A tensorial approach to the inversion of group-based phylogenetic models

Jeremy G Sumner*, Peter D Jarvis and Barbara R Holland

Abstract

Background: Hadamard conjugation is part of the standard mathematical armoury in the analysis of molecular phylogenetic methods. For group-based models, the approach provides a one-to-one correspondence between the so-called “edge length” and “sequence” spectrum on a phylogenetic tree. The Hadamard conjugation has been used in diverse phylogenetic applications not only for inference but also as an important conceptual tool for thinking about molecular data leading to generalizations beyond strictly tree-like evolutionary modelling.

Results: For general group-based models of phylogenetic branching processes, we reformulate the problem of constructing a one-one correspondence between pattern probabilities and edge parameters. This takes a classic result previously shown through use of Fourier analysis and presents it in the language of tensors and group representation theory. This derivation makes it clear why the inversion is possible, because, under their usual definition, group-based models are defined for abelian groups only.

Conclusion: We provide an inversion of group-based phylogenetic models that can implemented using matrix multiplication between rectangular matrices indexed by ordered-partitions of varying sizes. Our approach provides additional context for the construction of phylogenetic probability distributions on network structures, and highlights the potential limitations of restricting to group-based models in this setting.

Keywords: Groups, Representation theory, Symmetry, Markov chains

Background

Fundamental to evolutionary biology is the development and implementation of molecular phylogenetic methods [1]. These methods provide the means to reconstruct the past evolutionary history of biological entities given present-day molecular data, such as DNA. Considering Kimura’s neutral theory of molecular evolution, it is logical to apply a stochastic model at the level of DNA substitutions to construct probabilistic description of what molecular alignments are expected to be observed, given a proposed evolutionary history (tree topology and edge lengths). is commonly implemented assuming an IID (across sites in the alignment) and Markov process for DNA substitution, leading to a model that has a continuous-time Markov chain at its core (see Semple and Steel [2] for an introduction to the mathematics underlying modern phylogenetic methodology).

In a series of papers, Hendy and colleagues introduced the Hadamard conjugation as a novel tool for phylogenetic analyses [3-5]. They found an invertible relationship between a phylogenetic tree, as characterized by its edge length spectrum, and the probability distribution of site patterns (referred to as the sequence spectrum). Originally introduced only for the 2-state symmetric model, the Hadamard conjugation was later extended to the K3ST model [6-8] and further to any of the so-called “group-based” models [9]. Hadamard conjugation has been used as both a tool for simulation [10] and to look at statistical properties of methods, exploring the inconsistency of parsimony under a molecular clock [5,11]. For these sorts of applications, following the notation in Felsenstein [1], we can use the Hadamard transform $H$ to start with an edge length spectrum $\gamma$ and calculate the sequence spectrum $s = H^{-1} \log(H\gamma)$. The beauty of Hadamard conjugations is that one can also begin with an observed sequence spectrum $\hat{s}$ and perform the inverse of the conjugation to empirically obtain an edge length spectrum...
\[ \gamma = H^{-1} \log(H\hat{s}). \]

Although it is not expected that the \( \gamma \) spectrum will precisely match a tree, Hendy [12] proposed using an optimisation criterion to map from \( \gamma \) to the “closest tree”.

Several authors have commented that it is potentially a useful feature of Hadamard conjugation that data isn’t forced onto a fixed tree. The conflicting information can be retained and interpreted in the form of a “lentoplot” [13] or a splits-graph [14], with both of these methods implemented in Spectronet [15]. Schliep [16] gives some more statistical justification for such an approach by making a link to modern statistical techniques such as the Lasso and Ridge regression.

von Haeseler and Churchill [17] seems to be the first paper that explicitly suggests using Hadamard conjugation to provide a likelihood framework for networks. The principle idea in this work was to start with an edge length spectrum that encodes a set of incompatibilities, use the Hadamard transformation to get site probabilities and use these to determine a likelihood. This idea was further explored by Bryant [18], and Bryant [19] followed this through defining the “n-taxon process” for group-based models. It should be noted that likelihoods calculated via Hadamard are not equivalent to likelihoods calculated by taking a mixture of trees. Indeed, Matsen and Steel [20], Matsen et al. [21] used Hadamard methods in combination with phylogenetic invariants to show that mixtures of trees with the same topology can exactly mimic another tree under the 2-state model. Considering biological applications, thinking in terms of mixtures of trees or partitions where the data can be thought of as arising on a set of trees [22-24] seems more reasonable than the Hadamard conjugation. Strimmer and Moulton [25] suggested using split networks as a spring board to likelihood-based analyses on DAGs, but later identified several problems with the approach [26]; most notably, in split-networks internal nodes do not have a biological interpretation as an ancestor.

