.4: Diagram of the recursive computation of the upper bound of the expected memory vector norm at the root (grey circle) for a four-state alphabet and equilibrium distribution $\boldsymbol{\pi} = (1, 2, 3, 4)/10$, giving $H \approx 4, 56$. Blue numbers indicate branch length. Initially, the upper bound at the leaves is $\sqrt{3} \approx 1, 73$. Then we apply Eq. 4.32 recursively towards the root, giving e.g. $E(A_1) \leq \sqrt{3}e^{-2} + \sqrt{3}e^{-3} + 3e^{-5}H \approx 0.4$ and $E(R) \leq E(A_1)e^{-1} + E(A_2)e^{-1} + E(A_1)E(A_2)e^{-2} \leq 0.35$.

implying that Theorem 4.4.b reads

$$E(R) \leq e^{-t_1}E(A_1) + e^{-t_2}E(A_2) + e^{-(t_1+t_2)}HE(A_1)E(A_2). \tag{4.32}$$

Theorem 4.4 can be compared to the recursive formula of the Pruning algorithm of Eq. 4.6, namely $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = (e^{Qt_1}\boldsymbol{\alpha_1}) \circ (e^{Qt_2}\boldsymbol{\alpha_2})$. Analogously, we can recursively use Eq. 4.32 to bound the expected norm of the memory vector at the root of a tree. As shown in Figure 4.4, at the leaves we can use the bound $\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|] \leq \sqrt{K}$ [Prop. 4.3], and then recursively use Eq. 4.32 to bound the norm at the deeper parent nodes. By convention, no upper bound can be greater than $\sqrt{K}$, as implied by Prop. 4.3.

4.7 A bound of information flow on a d-ary tree

It is well known that, for $x \approx 0$, one can use the approximation $(1 + x)^n \approx 1 + nx$. In this sense, for numbers close to 1, multiplication can be approximated by addition. Similarly, to prove Theorem 4.6, we need to bound the entrywise product by a sum of vectors using the following lemma, proved in Appendix 4.B.

Lemma 4.5. *For any integer $d \geq 2$ and reals $S \in (0, 2)$ and $\epsilon > 0$, if $S \leq 4\epsilon/(1 + 2\epsilon)$, then*

$$(1 + \frac{S}{d})^d < 1 + (1 + \epsilon)S. \tag{4.33}$$

If memory vectors $\boldsymbol{m_c} = P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}$ approach $\mathbf{0}$, then $P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} \approx \mathbf{1}$ and the entrywise product $\bigcirc_{c\in[d]} P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}}$ can be approximated by $\mathbf{1}+\sum_{c\in[d]} \boldsymbol{m_c}$. A more subtle argument yields a linear bound of a memory vector given pattern ∂ using subpatterns ∂_c, as explained in Theorem 4.6. We use vectors $P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}}$ just to make the connection to other results clear, and we could write any normalized vectors $\tilde{\boldsymbol{\beta}}_{\boldsymbol{c}}$ instead.

Theorem 4.6 (Hadamard-product upper bounds of the memory vector)**.** *In the evolutionary process of Figure 4.2, assume the stationary prior $\boldsymbol{\mu} = \boldsymbol{\pi}$, as also that all matrices P_c have the same equilibrium distribution $\boldsymbol{\pi}$.*
Consider the normalized likelihood vector $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$ at R and the normalized likelihood vectors $P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} = \mathbf{1} + \boldsymbol{m_c}$ at nodes P_cA_c. Then, under the assumption that, for some $\epsilon > 0$, $\sum_{c\in[d]} \|\boldsymbol{m_c}\|_\infty \leq 4\epsilon/(1+2\epsilon)$ for all patterns ∂, it holds that

$$\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi] < (1+\epsilon)\sum_{c\in[d]} \mathbb{E}[\|\boldsymbol{m_c}\|_\pi]. \tag{4.34}$$

Proof. First of all, since in Eq. 4.24 we stated that $\Pr_\pi(\partial) = D(\partial)\Pr_{IND}(\partial)$, it holds that

$$\begin{aligned} D(\partial)\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi &= \|\frac{\boldsymbol{\rho_\partial}}{\Pr_\pi(\partial)}D(\partial) - \mathbf{1}D(\partial)\|_\pi = \\ &= \|\frac{\boldsymbol{\rho_\partial}}{\Pr_{IND}(\partial)} - \mathbf{1}D(\partial)\|_\pi \end{aligned} \tag{4.35}$$

Now we can use Equations 4.25 and 4.26, where we expanded $\boldsymbol{\rho_\partial}/\Pr_{IND}(\partial)$ and $D(\partial)$, giving

$$\begin{aligned} D(\partial)\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi &= \|\mathbf{1} + \sum_{p\in[d]}\sum_{C\in\binom{[d]}{p}} \bigcirc_{c\in C}\boldsymbol{m_c} - \mathbf{1}\Big(1 + \sum_{p\in[d]}\sum_{C\in\binom{[d]}{p}} \boldsymbol{\pi}\cdot\bigcirc_{c\in C}\boldsymbol{m_c}\Big)\|_\pi = \\ &= \|\sum_{p\in[d]}\sum_{C\in\binom{[d]}{p}} \bigcirc_{c\in C}\boldsymbol{m_c} - \mathbf{1}\Big(\boldsymbol{\pi}\cdot\sum_{p\in[d]}\sum_{C\in\binom{[d]}{p}} \bigcirc_{c\in C}\boldsymbol{m_c}\Big)\|_\pi \\ &= \|\sum_{p\in[d]}\sum_{C\in\binom{[d]}{p}} \bigcirc_{c\in C}\boldsymbol{m_c} - \mathbf{1}\langle\mathbf{1}, \sum_{p\in[d]}\sum_{C\in\binom{[d]}{p}} \bigcirc_{c\in C}\boldsymbol{m_c}\rangle_\pi\|_\pi. \end{aligned} \tag{4.36}$$

Using the centralizing inequality of Prop. 4.2, we obtain

$$D(\partial)\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi \leq \|\sum_{p\in[d]}\sum_{C\in\binom{[d]}{p}} \bigcirc_{c\in C}\boldsymbol{m_c}\|_\pi = \tag{4.37}$$

$$= \|\bigcirc_{c\in[d]}(\mathbf{1}+\boldsymbol{m_c}) - \mathbf{1}\|_\pi = \tag{4.38}$$

$$= \|\ |\bigcirc_{c\in[d]}(\mathbf{1}+\boldsymbol{m_c}) - \mathbf{1}|\ \|_\pi, \tag{4.39}$$

where $|\cdot|$ denotes the entrywise absolute value. We have the entrywise inequality

$$|\bigcirc_{c\in[d]}(\mathbf{1}+\boldsymbol{m_c})-\mathbf{1}|\leq\bigcirc_{c\in[d]}(\mathbf{1}+|\boldsymbol{m_c}|)-\mathbf{1}, \tag{4.40}$$

and the Arithmetic-Geometric Mean Inequality applied to $\bigcirc_{c\in[d]}(\mathbf{1}+|\boldsymbol{m_c}|)$ implies that, in an entrywise manner,

$$\mathbf{0}\leq\bigcirc_{c\in[d]}(\mathbf{1}+|\boldsymbol{m_c}|)-\mathbf{1}\leq\Big(\mathbf{1}+\frac{\sum_{c\in[d]}|\boldsymbol{m}^{\boldsymbol{c}}_{\boldsymbol{\partial}}|}{d}\Big)^d-\mathbf{1},$$

where the exponentiation occurs entrywise. Consequently

$$\|\bigcirc_{c\in[d]}(\mathbf{1}+\boldsymbol{m_c})-\mathbf{1}\|_\pi\leq\|\Big(\mathbf{1}+\frac{\sum_{c\in[d]}|\boldsymbol{m}^{\boldsymbol{c}}_{\boldsymbol{\partial}}|}{d}\Big)^d-\mathbf{1}\|_\pi. \tag{4.41}$$

We write $|\boldsymbol{m_c}|=(|m_c^0|,\cdots,|m_c^K|)$. Lemma 4.5 with $S=\sum_{c\in[d]}|m_c^i|$ for $i\in\mathbb{A}$ further implies that, if we have $\sum_{c\in[d]}\|\boldsymbol{m_c}\|_\infty\leq 4\epsilon/(1+2\epsilon)$, then the following entrywise inequality holds

$$\Big(\mathbf{1}+\frac{\sum_{c\in[d]}|\boldsymbol{m}^{\boldsymbol{c}}_{\boldsymbol{\partial}}|}{d}\Big)^d-\mathbf{1}<(1+\epsilon)\sum_{c\in[d]}|\boldsymbol{m}^{\boldsymbol{c}}_{\boldsymbol{\partial}}|. \tag{4.42}$$

Therefore, assuming that $\sum_{c\in[d]}\|\boldsymbol{m_c}\|_\infty\leq 4\epsilon/(1+2\epsilon)$ for all ∂, and combining Equations 4.39, 4.41 and 4.42, it holds that

$$\begin{aligned}D(\partial)\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}-\mathbf{1}\|_\pi<(1+\epsilon)\|\sum_{c\in[d]}|\boldsymbol{m}^{\boldsymbol{c}}_{\boldsymbol{\partial}}|\|_\pi\leq\\ \leq(1+\epsilon)\sum_{c\in[d]}\|\ |\boldsymbol{m_c}|\ \|_\pi=(1+\epsilon)\sum_{c\in[d]}\|\boldsymbol{m_c}\|_\pi,\end{aligned} \tag{4.43}$$

where we applied the triangle inequality. Recall that $\Pr(\partial)=D(\partial)\Pr_{IND}(\partial)$ due to stationarity. Taking the expected value of both sides using distribution $\Pr_{IND}(\partial)$, under which random variables $\boldsymbol{m_c}$ are independent, we get

$$\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}-\mathbf{1}\|_\pi]<(1+\epsilon)\sum_{c\in[d]}\mathbb{E}[\|\boldsymbol{m_c}\|_\pi], \tag{4.44}$$

as desired. □

Interestingly, using Eq. 4.43 we can upper bound the module $\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}-\mathbf{1}\|_\pi$ by bounding the dependence factor $D(\partial)$ close enough to 1. Such a bound is stated in Proposition 4.13 [Appendix 4.B]. This means that, along the proof of Theorem 4.6,

we implicitly bounded the module $\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi$ for a given pattern ∂.

Theorem 4.6 does not consider the child nodes A_c, but only the parent nodes $P_c A_c$. Moreover, we did not mention how to make the memory vectors at $P_c A_c$ small enough to satisfy the assumption of Theorem 4.6. These particularities are considered in Theorem 4.7, which is our core result bounding the information flow on trees from children to parent node. Intuitively, if many mutations have occurred, one expects that information flows somehow decreasingly from the leaves towards the root. Theorem 4.7 formalizes this intuition, sub-additively bounding the memory vector norm at the root R by the sum of norms at its child nodes A_c.

Theorem 4.7 (Root-Children upper bounds of the memory vector)**.** *In the evolutionary process of Figure 4.2, assume the* ***stationary prior*** $\boldsymbol{\mu} = \boldsymbol{\pi}$*, as also that all matrices P_c have the same equilibrium distribution $\boldsymbol{\pi}$.*
Consider moreover the normalized likelihood vector $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$ at R and the normalized likelihood vectors $\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}}$ at child nodes A_c. Moreover assume that, for some constants $C_c > 0$ and for any normalized likelihood vector $\tilde{\boldsymbol{\alpha}}$, we have the bound

$$\|P_c \tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi \leq C_c \, \|\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi.$$

In these conditions, if

$$\sum_{c \in [d]} C_c \leq \frac{\min_i \pi_i}{\sqrt{1 - \min_i \pi_i}} \frac{4\epsilon}{1 + 2\epsilon}$$

for some $\epsilon > 0$, then it holds that

$$\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi] < (1 + \epsilon) \sum_{c \in [d]} C_c \mathbb{E}[\|\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}\|_\pi].$$

Proof. To apply Theorem 4.6, we need the fulfillment for all patterns $\partial = (\partial_1, \cdots, \partial_d)$ of inequality

$$\sum_{c \in [d]} \|P_c \tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}\|_\infty \leq 4\epsilon/(1 + 2\epsilon).$$

Since $\|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_\pi / \min_i \sqrt{\pi_i}$ as stated in Eq. 4.10, it holds that

$$\sum_{c \in [d]} \|P_c \tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}\|_\infty \leq \frac{1}{\min_i \sqrt{\pi_i}} \sum_{c \in [d]} \|P_c \tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}\|_\pi. \tag{4.45}$$

We bounded the $L_2(\pi)$-norm of any memory vector in Eq. 4.15, implying

$$\sum_{c\in[d]} \|P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}\|_\pi \leq \sum_{c\in[d]} C_c \|\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} - \mathbf{1}\|_\pi \leq \tag{4.46}$$

$$\leq \sqrt{\frac{1}{\min_i \pi_i} - 1} \sum_{c\in[d]} C_c. \tag{4.47}$$

Consequently, to apply Theorem 4.6 it is enough that

$$\frac{1}{\min_i \sqrt{\pi_i}} \sqrt{\frac{1}{\min_i \pi_i} - 1} \sum_{c\in[d]} C_c \leq \frac{4\epsilon}{1+2\epsilon},$$

or equivalently

$$\sum_{c\in[d]} C_c \leq \frac{\min_i \pi_i}{\sqrt{1-\min_i \pi_i}} \frac{4\epsilon}{1+2\epsilon}. \tag{4.48}$$

Finally, Equation 4.46, combined with Theorem 4.6, yields the desired result.

□

4.8 Definition of unsolvability

The reconstruction problem consists in estimating the ancestral state at the root R given a pattern ∂ [Mossel, 2001b]. The tree, the prior $\boldsymbol{\mu}$ and transition matrix P are known. Notably, the ML estimate has the highest probability of a correct reconstruction among all reconstruction methods (Theorem 17.2 of Guiasu [1977]). After computing the posterior state distribution $\boldsymbol{r}_{\boldsymbol{\partial}}$ at the root given ∂, the ML estimate is the state with the maximum posterior probability, also called the Maximum A Posteriori (MAP) estimate.

If we assume a prior state distribution $\boldsymbol{\mu} = (\mu_0, \cdots, \mu_K)$ and the likelihood vector at the root R is $\boldsymbol{\rho}_{\boldsymbol{\partial}}$, then Bayes' theorem implies that the posterior state distribution is

$$\boldsymbol{r}_{\boldsymbol{\partial}} = \frac{\boldsymbol{\rho}_{\boldsymbol{\partial}} \circ \boldsymbol{\mu}}{\Pr(\partial)} = \frac{\boldsymbol{\rho}_{\boldsymbol{\partial}} \circ \boldsymbol{\mu}}{\boldsymbol{\rho}_{\boldsymbol{\partial}} \cdot \boldsymbol{\mu}}. \tag{4.49}$$

If we write $\boldsymbol{r}_{\boldsymbol{\partial}} = (r_0, \cdots, r_K)$, then the MAP estimate is a state $i \in \mathbb{A}$ such that $r_i = \max_i r_i$, and the MAP probability of a correct reconstruction is $\max_i r_i$.

In a evolutionary process on a d-ary tree, we want to describe when the root state cannot be confidently reconstructed, assuming that the tree has a large number of levels g. In general, the transition matrices of the evolutionary process can be

different, although in Section 4.9 we focus on the case when all transition matrices are equal to P. We introduce the following definition.

Definition 4.3. *When $\boldsymbol{r}_{\boldsymbol{\partial}} \neq \boldsymbol{\mu}$, we say that pattern ∂ is informative, because the observation of ∂ influences the posterior. In contrast, we say that pattern ∂ is uninformative when $\boldsymbol{r}_{\boldsymbol{\partial}} = \boldsymbol{\mu}$.*

Since $\boldsymbol{r}_{\boldsymbol{\partial}} = \boldsymbol{\rho}_{\boldsymbol{\partial}} \circ \boldsymbol{\mu}/(\boldsymbol{\rho}_{\boldsymbol{\partial}} \cdot \boldsymbol{\mu})$, pattern ∂ in uninformative iff $\boldsymbol{\rho}_{\boldsymbol{\partial}}$ is uniform, which occurs iff $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = \mathbf{1}$. Therefore the informal condition $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1} \approx \mathbf{0}$ is a compact way to summarize the lack of information provided by a pattern ∂.