In Sumner et al. [27], we gave some additional insight into the interpretation of applying the Hadamard conjugation in a network setting. We showed that permutation group structure inherent to the Hadamard transformation – as for any group-based model – restricts the resulting process from being capable of reproducing truly convergent processes. This is a serious limitation, as one of the biological motivations for explicit network models is the ability to model convergent processes. We also presented an alternative algebraic formalism for the general Markov model, analogous to the n-taxon process, but capable of reproducing convergent processes.

From the point of view of group representation theory, the inversion of group-based models relies on the fact that the irreducible representations of an abelian group are one-dimensional, and the model structure essentially reduces to analysing group characters – hence the standard presentation of a Fourier inversion. In this article, we make this connection concrete. For the general Markov model, it is then immediately apparent that an analogous inversion is not possible because the algebraic structure underlying the model is not abelian and hence the irreducible representations are not one-dimensional. In fact, to obtain one-dimensional representations for the general Markov model, it is necessary to apply higher-degree polynomial maps (beyond the degree 1, linear case), and define “Markov invariants” [28]. These invariants present one-dimensional representations but at the cost of the higher degree – degree 5 in the case of the general Markov model with four states on quartet trees [29,30]. This connection between Hadamard transformation and Markov invariants is an interesting one, but we do not discuss it further here.

In this paper we approach the inversion of group-based phylogenetic models by taking a representation-theoretic perspective and working explicitly with tensor indices. Our approach rests heavily on the formalism of “phylogenetic tensors”, as presented in Bashford et al. [31], for the binary-symmetric and K3ST model, and Sumner et al. [27,28], for the general Markov model.

Although the main inversion results presented here are not more general than those in in Székely et al. [7], we think it is important to reformulate them using the language of tensors and representation theory. This viewpoint has already led to new approaches for modeling convergent evolution [27] and for studying non-group-based models [28]. However, in none of our previous work was the link to Hadamard conjugation explicitly discussed. By presenting an old technique (Hadamard conjugation) in a new light we hope to introduce other researchers to the viewpoint of tensor analysis and representation theory.

**Methods**

**Group-based models**

We consider the continuous-time formulation of Markov processes, and show how to implement the inversion of a group-based phylogenetic model based on any abelian group \( G \). We note that such an inversion requires a map from tensor product space (where elements are indexed by ordered-\( n \)-partitions) to phylogenetic splits (where elements are indexed by bipartitions). We achieve this by finding canonical maps from bipartitions to ordered-\( n \)-partitions.

For a group \( G \) (not necessarily abelian) with order \( |G| = d \), we write \( G = \{\sigma_1, \sigma_2, \ldots, \sigma_d\} \), and, when necessary, write \( \epsilon \in G \) to specify the identity element of \( G \). We will discuss the “regular representation” of \( G \) shortly, but
skipping ahead we find that any rate matrix $Q$ occurring in a group-based Markov model can be written in the form

$$Q = -\lambda 1 + \sum_{\sigma \in G} \alpha^\sigma K_\sigma,$$  

(1)

where each $0 \leq \alpha^\sigma \in \mathbb{R}$, $\lambda = \sum_{\sigma \in G} \alpha^\sigma$ and the $K_\sigma$ are the permutation matrices corresponding to the (non-identity) group elements $\sigma \in G$.

For the reader interested in deriving this result, consider the $d$-dimensional vector space $\langle G \rangle \subset \mathbb{C} = \{v_1, \sigma_2, \ldots, \sigma_d\} = \{v = v_1 \sigma_1 + v_2 \sigma_2 + \ldots + v_d \sigma_d : v_i \in \mathbb{C}\}$, with scalar multiplication and vector addition defined via

$$v + v' = (v_1 \sigma_1 + v_2 \sigma_2 + \ldots + v_d \sigma_d) + (v'_1 \sigma_1 + v'_2 \sigma_2 + \ldots + v'_d \sigma_d) = (v_1 + v'_1) \sigma_1 + (v_2 + v'_2) \sigma_2 + \ldots + (v_d + v'_d) \sigma_d,$$

for all $v, v' \in \langle G \rangle$ and $\lambda \in \mathbb{C}$. The regular representation, $\rho_{\text{reg}} : G \rightarrow GL(d, \mathbb{C})$, is then defined by setting the group action

$$\sigma : v \mapsto \sigma v = v_1 (\sigma_1 \sigma_1) + v_2 (\sigma_2 \sigma_2) + \ldots + v_d (\sigma_d \sigma_d),$$

for all $v \in \langle G \rangle$ and $\sigma \in G$. If we fix $\{\sigma_1, \sigma_2, \ldots, \sigma_d\}$ as an ordered basis for $\langle G \rangle$, it is then clear – via Cayley’s theorem – that each group element $\sigma$ gets mapped to a permutation matrix $K_\sigma := \rho_{\text{reg}}(\sigma)$, with $K_\sigma^{-1} = \sum_{\sigma} [K_\sigma]_i^j = \sigma \sigma_i$. Thus $K_\sigma$ has matrix elements

$$[K_\sigma]_i^j = \begin{cases} 1, & \text{if } \sigma_i = \sigma_j, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Consider the unit column vectors

$$\xi_1 = (1, 0, 0, \ldots, 0)^T, \quad \xi_2 = (0, 1, 0, \ldots, 0)^T, \quad \ldots$$

$$\xi_d = (0, 0, \ldots, 1, 0)^T,$$

and identify each $\sigma_i \in \langle G \rangle$ with $\xi_i \in \mathbb{C}^d$, so that the group action becomes $\sigma : \xi \mapsto K_\sigma \xi = \xi_j$ where $j = \sigma_i$. Thus the matrix elements $[K_\sigma]_i^j$ have $i$ as the column label and $j$ as the row label.

A group-based Markov model is then obtained by taking a continuous-time Markov chain with state space $G \subset \{\sigma_1, \sigma_2, \ldots, \sigma_d\}$ and using the group multiplication in $G$ to assign a rate $\alpha_\theta$ to all substitutions $\sigma_1 \mapsto \sigma_2$ where $\theta = \sigma_1 \sigma_2$. Following this through (as is done in detail in [32]) we are led to the formula (1) for rate matrices in any group-based model.

The regular representation is one example of the general concept of a representation of $G$ on a vector space $V$, defined as a homomorphism $\rho : G \rightarrow GL(V)$ satisfying

$$\rho(g_1 g_2) = \rho(g_1) \rho(g_2)$$

for all $g_1, g_2 \in G$. A representation is said to be reducible if there exists a proper subspace $U \subset V$ satisfying $\rho(g) U \subset U$, i.e. the set of matrices $\rho(G)$ send vectors in $U$ back to $U$. In this case, $U$ is called an invariant subspace. The representation $\rho$ is then called irreducible if $V$ does not contain any invariant subspaces.

The reader should note that the usual construction of a “group-based” model [2] stipulates that $G$ be abelian. Although the construction just given using the regular representation allows for non-abelian $G$, we will nonetheless only consider the abelian case in this paper, because, as discussed in the introduction, it is only in the abelian case that a (linear) inversion of phylogenetic models is possible. In this case the irreducible representations of $G$ are all one-dimensional [33], and hence the analysis reduces to computations with group characters, as is exploited in the previous approaches using Fourier analysis [9,34].

**Phylogenetic tensors**

We denote $[d] := \{1, 2, \ldots, d\}$ as the state space for a continuous-time Markov chain. Consider an $n$-taxa phylogenetic tree and a $d$-state phylogenetic pattern distribution $[p_{i_1, i_2, \ldots, i_d}]_{i_1, i_2, \ldots, i_d \in [d]}$ with the interpretation that $p_{i_1, i_2, \ldots, i_d}$ is the probability that the observed state at the $k$th leaf on the tree is $i_k$. As is shown in Sumner and Jarvis [35] and in more detail in Sumner et al. [27], such phylogenetic pattern distributions can be represented abstractly as tensors in the $n$-fold tensor product space $\otimes^n \mathbb{C}^d := \mathbb{C}^d \otimes \mathbb{C}^d \otimes \ldots \otimes \mathbb{C}^d$, as follows. If we choose $\{\xi_1, \xi_2, \ldots, \xi_d\}$ as an ordered basis for $\mathbb{C}^d$, and ordered basis $\{\xi_{i_1} \otimes \xi_{i_2} \otimes \ldots \otimes \xi_{i_d}\}_{i_1, i_2, \ldots, i_d \in [d]}$ for the tensor product space, a “phylogenetic tensor” $P \in \otimes^n \mathbb{C}^d$ is then defined as

$$P = \sum_{i_1, i_2, \ldots, i_d \in [d]} p_{i_1, i_2, \ldots, i_d} \xi_{i_1} \otimes \xi_{i_2} \otimes \ldots \otimes \xi_{i_d}.$$
focus heavily on the indexing of tensor components in the various cases discussed.