Instead of focusing on a single pattern, we will use an average of the norms of $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}$ over all possible patterns Δ_g at the leaves of a g-level tree, which grows with g. Explicitly, we consider the expected value

$$\mathbb{E}[\, \|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi] = \sum_{\partial \in \Delta_g} \Pr(\partial) \|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi, \tag{4.50}$$

which has a growing number of summands as $g \to \infty$. Recall that finite-dimensional norms are equivalent (see Steven G. Johnson [2020]), meaning that they differ at most by a multiplicative constant. Thus we can state the following definition.

Definition 4.4. *We say that the reconstruction problem is unsolvable when, for some norm $\|\cdot\|_*$,*

$$\mathbb{E}[\, \|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_*] \to 0 \text{ as } g \to \infty.$$

Differently explained, the reconstruction problem is unsolvable when patterns are expectedly uninformative as g grows.

This definition of unsolvability is equivalent to the definition studied by Mossel [2001a], as stated in Appendix 4.C, where we also prove that the prior $\boldsymbol{\mu}$ does not influence unsolvability as long as $\boldsymbol{\mu} > 0$. Thus we will assume $\boldsymbol{\mu} = \boldsymbol{\pi} > 0$, that is, a stationary process.

The unsolvability of the reconstruction problem has been studied repeatedly [Mossel, 2001b]. As proven by Mossel and Peres [2003], assuming any 2×2 transition matrix over a binary alphabet, the reconstruction problem is unsolvable if $d|\theta_1| < 1$. Similarly, assuming a Jukes-Cantor (JC) transition matrix over any alphabet, the reconstruction problem is unsolvable if $d|\theta_1| < 1$ [Mossel and Peres, 2003]. Recall that a JC transition matrix has all its off-diagonal entries identical. Solvable examples when $d|\theta_1| > 1$ are described by Mossel [2001a].

4.9 Bounds of unsolvability

In this section, we state some bounds of the unsolvability of the reconstruction problem. The bound of Prop. 4.8 deals with a binary tree and a reversible transition matrix, while the bound of Prop. 4.9 applies to any d-ary tree and nearly any transition matrix.

Proposition 4.8. *Consider an irreducible, aperiodic and* ***reversible*** *Markov matrix P over $K+1$ states with absolutely largest non-unitary eigenvalue θ_1. Given the equilibrium distribution $\boldsymbol{\pi}$ of P, set $H := \sqrt{\sum_i 1/\pi_i}$.*
Then, on a binary tree, the reconstruction problem using transition matrix P is unsolvable if

$$|\theta_1| < \frac{-1+\sqrt{1+H\sqrt{K}}}{H\sqrt{K}}.$$

Proof. We can assume a stationary prior due to Prop. 4.14 We can use Theorem 4.4.a with a reversible transition matrix P for both child nodes of R. The bound of Eq. 4.20 gives $\|P\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}}-\mathbf{1}\|_\pi \le |\theta_1|\|\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}}-\mathbf{1}\|_\pi$. Define $E(A) := \max_{c\in\{1,2\}} \mathbb{E}[\|\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}}-\mathbf{1}\|_\pi]$. Thus Theorem 4.4.a gives

$$E(R) \le 2|\theta_1|E(A) + |\theta_1|^2 H E(A)^2. \tag{4.51}$$

Applying Eq. 4.51 recursively, a sufficient condition for unsolvability is that the upper bound of the expected memory vector norm monotonously decreases towards zero as we go deeper in the tree towards the root. Define function $f(x) = 2x|\theta_1| + H|\theta_1|^2x^2$ and note that Eq. 4.51 can be stated as $E(R) \le f(E(A))$. Function $f(x)$ has only two fixed points satisfying $f(x) = x$, namely $x_1 = 0$ and

$$x_2 = \frac{1-2|\theta_1|}{|\theta_1|^2 H}. \tag{4.52}$$

We assume $|\theta_1| < 1/2$, implying that $x_2 > 0$ and $f(x) < x$ when $x \in (x_1, x_2)$. WLOG we can assume $E(A) > 0$, and thus inequality $f(E(A)) < E(A)$ holds if

$$E(A) < x_2. \tag{4.53}$$

Since $E(A) \le \sqrt{K}$ from Prop. 4.3, a sufficient condition is $\sqrt{K} < x_2$. Substituting x_2 using Eq. 4.52 and solving for $|\theta_1|$, under the assumption $|\theta_1| < 1/2$, a sufficient condition for unsolvability is therefore

$$|\theta_1| < \frac{-1+\sqrt{1+H\sqrt{K}}}{H\sqrt{K}}. \tag{4.54}$$

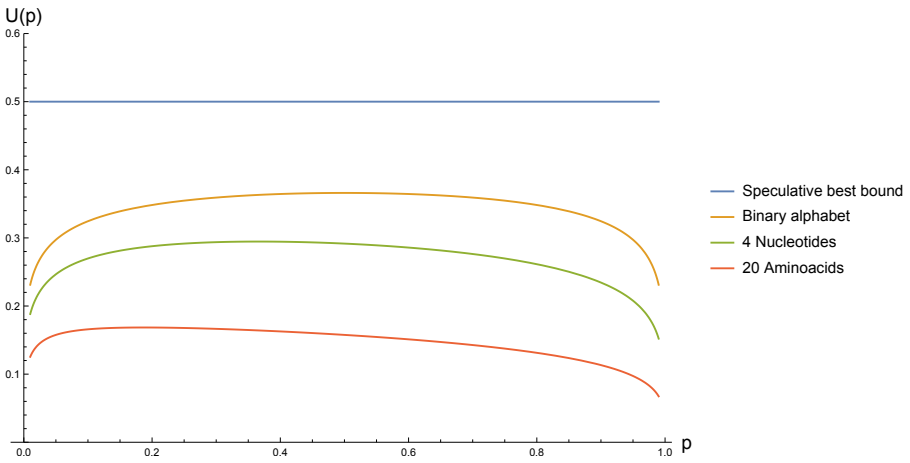

Figure 4.5: Plots of the unsolvability bound $U(p)$ of Eq. 4.57 as a function of the first entry $p \in [0.01, 0.99]$ of the equilibrium distribution $\boldsymbol{\pi}(p)$ defined in Eq. 4.56. For any reversible transition matrix with equilibrium distribution $\boldsymbol{\pi}(p)$, if $|\theta_1|<$ $U(p)$, then the reconstruction problem is unsolvable. For an arbitrary alphabet, the speculative best bound of $1/2$ has been proven only for a JC transition matrix [Mossel, 2001b].

It remains to show that Eq. 4.54 implies that $|\theta_1| < 1/2$. Writing $H\sqrt{K} = \psi > 0$, this implication is equivalent to

$$\sqrt{1+\psi} < 1 + \frac{\psi}{2}, \tag{4.55}$$

which squaring both sides yields the obvious $0 < \psi^2/4$, as desired.

□

Our bound is not sharp, although it is more general than the bound of Mossel [2001b]. To visualize this, consider the stationary distribution

$$\boldsymbol{\pi}(p) = (p, \frac{1-p}{K}, \cdots, \frac{1-p}{K}). \tag{4.56}$$

Then $H(p) = \sqrt{1/p + K/(1-p)}$ and the upper bound of Prop. 4.8 is

$$U(p) = \frac{-1 + \sqrt{1 + H(p)\sqrt{K}}}{H(p)\sqrt{K}}, \tag{4.57}$$

represented in Figure 4.5 for $p \in [0.01, 0.99]$ respectively assuming a binary, nucleotide and amino acid alphabet. This result can be compared with the speculative best possible bound on a binary tree, which is $1/2$ assuming a JC transition matrix [Mossel, 2001b].

Now we will use Theorem 4.7 to obtain an upper bound for unsolvability on a

d-ary tree that admits nearly any transition matrix.