Suppose $\pi = \sum_{i \in [d]} \pi_i \xi_i \in \mathbb{C}^d$ represents the state distribution of a single taxa, i.e. $\pi_i$ is the probability that a randomly chosen site in the sequence will be in state $i$. Now suppose a phylogenetic branching event occurs and the sequence is copied. The corresponding phylogenetic tensor $P = \sum_{i_1,i_2 \in [d]} p_{i_1,i_2} \xi_{i_1} \otimes \xi_{i_2}$ representing the joint distribution of the two-taxa just after the branching event then has the property that $p_{i_1,i_2} = \pi_{i_1}$ if $i_2 = i_1$ and is zero otherwise. Thinking in terms of tensor operations, we find that phylogenetic branching events can be generated by a linear operator $\delta : \mathbb{C}^d \rightarrow \mathbb{C}^d \otimes \mathbb{C}^d$ determined by $\delta(\pi) = P$ and defined in general using our chosen basis as

$$
\delta(\xi_i) := \xi_i \otimes \xi_i,
\delta(\pi) = \delta \left( \sum_i \pi_i \xi_i \right) = \sum_i \pi_i \delta(\xi_i)
= \sum_i \pi_i \xi_i \otimes \xi_i.
$$

The remarkable fact for group-based models, central to the present article, is that the permutation matrices “intertwine” particularly simply with the branching operator:

$$
\delta(K_{\sigma} \xi_i) = \delta(\xi_{\sigma(i)}) = \xi_{\sigma(i)} \otimes \xi_{\sigma(i)} = K_{\sigma} \otimes K_{\sigma} \cdot \delta(\xi_i).
$$

Thus, for any rate matrix $Q$ arising from a group-based model, we have (via the linearity of $\delta$):

$$
\delta \cdot Q = \left( -\lambda_e 1 + \sum_{\alpha \neq \sigma \in G} \alpha^\sigma K_{\sigma} \otimes K_{\sigma} \right) \cdot \delta.
$$

(3)

We also note that, since $Q$ can be expressed a linear combination of permutation matrices representing elements in a group $G$, the matrix powers $Q^2, Q^3, Q^4$ . . . will also be expressible as linear combinations of the same permutation matrices (although precise expressions for the relevant coefficients may or may not be easily computable). Together with (3), this implies that, for any substitution matrix $e^{Qt}$ arising from matrix exponentiation,

$$
\delta \cdot e^{Qt} = e^{-\lambda t} \exp \left( \sum_{\alpha \neq \sigma \in G} \alpha^\sigma K_{\sigma} \otimes K_{\sigma} \right) \cdot \delta.
$$

(4)

This relation shows that mathematically, and hence conceptually, “Markov evolution on a single followed by a branching event” can be replaced with “Branching event on a single taxon followed by (correlated) Markov evolution of two taxa.” This equivalence is illustrated in Figure 1, and should be compared to the equivalent discussion of the “$n$-taxa process” given in [18] and [19].

In Sumner et al. [27] we showed how to generalise this intertwining action to the case of the general Markov model. Interestingly, for the general Markov model the appropriate intertwining has quite a different structure from what occurs in group-based models, and hence the simplicity of (4) is somewhat misleading in general. We refer the reader to Sumner et al. [27] for more discussion on this point.

Returning to the case of group-based models, for each subset $A \subseteq [n]$, we define a linear map on $\otimes^n \mathbb{C}^d$ as the tensor product $K_{\sigma}^{(A)} := K_{\sigma^1} \otimes K_{\sigma^2} \otimes \ldots \otimes K_{\sigma^n}$ where $a_i = 1$ if $i \in A$ and 0 otherwise. For example, if $n = 5$, we have

$$
K_{\sigma}^{(1,2,4)} = K_{\sigma} \otimes K_{\sigma} \otimes 1 \otimes K_{\sigma} \otimes 1.
$$

To develop a phylogenetic tensor on a tree, we root the phylogenetic tree at taxon $n$, and label edges by subsets $\emptyset \neq e \subseteq [n-1]$, where $i \in e$ if the path from taxon $n$ to taxon $i$ crosses the edge labelled by $e$. A five taxon tree with this labelling, is presented in Figure 2. To each edge labelled by $\emptyset \neq e \subseteq [n-1]$, we assign the rate matrix

$$
Q_e := -\lambda_e 1 + \sum_{\alpha \neq \sigma \in G} \alpha^\sigma K_{\sigma},
$$

where each $\alpha^\sigma \geq 0$ is the rate of substitution for all states $\sigma_1$ to $\sigma_2$ satisfying $\sigma = \sigma_2 \sigma_1^{-1}$, and $\lambda_e = \sum_{\sigma \in G} \alpha^\sigma$. Each edge is then assigned substitution matrix $M_e = e^{Q_e}$, so
that the time parameter for each edge is absorbed into the definition of $Q_e$.