Proposition 4.9. *Given any irreducible and aperiodic Markov matrix P, assume that, for some $C > 0$ and for any normalized likelihood vector $\tilde{\boldsymbol{\alpha}}$, we have the bound*

$$\|P\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi \leq C \, \|\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi.$$

Then, on a d-ary tree, the reconstruction problem using transition matrix P is unsolvable if

$$Cd < \min\left\{\frac{1}{3}, \frac{8}{5}\frac{\min_i \pi_i}{\sqrt{1 - \min_i \pi_i}}\right\}.$$

Proof. Recall that we can assume a stationary prior due to Prop. 4.14. We want to apply Theorem 4.7 with all constants C_c equal to C. WLOG we can assume that the d-arity of the tree is complete (every children makes exactly d-children), because the conditions of Theorem 4.7 apply also for a parent node with less than d children.

Consider the normalized likelihood vectors at the leaves, from now on $\tilde{\boldsymbol{l}}_{\boldsymbol{c}}$ where $c \in \{1, \cdots, d^g\}$. The recursive application of Theorem 4.7 yields that

$$\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - 1\|_\pi] < (1+\epsilon)^g C^g \sum_{c \in [d^g]} \mathbb{E}[\|\tilde{\boldsymbol{l}}_{\boldsymbol{c}} - \mathbf{1}\|_\pi], \tag{4.58}$$

while the upper bound of the norm of any memory vector (Eq. 4.15) gives

$$\begin{aligned}\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - 1\|_\pi] &< (1+\epsilon)^g C^g \sum_{c \in [d^g]} \sqrt{\frac{1}{\min_i \pi_i} - 1} = \\ &= (1+\epsilon)^g C^g d^g \sqrt{\frac{1}{\min_i \pi_i} - 1}.\end{aligned} \tag{4.59}$$

Recall that the reconstruction problem is unsolvable iff $\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - 1\|_\pi]$ tends to 0 as g grows. Thus the reconstruction problem is unsolvable if the condition of Theorem 4.7 holds and moreover

$$(1+\epsilon)Cd < 1, \tag{4.60}$$

as implied by the bound of Eq. 4.59. To make the following bounds hold simultaneously,

$$\begin{aligned}Cd &< \frac{1}{1+\epsilon} \\ Cd &\leq \frac{\min_i \pi_i}{\sqrt{1 - \min_i \pi_i}} \frac{4\epsilon}{1+2\epsilon},\end{aligned} \tag{4.61}$$

the optimal ϵ is the one making the two bounds coincide, although this intersection depends on $\min_i \pi_i$. For the sake of clarity, set $\epsilon = 2$, yielding the result.

□

Prop. 4.9 assumes that any memory vector decreases at least by a factor of C under the action of P. We have inferred such bounds in Equations 4.20 and 4.65 (Appendix 4.A), giving the following corollaries.

Corollary 4.10. *Consider any irreducible, aperiodic and* ***reversible*** *Markov matrix P with non-unitary absolutely largest eigenvalue θ_1. Then, on a d-ary tree, the reconstruction problem using transition matrix P is unsolvable if*

$$|\theta_1|d < \min\left\{\frac{1}{3}, \frac{8}{5}\frac{\min_i \pi_i}{\sqrt{1-\min_i \pi_i}}\right\}.$$

Corollary 4.11. *Consider any irreducible and aperiodic Markov matrix P with largest singular value σ_1. Then, on a d-ary tree, the reconstruction problem using transition matrix P is unsolvable if*

$$\sigma_1 d < \frac{\min_i \sqrt{\pi_i}}{\max_i \sqrt{\pi_i}} \min\left\{\frac{1}{3}, \frac{8}{5}\frac{\min_i \pi_i}{\sqrt{1-\min_i \pi_i}}\right\}.$$

The bounds of Corollaries 4.10 and 4.11 are weak compared to the speculative best possible bound $|\theta_1|d < 1$, although this speculative bound has been proven only for JC transition matrices [Mossel, 2001b].

In this section we assumed the same transition matrix P for all edges, but Theorem 4.7 is very versatile and can be applied in multiple scenarios. We can, for example, allow a different transition matrix for every child node as long as $\boldsymbol{\pi}$ is the equilibrium distribution. This is specially useful for continuous Markov models of evolution, where every transition matrix e^{Qt_e} has the same equilibrium distribution $\boldsymbol{\pi}$ as the rate matrix Q.

4.10 Conclusion

As we have shown, the $L_2(\pi)$-norm of the memory vector is a natural formalization of the concept of "amount of identification". Importantly, the $L_2(\pi)$-norm of the memory vector has an intuitive behaviour, growing with the number of children available and decreasing exponentially with the evolutionary time.

Future theoretical work could focus on making sharper our bounds of identification flow. With our approach, we compute bounds for every pattern (as Eq. 4.43)

and then take the expected value. Directly bounding the expected value could lead to some improvements. Another alternative strategy could be to avoid the usage of different norms, which debilitates e.g. Theorem 4.7 (concretely in Eq. 4.45).

From a practical perspective, the $L_2(\pi)$-norm of the memory vector can be used to distinguish poorly identified ancestors in a phylogeny. However, the memory vector is specially interesting for stationary and reversible processes, whose root is unidentifiable due to the Pulley principle [Felsenstein, 1981]. Any possible root R determines an expected norm $\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi]$, which can be estimated using a large number of observed patterns. Therefore, given an alignment of DNA or amino acid sequences, we can distinguish the least identified root of the reconstructed tree. This distinction among roots can be valuable to assess other rooting methods. Alternatively, to benefit from the linear properties of the inner product, also the expected squared norm $\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi^2]$ could be employed.

Appendix 4.A A general bound of the norm growth

The bounds of norm growth as the one of Eq. 4.20 are useful for our proofs. Without the assumption that matrix P be reversible, a general bound similar to Eq. 4.20 as

$$\|P\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi \leq C \, \|\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi \tag{4.62}$$

for some constant $C > 0$ is more difficult to obtain, since matrix P may have complex eigenvalues or not diagonalize. A weak bound using well known properties of the singular value decomposition (explained in Horn and Johnson [1991]) can be obtained as follows. Fist of all recall that, at a node PR, the normalized likelihood vector is $P\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$. Therefore the memory vector at node PR is $P\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}$. Moreover, since $P\mathbf{1} = \mathbf{1}$, it holds that

$$P\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1} = P(\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}),$$

that is, matrix P is an endomorphism of memory vectors. Now let us define $\Pi := \mathbf{1}\boldsymbol{\pi}^T$ and $\Psi := \mathrm{Diag}(\boldsymbol{\pi})$. Since $P\mathbf{1} = \Pi\mathbf{1} = \Pi\tilde{\boldsymbol{\alpha}} = \mathbf{1}$ for any normalized likelihood vector $\tilde{\boldsymbol{\alpha}}$, it holds that

$$P\tilde{\boldsymbol{\alpha}} - \mathbf{1} = (P - \Pi)(\tilde{\boldsymbol{\alpha}} - \mathbf{1}). \tag{4.63}$$

Notably, matrix $P - \Pi$ has eigenvalues $0, \theta_1, \cdots, \theta_K$, although we only need its singular values. Consider the largest singular value σ_1 of $P - \Pi$, which satisfies $\sigma_1 \geq |\theta_1|$ and thus may be larger than 1 (see Horn and Johnson [1991]). Let $\|\cdot\|$

denote the matrix norm induced by the Euclidean norm. It holds that

$$\begin{aligned}\|P\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi &= \|(P-\Pi)(\tilde{\boldsymbol{\alpha}} - \mathbf{1})\|_\pi = \\ &= \|\sqrt{\Psi}(P-\Pi)(\tilde{\boldsymbol{\alpha}} - \mathbf{1})\| \leq \|\sqrt{\Psi}\|\,\|P-\Pi\|\,\|\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|,\end{aligned} \tag{4.64}$$

where the inequality follows because induced norms are sub-multiplicative. Since the Euclidean matrix norm coincides with the spectral norm, $\|\sqrt{\Psi}\| = \max_i \sqrt{\pi_i}$ and $\|P - \Pi\| = \sigma_1$. This together with $\|\boldsymbol{x}\| \leq \|\boldsymbol{x}\|_\pi / \min_i \sqrt{\pi_i}$ gives

$$\|P\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi \leq \frac{\max_i \sqrt{\pi_i}}{\min_i \sqrt{\pi_i}} \sigma_1 \|\tilde{\boldsymbol{\alpha}} - \mathbf{1}\|_\pi, \tag{4.65}$$

and thus the bound of Eq. 4.62 holds for $C = \sigma_1 \max_i \sqrt{\pi_i} / \min_i \sqrt{\pi_i}$. This bound is specially weak when $\min_i \sqrt{\pi_i} << \max_i \sqrt{\pi_i}$.

Appendix 4.B Technical results

Proposition (Proof of Prop. 4.3)**.** *Consider a stationary process on a tree with m leaves. If the alphabet $\mathbb{A}$ has $K+1$ states, then it holds that*

$$\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi] := \sum_{\partial \in \mathbb{A}^m} \Pr(\partial) \|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi \leq \sqrt{K}.$$

Proof. Since $\mathrm{Var}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi] \geq 0$, we have

$$\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi]^2 \leq \mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi^2]. \tag{4.66}$$

Using Eq. 4.14, we know that $\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi^2 \leq \max_k \tilde{\rho}_\partial^k - 1$. If we define $\boldsymbol{\rho}_{\boldsymbol{\partial}} := (\rho_\partial^k)$, substituting $\Pr(\partial) = \boldsymbol{\pi} \cdot \boldsymbol{\rho}_{\boldsymbol{\partial}}$ due to stationarity, we can do

$$\begin{aligned}\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi^2] &\leq -1 + \sum_\partial \Pr(\partial) \max_k \tilde{\rho}_\partial^k = \\ &= -1 + \sum_\partial \max_k \rho_\partial^k \leq -1 + \sum_\partial \sum_k \rho_\partial^k \leq K,\end{aligned} \tag{4.67}$$

where we used the fact that $\sum_\partial \boldsymbol{\rho}_{\boldsymbol{\partial}} = \mathbf{1}$. All in all, combining Equations 4.66 and 4.67 we infer

$$\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi] \leq \sqrt{K}, \tag{4.68}$$

as desired. Note that this bound is sharper than the upper bound $\sqrt{1/\min_i \pi_i - 1}$ obtainable from Eq. 4.14, since $1/\min_i \pi_i \geq K+1$.

□

Lemma (Proof of Lemma 4.5)**.** *For any integer $d \geq 2$ and reals $S \in (0,2)$ and $\epsilon > 0$, if $S \leq 4\epsilon/(1+2\epsilon)$, then*

$$(1+\frac{S}{d})^d < 1 + (1+\epsilon)S. \tag{4.69}$$

Proof. Expanding $(1+\frac{S}{d})^d$, Eq. 4.69 is equivalent to

$$\sum_{c\in[2,d]} \binom{d}{c} \frac{S^c}{d^c} < \epsilon S. \tag{4.70}$$

Since $0 < S < 2$, we have the bound

$$\sum_{c\in[2,d]} \binom{d}{c} \frac{S^c}{d^c} < \sum_{c\in[2,d]} \frac{S^c}{c!} < \sum_{c\in[2,\infty]} \frac{S^c}{2^c} = \frac{S^2}{2(2-S)}. \tag{4.71}$$

It is easy to see that, for $S \in (0,2)$, inequality

$$\frac{S^2}{2(2-S)} \leq \epsilon S \tag{4.72}$$

holds iff

$$S \leq \frac{4\epsilon}{1+2\epsilon}, \tag{4.73}$$

and the result follows.

□

Lemma 4.12. *Given vectors $\boldsymbol{x_1}, \cdots, \boldsymbol{x_d}$, let $\|\cdot\|$ denote the standard Euclidian norm. It holds that*

$$\|\bigcirc_{c\in[d]} \boldsymbol{x_c}\| \leq \prod_{c\in[d]} \|\boldsymbol{x_c}\|, \tag{4.74}$$

with equality for $d \geq 2$ iff there exists a vector $\boldsymbol{e}$ with a single nonzero entry such that $\boldsymbol{x_c} \propto \boldsymbol{e}$ for all $c \in [d]$.

Proof. The case $d = 1$ is trivial, and the case $d = 2$ is

$$\|\boldsymbol{x_1} \circ \boldsymbol{x_2}\| \leq \|\boldsymbol{x_1}\|\|\boldsymbol{x_2}\|, \tag{4.75}$$

which assuming $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ have respectively entries x_1^k and x_2^k is equivalent to

$$\sum_k (x_1^k)^2 (x_2^k)^2 \leq (\sum_k (x_1^k)^2)(\sum_k (x_2^k)^2). \tag{4.76}$$

Last inequality is clearly true since the summands of the LHS are a subset of the summands of the RHS. To apply induction, we assume that the inequality is true

for some $d \geq 2$ and prove it for $d+1$. It holds that

$$\| \bigcirc_{c\in[d+1]} \boldsymbol{x_c}\| = \|(\boldsymbol{x_1} \circ \boldsymbol{x_2}) \circ \bigcirc_{c\in[3,d+1]} \boldsymbol{x_c}\| \leq \|\boldsymbol{x_1} \circ \boldsymbol{x_2}\| \prod_{c\in[3,d+1]} \|\boldsymbol{x_c}\|, \tag{4.77}$$

where we used the induction hypothesis with the d vectors $\boldsymbol{x_1} \circ \boldsymbol{x_2}$ and $\boldsymbol{x_c}$ for $c \in [3, d+1]$. using Eq. 4.75, the inequality follows.

Note that, for the equality in Eq. 4.75 to occur, it is necessary that vectors $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ have only one and the same nonzero entry. Using a symmetric argument, for the equality in Eq. 4.74 to occur, it is necessary that $\boldsymbol{x_1}, \cdots, \boldsymbol{x_d}$ have only one and the same nonzero entry. It is easy to see that this is also a sufficient condition. □

Proposition 4.13 (Mixing of the dependence factor)**.** *In the evolutionary process of Figure 4.2, assume that all matrices P_c have the same equilibrium distribution $\boldsymbol{\pi}$. Consider moreover the normalized likelihood vector $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$ at R and the normalized likelihood vectors $P_c\tilde{\boldsymbol{\alpha}}_{\boldsymbol{c}} = \mathbf{1} + \boldsymbol{m_c}$ at nodes P_cA_c. Then, under the assumption that $\sum_{c\in[d]} \|\boldsymbol{m_c}\|_\infty \leq 4\epsilon/(1+2\epsilon)$, the dependence factor satisfies*

$$|D(\partial) - 1| < \epsilon \sum_{c\in[d]} \|\boldsymbol{m_c}\|_\infty \leq \frac{4\epsilon^2}{1+2\epsilon}.$$

Proof. Let us bound the dependence factor $D(\partial)$ using the expansion of Eq. 4.26. First, we denote by $|\boldsymbol{m_c}|$ the vector obtained by taking the absolute value of each entry of $\boldsymbol{m_c}$, and thus write $|\boldsymbol{m_c}| = (|m_c^0|, \cdots, |m_c^K|)$. Using the fact that a weighted sum is smaller than its biggest term, plus the Arithmetic-Geometric Mean inequality, we obtain

$$\begin{aligned}
|D(\partial) - 1| &= |\boldsymbol{\pi} \cdot \sum_{p\in[2,d]} \sum_{C\in\binom{[d]}{p}} \bigcirc_{c\in C} \boldsymbol{m_c}| \\
&\leq \max_i \sum_{p\in[2,d]} \sum_{C\in\binom{[d]}{p}} \prod_{c\in C} |m_c^i| = \\
&= \max_i \prod_{i\in[c]} (1 + |m_c^i|) - \sum_{c\in[d]} |m_c^i| - 1 \leq \\
&\leq \max_i \ (1 + \frac{\sum_{c\in[d]} |m_c^i|}{d})^d - \sum_{c\in[d]} |m_c^i| - 1.
\end{aligned} \tag{4.78}$$

Using Lemma 4.5, if $\sum_{c\in[d]} |m_c^i| < 4\epsilon/(1+2\epsilon)$ for all i, then the expression of Eq.