Now iterating (4) multiple times, Bashford et al. [27,31] show that any phylogenetic tensor can be written as

$$P = e^{-\lambda} \exp \left( \sum_{\varnothing \neq e \subseteq [n-1]} \alpha^e_{\sigma} K_{\sigma}^{(e)} \right) \cdot \delta^{n-1}\pi. \quad (5)$$

where $\lambda = \sum_{\varnothing \neq e \subseteq [n-1]} \lambda_{\sigma}$, and $\delta^{n-1}\pi$ is the $d \times d \times \ldots \times d$ tensor that represents the "zero edge-length star tree" distribution on $n$ taxa. It is this form of phylogenetic tensors that will do a lot of the heavy lifting in the discussion that follows. The reader should note that under this representation, there is no need for the edge parameters $\{\alpha^e_{\sigma} : \varnothing \neq e \subseteq [n-1], \sigma \in G\}$ to be chosen to be compatible with a particular tree, hence the possibilities for generalising to non-tree-like or network models, as discussed in the introduction.

The stationary distribution for group-based models is uniform (because the rate matrices are doubly stochastic). In this paper we always assume a stationary distribution, so that:

$$\pi = \frac{1}{d} (1, 1, \ldots, 1)^T,$$

and $\delta^{n-1}\pi$ has tensor components

$$[\delta^{n-1}\pi]_{i_1 i_2 \ldots i_n} = \begin{cases} \frac{1}{d}, & \text{if } i_1 = i_2 = \ldots = i_n, \\ 0, & \text{otherwise.} \end{cases}$$

This concludes our discussion of the tensor presentation of phylogenetic probability distributions under group-based models. It is important to note that everything discussed so far works for any group-based model, with no requirement that the underlying group $G$ be abelian.

In what follows, we discuss the inversion of abelian group-based models. We present the simplest case with $G = \mathbb{Z}_2$; the $G = \mathbb{Z}_3$ case; the $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ case; the general $G = \mathbb{Z}_r$ case; and finally we discuss the case of any abelian group.

**Results**

**The binary-symmetric case**

We begin with the inversion of the so-called "binary-symmetric" model. Consider $\mathbb{C}^2$ with standard basis

$$\{\xi_0 = (1, 0), \xi_1 = (0, 1)\}.$$

As a group-based model, the binary-symmetric model arises by taking the group

$$G := \mathbb{Z}_2 = \{0, 1\}_{\text{mod 2}} \cong \{\sigma |\sigma^2 = \epsilon\},$$

with a generic rate matrix given by

$$Q = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = -1 + K,$$

where $K = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is the permutation matrix representing $\sigma$ in the standard basis.

Now $\rho_{\text{reg}} : \mathbb{Z}_2 \rightarrow M_2(\mathbb{C})$, with $\sigma \mapsto K$, is the regular representation of $\mathbb{Z}_2$, and the character table of $\mathbb{Z}_2$ given in Table 1 is easily recognised to be the Hadamard matrix

$$h = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$  

As $\mathbb{Z}_2$ is an abelian group, the irreducible representations are one-dimensional.

The corresponding projection operators can be read off from the columns of the character table. That is, the operators

$$\Theta_{\text{id}} := \frac{1}{2} (\epsilon + \sigma), \quad \Theta_{\text{sgn}} := \frac{1}{2} (\epsilon - \sigma);$$

project $\rho_{\text{reg}} = \text{id} \oplus \text{sgn}$ onto the id and sgn representations of $\mathbb{Z}_2$, respectively.

**Table 1 The character table of $\mathbb{Z}_2$**

| $\epsilon$ | $\text{id}$ | $\text{sgn}$ |
|-----------|-------------|-------------|
| $\sigma$  | 1           | -1          |

Figure 2 A six taxa tree rooted at taxon 6 with edges labelled by subsets of $\{1, 2, 3, 4, 5\}$. 

$\{1, 2, 3, 4, 5\}$

$\{1\}$

$\{2\}$

$\{3\}$

$\{4\}$

$\{5\}$

$\{1, 2\}$

$\{3, 4, 5\}$

$\{4, 5\}$

$\{\} \times \cdots \times \{\}$

$\{\}$

$\{\}$

$\{\}$

$\{\}$
This observation prompts us to work in the alternative basis:

\[
\begin{align*}
f_0 : &= \Theta_{id} \cdot \xi_0 = \Theta_{id} \cdot \xi_1 = h\xi_0 = \xi_0 + \xi_1, \\
f_1 : &= \Theta_{sgn} \cdot \xi_0 = -\Theta_{sgn} \cdot \xi_1 = h\xi_1 = \xi_0 - \xi_1.
\end{align*}
\]

In this basis the permutation matrix is diagonal:

\[
\hat{K} := hKh^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \hat{Q} := -1 + \hat{K} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.
\]

The representation-theoretic perspective on \( \hat{K} \) is to observe that \( id(\sigma) = 1 \) and \( \text{sgn}(\sigma) = -1 \).