4.78 can be upperly, strictly bounded by $\epsilon \max_i \sum_{c\in[d]} |m_c^i|$. Moreover, it is clear that

$$\sum_{c\in[d]} |m_c^i| \leq \sum_{c\in[d]} \|\boldsymbol{m_c}\|_\infty.$$

Therefore, assuming that the last sum is smaller than $4\epsilon/(1+2\epsilon)$, the dependence factor $D(\partial)$ satisfies

$$\begin{aligned} |D(\partial) - 1| &< \epsilon \max_i \sum_{c\in[d]} |m_c^i| \leq \\ &\leq \epsilon \sum_{c\in[d]} \|\boldsymbol{m_c}\|_\infty, \end{aligned} \tag{4.79}$$

proving the first inequality of the statement. The second inequality follows by reusing the assumption $\sum_{c\in[d]} \|\boldsymbol{m_c}\|_\infty \leq 4\epsilon/(1+2\epsilon)$.

□

Appendix 4.C Equivalent definitions of unsolvability

Unsolvability for 2-dimensional or highly symmetric transition matrices was studied in Mossel [2001a] using the Total Variation (TV) distance.

Definition 4.5 (Unsolvability in Mossel [2001a])**.** *We say that the reconstruction problem is TV-unsolvable if the likelihood vector* $\boldsymbol{\rho_\partial} = (\rho_\partial^0, \cdots, \rho_\partial^K)$ *satisfies, for all* $i \neq j \in \mathbb{A}$,

$$\sum_{\partial\in\Delta_g} |\rho_\partial^i - \rho_\partial^j| \to 0 \text{ as } g \to \infty.$$

Now define $\mathbb{E}_\partial^\pi[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - 1\|_*]$ as the expected value of $\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - 1\|_*$ with prior $\boldsymbol{\mu} = \boldsymbol{\pi}$. Recall that $\boldsymbol{\mu} > 0$ by definition and $\boldsymbol{\pi} > 0$ due to Prop. 4.1.a. In Prop. 4.14, we prove that Definitions 4.4 and 4.5 are equivalent.

Proposition 4.14. *The following statements are equivalent:*

a) The reconstruction problem is unsolvable.

b) The reconstruction problem is unsolvable with stationary prior $\boldsymbol{\mu} = \boldsymbol{\pi}$.

c) The reconstruction problem is TV-unsolvable.

Proof.

- *a)* $\iff$ *b)*: Since $\Pr_\pi(\partial) = \boldsymbol{\pi} \cdot \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$ and $\Pr(\partial) = \boldsymbol{\mu} \cdot \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}}$, it holds that

$$\Pr(\partial) \min_i \pi_i \leq \Pr_\pi(\partial) \leq \frac{\Pr(\partial)}{\min_i \mu_i}, \tag{4.80}$$

and consequently

$$\mathbb{E}_\partial[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - 1\|_*] \min_i \pi_i \leq \mathbb{E}^\pi_\partial[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - 1\|_*] \leq \frac{\mathbb{E}_\partial[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - 1\|_*]}{\min_i \mu_i}. \tag{4.81}$$

These chain of inequalities implies that $\mathbb{E}_\partial[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - 1\|_*] \to 0$ iff $\mathbb{E}^\pi_\partial[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - 1\|_*] \to 0$, as desired. Note that we can chose any prior as long as it has positive entries.

- *b)* $\iff$ *c)*: If we write $\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = (\tilde{\rho}^0_\partial, \cdots, \tilde{\rho}^K_\partial)$, we can do

$$\sum_{\partial \in \Delta_g} |\rho^i_\partial - \rho^j_\partial| = \sum_{\partial \in \Delta_g} \Pr_\pi(\partial) |\tilde{\rho}^i_\partial - \tilde{\rho}^j_\partial| = \mathbb{E}^\pi_\partial[|\tilde{\rho}^i_\partial - \tilde{\rho}^j_\partial|] \tag{4.82}$$

We have the inequality

$$|\tilde{\rho}^i_\partial - \tilde{\rho}^j_\partial| \leq |\tilde{\rho}^i_\partial - 1| + |\tilde{\rho}^j_\partial - 1|, \tag{4.83}$$

and consequently unsolvability in expectation (using the L_1-norm) implies TV-unsolvability. Conversely, since $\boldsymbol{\pi}$ is a distribution and $\boldsymbol{\pi} \cdot \tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} = 1$, it holds that

$$\begin{aligned} |\tilde{\rho}^i_\partial - 1| &= |\sum_j \tilde{\rho}^i_\partial \pi_j - \sum_j \tilde{\rho}^j_\partial \pi_j| = \\ &= |\sum_{j \neq i} \pi_j (\tilde{\rho}^i_\partial - \tilde{\rho}^j_\partial)| \leq \sum_{j \neq i} \pi_j |(\tilde{\rho}^i_\partial - \tilde{\rho}^j_\partial)|. \end{aligned} \tag{4.84}$$

Thus if $\mathbb{E}^\pi_\partial[|\tilde{\rho}^i_\partial - \tilde{\rho}^j_\partial|] \to 0$ for all $i \neq j$ then $\mathbb{E}_\partial[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - 1\|_1] \to 0$, as desired.

□

References in this chapter

J. Felsenstein. Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, 17(6):368–376, 1981. ISSN 1432-1432. doi: 10.1007/BF01734359.

O. Gascuel and M. Steel. Predicting the Ancestral Character Changes in a Tree is Typically Easier than Predicting the Root State. *Systematic Biology*, 63(3): 421–435, 02 2014. ISSN 1063-5157. doi: 10.1093/sysbio/syu010. URL `https://doi.org/10.1093/sysbio/syu010`.

S. Guiasu. *Information Theory with Applications*. McGraw-Hill Companies, 1977.

R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*, chapter Chapter 3, pages –. Cambridge University Press, 1991. doi: 10.1017/CBO9780511840371.

D. A. Levin and Y. Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

C. Manuel. An upper bound of the information flow from children to parent node on trees. *arXiv preprint arXiv:2204.10618*, 2022.

B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, and R. Lanfear. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, 37 (5):1530–1534, 02 2020. ISSN 0737-4038. doi: 10.1093/molbev/msaa015. URL `https://doi.org/10.1093/molbev/msaa015`.

E. Mossel. Reconstruction on Trees: Beating the Second Eigenvalue. *The Annals of Applied Probability*, 11(1):285 – 300, 2001a. doi: 10.1214/aoap/998926994. URL `https://doi.org/10.1214/aoap/998926994`.

E. Mossel. Survey: Information flow on trees. In *Graphs, Morphisms and Statistical Physics*, 2001b.

E. Mossel and Y. Peres. Information flow on trees. *The Annals of Applied Probability*, 13(3):817 – 844, 2003. doi: 10.1214/aoap/1060202828. URL https://doi.org/10.1214/aoap/1060202828.

E. Mossel and M. Steel. How much can evolved characters tell us about the tree that generated them? In O. Gascuel, editor, *Mathematics of Evolution and Phylogeny*. Oxford University Press, USA, 2007.

A. Stamatakis. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313, 01 2014. ISSN 1367-4803. doi: 10.1093/bioinformatics/btu033. URL https://doi.org/10.1093/bioinformatics/btu033.

Steven G. Johnson. Notes on the equivalence of norms. https://math.mit.edu/~stevenj/18.335/norm-equivalence.pdf, 2020.

Chapter 5

Conclusion and Outlook

This thesis emphasised the fundamental importance of theory. We did so because a solid theory provides both a constructive way to infer new results and an efficient way to dismiss unrealistic assumptions.

In Chapter 2, we characterized the connectivity of taboo-free graphs whose taboos do not change in time. However, taboo-sets often vary along evolution. As shown by Rusinov et al. [2015], the lifespan of the taboo-set determines up to which extent taboos are actually avoided. Thus a natural follow-up of Chapter 2 would be the simulation of taboo-free evolution. Manuel et al. [unpublished] prepared such simulations, but the quantitative influence of taboo-free evolution on sequence evolution seemed too weak to be significant for real data. All in all, we can argue that taboo-sets typically have little impact on phylogenetic inference. This is reassuring for all applications.