Referring to (5), we know that we can write a generic phylogenetic tensor as

\[
P = e^{-\lambda} \exp \left( \sum_{\emptyset \neq e \subseteq [n-1]} \alpha_e K^{(e)} \right) \cdot \delta^{n-1}\pi,
\]

where \( \lambda = \sum_{\emptyset \neq e \subseteq [n-1]} \alpha_e \).

We index matrix and tensor indices by using \( i, j, k = 0, 1 \in \mathbb{Z}_2 \) and allow multiplication \( \times \) in the ring of integers \( \mathbb{Z} \). The Hadamard matrix then has matrix elements \( [H]^j_i = (-1)^{i \cdot j} \) where \( i \) is the row index and \( j \) is the column index. Observe that in the diagonal basis, the permutation matrix has elements

\[
[K]^j_i = \delta_{ij}(-1)^i.
\]

Thus we have expressions such as

\[
[K^{(2,3)}]^i_{i'j} = \delta_{i'i} \delta_{j'j} \delta_{i'j}(-1)^{i' + ij},
\]

where \( K^{(2,3)} = 1 \otimes \hat{K} \otimes \hat{K} \).

As we are dealing with tensors of arbitrary size, it is convenient to represent a string such as \( i_1i_2 \ldots i_n \) as an ordered-bipartition \( \mu = \mu_0 \sqcup \mu_1 \) of the set \( [n] \), where \( \mu_0, \mu_1 \subseteq [n] \) with \( j \in \mu_k \) if and only if \( i_j = k \). For example we have the following equivalences:

\[
00110 \equiv \{1, 2, 5\} : \{3, 4\}, \quad 01111 \equiv \{1\} : \{2, 3, 4, 5\}, \\
10001 \equiv \{2, 3, 4\} : \{1, 5\}
\]

and inequivalence:

\[
01010 \equiv \{1, 3, 5\} : \{2, 4\} \neq \{2, 4\} : \{1, 3, 5\} \equiv 10101.
\]

We then have

\[
[K^{(e)}]^i_{i'j} = \delta_{i'i} \delta_{j'j} \delta_{i'j}(-1)^{i' + ij},
\]

Defining \( h^{(n)} := h^{(n-1)} \otimes h \) where \( h^{(1)} := h \), in the diagonal basis \( \hat{P} := h^{(n)} \cdot \hat{P} \) and using our notation \( \hat{h}^{(n)} \) has tensor components

\[
\left[h^{(n)}\right]_{\mu \nu} = \left[h^{(n)}\right]_{\mu_0 \nu_1} = \left[h^{(n)}\right]_{\nu_0 \mu_1} = \left[h^{(n)}\right]_{\mu_0 \nu_1}. \quad (|\mu| + |\nu| = n)
\]

The zero edge-length star-tree initial distribution has tensor components

\[
\left[\hat{\delta}^{n-1}\pi\right]_{i_1i_2 \ldots i_n} = \frac{1}{2} \delta_{i_1i_2} \delta_{i_1i_3} \ldots \delta_{i_1i_n},
\]

(once we have given preference to taxon 1 in this expression, there are many ways that this distribution can be expressed using the \( \delta_{ij} \)). In the diagonal basis with \( \hat{\delta}^{n-1}\pi := h^{(n)} \cdot \hat{\delta}^{n-1}\pi \), we have components

\[
\left[\hat{\delta}^{n-1}\pi\right]_{\mu_0 \mu_1} = \frac{1}{2} \left(1 + (-1)^{|\mu|}\right).
\]

Since \( \hat{K} \) is diagonal in the transformed basis, we can conclude that

\[
\hat{P} = \hat{P} \cdot \hat{h}^{(n)} \cdot \hat{P} = e^{-\lambda} \exp \left( \sum_{\emptyset \neq e \subseteq [n-1]} \alpha_e \left[K^{(e)}\right]_{\mu_0 \mu_1} \right) \frac{1}{2} \left(1 + (-1)^{|\mu|}\right).
\]

Of course many of these tensor components will be zero and we would like to ignore these.

Take \( u = u_0 : u_1 \) as an ordered bipartition of the reduced set \( [n-1] \), so that \( u \equiv i_1i_2 \ldots i_{n-1} \) where \( j \in u_k \) if and only if \( i_j = k \), and define

\[
\gamma(u) = \begin{cases} 
0 \quad &\text{if } |\mu_1| \text{ is even}, \\
1 \quad &\text{if } |\mu_1| \text{ is odd};
\end{cases}
\]

and interpret \( u \cdot \gamma(u) \) as a string: \( u \cdot \gamma(u) = i_1i_2 \ldots i_{n-1} \gamma(u) \).

If we make the definitions

\[
P_u := \hat{P} \cdot u \cdot \gamma(u), \quad \eta_u := \frac{1}{2} \sum_{\emptyset \neq e \subseteq [n-1]} \alpha_e \left[K^{(e)}\right]_{u \cdot \gamma(u)},
\]

then...
then we can write the non-zero components as
\[ P_u = e^{-\lambda} \exp (\eta_u), \]
with inverses
\[ \eta_u = \ln (P_u) + \lambda. \] (6)

This is the first part of the inversion.

We would like to go further and actually recover the individual edge weights \( \alpha_e \). To do this we define the (square) \( 2^{n-1} \times 2^{n-1} \) matrix \( F \) with components
\[ [F]_u^e := \left[ F(e) \right]_{u}^{y(u)} = (-1)^{e \cap u} = \left[ h^{(n-1)} \right]_u, \]
with \( e \) a subset and \( u \) an ordered-bipartition of \([n-1]\). As \( (h^{(n-1)})^2 = \frac{1}{2^{n-1}} \mathbb{I} \), we see that \( F \) provides its own inverse \( F^{-1} \)
with components
\[ [F^{-1}]_e^u = \frac{1}{2^{n-1}} [F]_e^u. \]

Defining the column vectors \( \tilde{\alpha} = \{ \alpha_e \} \) and \( \tilde{\eta} = \{ \eta_u \} \), we can write the matrix equations
\[ \tilde{\eta} = F \tilde{\alpha}, \quad \tilde{\alpha} = F^{-1} \tilde{\eta}. \]

Together with the first part of the inversion (6), these equations give a one-one map between pattern probabilities and edge weights for the binary-symmetric model.

**Inversion of the \( \mathbb{Z}_3 \) model**

Taking confidence from the previous case we now discuss the inversion of the group-based phylogenetic model with \( G = \mathbb{Z}_3 \). We take
\[ \mathbb{Z}_3 = \{0, 1, 2\}_{+,(\text{mod } 3)} \cong \langle \sigma | \sigma^3 = 1 \rangle, \]
and, by analogy to the \( \mathbb{Z}_2 \) case, index tensors with indices \( i,j \in 0, 1, 2 \) and allow multiplication \( \times \) by extending \( \mathbb{Z}_3 \) to the ring \( \mathbb{F}_3 = \{0, 1, 2\}_{+,(\text{mod } 3)} \).

In this case a generic rate matrix is given by
\[
Q = \begin{pmatrix}
-(\alpha + \beta) & \beta & \alpha \\
\alpha & -(\alpha + \beta) & \beta \\
\beta & \alpha & -(\alpha + \beta)
\end{pmatrix}
= -(\alpha + \beta) \mathbf{1} + \alpha K_1 + \beta K_2,
\]
where
\[
K_1 = \begin{pmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 1 & 0
\end{pmatrix}, \quad K_2 = \begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0
\end{pmatrix},
\]
are the matrices representing the permutations \( \sigma \cong (123) \) and \( \sigma^2 \cong (132) \) under the regular representation, respectively.

We define \( \omega = e^{2\pi i/3} \), and present the character table of \( \mathbb{Z}_3 \) in Table 2. The decomposition of the regular representation is \( \rho_{\text{reg}} = \pi \oplus \omega \oplus \omega^2 \), and the columns of the character table give the projection operators onto the (one-dimensional) irreducible subspaces:
\[
\Theta_{\pi} : = \frac{1}{3} (\epsilon + \sigma + \sigma^2), \quad \Theta_{\omega} : = \frac{1}{3} (\epsilon + \omega \sigma + \omega^2 \sigma^2), \quad \Theta_{\omega^2} : = \frac{1}{3} (\epsilon + \omega^2 \sigma + \omega \sigma^2).
\]