In Chapter 3 we proposed new measures of phylogenetic information. Although the spectral decomposition of the coherence and the memory seems technical (Sec. 3.7), our results show the big potential of these new measures for phylogenetics. Indeed, we obtained not only an unprecedentedly simple estimate of branch length in Eq. 3.45, but also the asymptotic test for saturation, which has nearly optimal power and does not depend on the estimated length of the branch tested (Sec. 3.10).

Leaving aside our new measures of information, the statistical definition of saturation using hypothesis testing (Subsec. 3.10.1) could be applied to more conventional measures, as the entropy. Thus we expect this formalization to be useful for future research. Another important observation is given in Subsec. 3.10.5: In a reconstructed tree, saturated branches can be relocated anywhere in the tree without significantly affecting the tree log-likelihood. Thus every saturated branch enlarges the space of equally good tree topologies. This is a similar phenomenon to that of terraces, where the lack of genes by some species leads to a large space of true

topologies explaining the data equally well [Sanderson et al., 2011].

Future work should elucidate whether conflicting phylogenies using real data are a consequence of branch saturation. A significant conflict would be the placement of the Last Universal Common Ancestor (LUCA) in the tree of life, typically on the branch between Bacteria and the rest of living organisms [Brown and Doolittle, 1995], but for a minority studies elsewhere, e.g. inside Bacteria [Cavalier-Smith, 2006].

Regarding Chapter 4, Theorems 4.4 and 4.7 are very general and allow a big space of rate matrices. This generality shows how naturally the norm of the memory vector can quantify identification, giving good reasons to explore its usage. Considering the results of Chapter 3, we believe that the memory at the root $M(R) := \mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi^2]$ better adapts to phylogenetic reconstruction than the expected norm $\mathbb{E}[\|\tilde{\boldsymbol{\rho}}_{\boldsymbol{\partial}} - \mathbf{1}\|_\pi]$ studied in Chapter 4. Indeed, the asymptotics of the log-likelihood stated in Section 3.9 show the importance of the linear properties of the coherence and the memory. If we define *the saturation center of a tree* as the root with the minimum sample memory $\hat{M}(R)$, future work could compare the saturation center with the root obtained from some rooting method. The relative location of the saturation center and the root could provide valuable insights about the evolutionary process underwent by past species, for example a higher rate of substitution in a particular clade of the tree.

References in this chapter

J. R. Brown and W. F. Doolittle. Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. *Proceedings of the National Academy of Sciences*, 92(7):2441–2445, 1995.

T. Cavalier-Smith. Rooting the tree of life by transition analyses. *Biology direct*, 1 (1):1–83, 2006.

C. Manuel, S. Pfannerer, and A. von Haeseler. Etaboo: Modelling and measuring taboo-free evolution. Unpublished, unpublished.

I. Rusinov, A. Ershova, A. Karyagina, S. Spirin, and A. Alexeevski. Lifespan of restriction-modification systems critically affects avoidance of their recognition sites in host genomes. *BMC Genomics*, 16(1):1084, 2015. ISSN 1471-2164. doi: 10.1186/s12864-015-2288-4.

M. J. Sanderson, M. M. McMahon, and M. Steel. Terraces in phylogenetic tree space. *Science*, 333(6041):448–450, 2011. doi: 10.1126/science.1206357. URL `https://www.science.org/doi/abs/10.1126/science.1206357`.

References (combined)

F. Ailloud, X. Didelot, S. Woltemate, G. Pfaffinger, J. Overmann, R. C. Bader, C. Schulz, P. Malfertheiner, and S. Suerbaum. Within-host evolution of Helicobacter pylori shaped by niche-specific adaptation, intragastric migrations and selective sweeps. *Nature Communications*, 10(1):2273, 2019. ISSN 2041-1723. doi: 10.1038/s41467-019-10050-1.

B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. Watson. *Molecular Biology of the Cell*, chapter 8, pages 532–534. Garland, 5th edition, 2004.

J. W. Archie. A Randomization Test for Phylogenetic Information in Systematic Data. *Systematic Biology*, 38(3):239–252, 09 1989. ISSN 1063-5157. doi: 10.2307/2992285. URL `https://doi.org/10.2307/2992285`.

A. Asinowski, A. Bacher, C. Banderier, and B. Gittenberger. *Analytic Combinatorics of Lattice Paths with Forbidden Patterns: Enumerative Aspects*, pages 195–206. ., 01 2018. doi: 10.1007/978-3-319-77313-1-15.

A. Asinowski, A. Bacher, C. Banderier, and B. Gittenberger. Analytic combinatorics of lattice paths with forbidden patterns, the vectorial kernel method, and generating functions for pushdown automata. *Algorithmica*, pages 1–43, 10 2019. doi: 10.1007/s00453-019-00623-3.

J. R. Brown and W. F. Doolittle. Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. *Proceedings of the National Academy of Sciences*, 92(7):2441–2445, 1995.

T. Cavalier-Smith. Rooting the tree of life by transition analyses. *Biology direct*, 1 (1):1–83, 2006.

M. M. Collery, S. A. Kuehne, S. M. McBride, M. L. Kelly, M. Monot, A. Cockayne, B. Dupuy, and N. P. Minton. What's a SNP between friends: The influence of single nucleotide polymorphisms on virulence and phenotypes of Clostridium difficile strain 630 and derivatives. *Virulence*, 8(6):767–781, Aug 2017. ISSN 2150-5608. doi: 10.1080/21505594.2016.1237333.

D. A. Duchêne, N. Mather, C. Van Der Wal, and S. Y. W. Ho. Excluding Loci With Substitution Saturation Improves Inferences From Phylogenomic Data. *Systematic Biology*, 71(3):676–689, 09 2021. ISSN 1063-5157. doi: 10.1093/sysbio/syab075. URL `https://doi.org/10.1093/sysbio/syab075`.

B. Efron. Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer, 1992.

Encyclopedia of Mathematics. Encyclopedia of Mathematics, by Springer Verlag GmbH and European Mathematical Society. URL: https://encyclopediaofmath.org/wiki/Newton_method. Accessed on 2021-09-09, 2021a.

Encyclopedia of Mathematics. Encyclopedia of Mathematics, by Springer Verlag GmbH and European Mathematical Society. URL: https://encyclopediaofmath.org/wiki/Maximum-likelihood_method. Accessed on 2021-09-09, 2021b.

J. Felsenstein. Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, 17(6):368–376, 1981. ISSN 1432-1432. doi: 10.1007/BF01734359.

J. Felsenstein. Confidence limits on phylogenies: an approach using the bootstrap. *evolution*, 39(4):783–791, 1985.

J. Felsenstein. *Inferring phylogenies*, chapter 16. Sinauer Associates, 2004.

W. M. Fitch and E. Margoliash. A method for estimating the number of invariant amino acid coding positions in a gene using cytochrome c as a model case. *Biochemical Genetics*, 1(1):65–71, Jun 1967. ISSN 1573-4927. doi: 10.1007/BF00487738.

B. Foley, T. Leitner, C. Apetrei, B. Hahn, I. Mizrachi, J. Mullins, A. Rambaut, S. Wolinsky, and B. Korber. HIV Sequence Compendium 2018. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, NM, LA-UR 18-25673, 2020.

O. Gascuel and M. Steel. Predicting the Ancestral Character Changes in a Tree is Typically Easier than Predicting the Root State. *Systematic Biology*, 63(3): 421–435, 02 2014. ISSN 1063-5157. doi: 10.1093/sysbio/syu010. URL `https://doi.org/10.1093/sysbio/syu010`.

M. Gelfand and E. Koonin. Avoidance of palindromic words in bacterial and archaeal genomes: A close connection with restriction enzymes. *Nucleic acids research*, 25: 2430–9, 07 1997. doi: 10.1093/nar/25.24.0.

X. Gu, Y. X. Fu, and W. H. Li. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Molecular Biology and Evolution*, 12(4):546–557, 07 1995. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a040235. URL `https://doi.org/10.1093/oxfordjournals.molbev.a040235`.

S. Guiasu. *Information Theory with Applications*. McGraw-Hill Companies, 1977.

D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, and L. S. Vinh. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, 35(2):518–522, 10 2017. ISSN 0737-4038. doi: 10.1093/molbev/msx281.

R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*, chapter Chapter 3, pages –. Cambridge University Press, 1991. doi: 10.1017/CBO9780511840371.

W. J. Hsu and M. J. Chung. Generalized Fibonacci Cubes. *1993 International Conference on Parallel Processing - ICPP'93*, 1:299–302, Aug 1993. doi: 10.1109/ICPP.1993.95.

A. Ilić, S. Klavžar, and Y. Rho. Generalized Fibonacci cubes. *Discrete Mathematics*, 312:2 – 11, 2012.

M. Kimura. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution*, 16(2):111–120, 1980.

S. Klavžar. Structure of Fibonacci cubes: a survey. *Journal of Combinatorial Optimization*, 25:505–522, 2013. doi: 10.1007/s10878-011-9433-z.

V. Kommireddy and V. Nagaraja. Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiology and molecular biology reviews : MMBR*, 77:53–72, 03 2013. doi: 10.1128/MMBR.00044-12.

LANL. Los Alamos National Laboratory. http://www.hiv.lanl.gov/, 2020. Accessed: 2020-12-07.

D. A. Levin and Y. Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

C. Manuel. An upper bound of the information flow from children to parent node on trees. *arXiv preprint arXiv:2204.10618*, 2022.

C. Manuel and A. von Haeseler. Structure of the space of taboo-free sequences. *Journal of Mathematical Biology*, 81(4):1029–1057, 2020.

C. Manuel, S. Pfannerer, and A. von Haeseler. Etaboo: Modelling and measuring taboo-free evolution. Unpublished, unpublished.

B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, and R. Lanfear. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, 37 (5):1530–1534, 02 2020. ISSN 0737-4038. doi: 10.1093/molbev/msaa015. URL `https://doi.org/10.1093/molbev/msaa015`.

E. Mossel. Reconstruction on Trees: Beating the Second Eigenvalue. *The Annals of Applied Probability*, 11(1):285 – 300, 2001a. doi: 10.1214/aoap/998926994. URL `https://doi.org/10.1214/aoap/998926994`.

E. Mossel. Survey: Information flow on trees. In *Graphs, Morphisms and Statistical Physics*, 2001b.

E. Mossel and Y. Peres. Information flow on trees. *The Annals of Applied Probability*, 13(3):817 – 844, 2003. doi: 10.1214/aoap/1060202828. URL `https://doi.org/10.1214/aoap/1060202828`.

E. Mossel and M. Steel. How much can evolved characters tell us about the tree that generated them? In O. Gascuel, editor, *Mathematics of Evolution and Phylogeny*. Oxford University Press, USA, 2007.

J. Neyman and E. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London*, IX(231): 289–337, 1933. doi: 10.1098/rsta.1933.0009.

T. Popoviciu. Sur les équations algébriques ayant toutes leurs racines réelles. *Mathematica*, 9(129-145):20, 1935.

REBASE. The Restriction Enzyme Database: Archea. http://rebase.neb.com/rebase/arcbaclistA.html, 2020a. Accessed: 2020-06-17.

REBASE. The Restriction Enzyme Database: Bacteria. http://rebase.neb.com/rebase/arcbaclistB.html, 2020b. Accessed: 2020-06-17.

R. J. Roberts, T. Vincze, J. Posfai, and D. Macelis. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Research*, 43(D1):D298–D299, 11 2014. ISSN 0305-1048. doi: 10.1093/nar/gku1046.

E. Rocha, A. Danchin, and A. Viari. Evolutionary role of restriction/modification systems as revealed by comparative genome analysis. *Genome Research*, 11, 06 2001. doi: 10.1101/gr.GR-1531RR.

I. Rusinov, A. Ershova, A. Karyagina, S. Spirin, and A. Alexeevski. Lifespan of restriction-modification systems critically affects avoidance of their recognition sites in host genomes. *BMC Genomics*, 16(1):1084, 2015. ISSN 1471-2164. doi: 10.1186/s12864-015-2288-4.

I. S. Rusinov, A. S. Ershova, A. S. Karyagina, S. A. Spirin, and A. V. Alexeevski. Avoidance of recognition sites of restriction-modification systems is a widespread but not universal anti-restriction strategy of prokaryotic viruses. *BMC Genomics*, 19(1):885, 2018a. ISSN 1471-2164. doi: 10.1186/s12864-018-5324-3.

I. S. Rusinov, A. S. Ershova, A. S. Karyagina, S. A. Spirin, and A. V. Alexeevski. Comparison of methods of detection of exceptional sequences in prokaryotic genomes. *Biochemistry (Moscow)*, 83(2):129–139, 2018b. ISSN 1608-3040. doi: 10.1134/S0006297918020050.

M. Salemi. Genetic distances and nucleotide substitution models: practice. In P. Lemey, M. Salemi, and V. Anne-Mieke, editors, *The Phylogenetic Handbook: a Practical Approach to Phylogenetic Analysis and Hypothesis Testing*, pages 125–141. Cambridge University Press, Cambridge, 2 edition, 2009. doi: 10.1017/CBO9780511819049.006.

P. Sanders and C. Schulz. High quality graph partitioning. *Proceedings of the 10th DIMACS implementation challenge workshop*, 03 2013. doi: 10.1090/conm/588.

M. J. Sanderson, M. M. McMahon, and M. Steel. Terraces in phylogenetic tree space. *Science*, 333(6041):448–450, 2011. doi: 10.1126/science.1206357. URL `https://www.science.org/doi/abs/10.1126/science.1206357`.

C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.

J. S. Shoemaker and W. M. Fitch. Evidence from nuclear sequences that invariable sites should be considered when sequence divergence is calculated. *Molecular Biology and Evolution*, 6(3):270–289, 05 1989. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a040550.

M. E. Siddall. Success of Parsimony in the Four-Taxon Case: Long-Branch Repulsion by Likelihood in the Farris Zone. *Cladistics*, 14(3):209–220, 1998. ISSN 0748-3007. doi: https://doi.org/10.1006/clad.1998.0063. URL `https://www.sciencedirect.com/science/article/pii/S0748300798900639`.

A. Stamatakis. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313, 01 2014. ISSN 1367-4803. doi: 10.1093/bioinformatics/btu033. URL `https://doi.org/10.1093/bioinformatics/btu033`.

Steven G. Johnson. Notes on the equivalence of norms. `https://math.mit.edu/~stevenj/18.335/norm-equivalence.pdf`, 2020.

K. Strimmer and A. von Haeseler. Genetic distances and nucleotide substitution models. In P. Lemey, M. Salemi, and V. Anne-Mieke, editors, *The Phylogenetic Handbook: a Practical Approach to Phylogenetic Analysis and Hypothesis Testing*, pages 111–141. Cambridge University Press, Cambridge, 2 edition, 2009. doi: 10.1017/CBO9780511819049.006.

D. W. Ussery, T. M. Wassenaar, and S. Borini. *Computing for Comparative Microbial Genome: Bioinformatics for Microbiologists*. Springer., 1st edition, 2008. ISBN 978-1-84800-254-8.

N. D. Weber, M. Aubert, C. H. Dang, D. Stone, and K. R. Jerome. DNA cleavage enzymes for treatment of persistent viral infections: recent advances and the pathway forward. *Virology*, 454-455:353–361, Apr 2014. doi: 10.1016/j.virol.2013.12.037.

R. J. Wilson. *Introduction to Graph Theory*. John Wiley & Sons, Inc., New York, NY, USA, 1986. ISBN 0-470-20616-0.

X. Xia, Z. Xie, M. Salemi, L. Chen, and Y. Wang. An index of substitution saturation and its application. *Molecular phylogenetics and evolution*, 26(1):1–7, 2003.

Z. Yang. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular evolution*, 39(3):306–314, 1994.

L. Yuan, X.-Y. Huang, z.-y. Liu, F. Zhang, X. L. Zhu, J.-Y. Yu, X. Ji, Y. Xu, G. Li, C. Li, H.-J. Wang, Y.-Q. Deng, M. Wu, M.-L. Cheng, Q. Ye, D.-Y. Xie, X.-F. Li, X. Wang, W. Shi, and C.-F. Qin. A single mutation in the prM protein of Zika virus contributes to fetal microcephaly. *Science*, 358, 09 2017. doi: 10.1126/science.aam7120.