Therefore, the matrix
\[
f = fQf^{-1} = \begin{pmatrix}
0 & 0 & 0 \\
0 & \omega + \beta \omega^2 & 0 \\
0 & 0 & \alpha \omega^2 + \beta \omega
\end{pmatrix},
\]
or, equivalently,
\[
\hat{K}_1 = fK_1 f^{-1} = \begin{pmatrix}
1 & 0 & 0 \\
0 & \omega & 0 \\
0 & 0 & \omega^2
\end{pmatrix},
\quad \hat{K}_2 = fK_2 f^{-1} = \begin{pmatrix}
1 & 0 & 0 \\
0 & \omega^2 & 0 \\
0 & 0 & \omega
\end{pmatrix}.
\]

We recall our basic result (5) that for group-based models, a generic phylogenetic tensor can be expressed as
\[
P = e^{-\lambda} \exp \left( \sum_{\emptyset \neq \sigma \subseteq [n-1]} (\alpha_e K_1^{(e)} + \beta_e K_2^{(e)}) \right) \cdot \delta^{n-1} \pi,
\]
where \( \lambda = \sum_{\emptyset \neq \sigma \subseteq [n-1]} (\alpha_e + \beta_e) \). We take the stationary distribution as initial distribution, so \( \pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T \).

The matrix elements of \( f \) can be expressed as \( [f]_i^j = \omega^{i \cdot j} \), we extend \( i,j \in \mathbb{Z}_3 \) to include multiplication \( \times \) from the ring of integers \( \mathbb{Z} \). Similarly,
\[
[\hat{K}_1]_i^j = \delta_{ij} \omega^i, \quad [\hat{K}_2]_i^j = \delta_{ij} (\omega^2)^i.
\]

**Table 2** The character table of \( \mathbb{Z}_3 \)

|   | \( \iota \) | \( \omega \) | \( \omega^2 \) |
|---|---|---|---|
| \( \sigma \) | 1 | \( \omega \) | \( \omega^2 \) |
| \( \sigma^2 \) | 1 | \( \omega^2 \) | \( \omega \) |
More generally, tensorial components can be expressed as

$$[\hat{K}_1 \otimes \hat{K}_1]^{1j_2 j_3}_{i_1 i_2 i_3} = \delta_{i_1 i_2} \delta_{j_2 j_3} \omega^2 i_1 + i_3.
$$

We represent a string $i_1 i_2 \ldots i_n$ as an ordered-tripartition, $i_1 i_2 \ldots i_n \equiv \mu = \mu_0 ; \mu_1 ; \mu_2$, of the set $[n]$, where $j \in \mu_k$ if and only if $i_j = k$. For example, if we take $n = 5$, we have:

00000 $\equiv \{1, 2, 3, 4, 5\}$

01210 $\equiv \{2, 5\} ; \{3\} ; \{1, 4\}$

01122 $\equiv \{1\} ; \{2, 3\} ; \{4, 5\}$.

Taking $n = 3$, we have

$$[\hat{K}^{(2,3)}_1]^{\mu}_{i_1 i_2 i_3} = \left[ 1 \otimes \hat{K}_1 \otimes \hat{K}_1 \right]^{\nu}_{\mu_1 \mu_2 \mu_3} = \delta_{\mu_1 \omega_1} \delta_{\mu_1(2,3)\mu_3},
$$

and in general:

$$[\hat{K}^{(e)}_1]^{\mu}_{i_1 i_2 i_3} = \delta_{\mu_1 \omega_1} \delta_{\mu_1(2,3)\mu_3},$$

$$[\hat{K}^{(e)}_2]^{\mu}_{i_1 i_2 i_3} = \delta_{\mu_2 \omega_1} \delta_{\mu_2(2,3)\mu_3}.$$

Taking the uniform distribution as initial distribution, the initial star-tree distribution can be written as

$$[\hat{\delta}^{n-1}\pi]_{i_1 i_2 \ldots i_n} = \frac{1}{2} \delta_{i_1 i_2} \delta_{i_3 i_4} \ldots \delta_{i_{n-1} i_n}.$$

Defining $f^{(n)} = f^{(n-1)} \otimes f$ where $f^{(1)} = f$, we have

$$f^{(n)} = f^{(n-1)} f^{(2)} \ldots f^{(n-1)} f^{(1)} = \omega^{j_1 + j_2 + \ldots + j_n} f^{(n)}.$$

and in the transformed basis, where $\hat{\delta}^{n-1} \pi := f^{(n)} \cdot \hat{\delta}^{n-1} \pi$, we have

$$[\hat{\delta}^{n-1}\pi]_{i_1 i_2 \ldots i_n} = \frac{1}{2} \sum_{j_1, j_2 \ldots, j_n} \omega^{j_1 + j_2 + \ldots + j_n} f^{(n)}.$$

Indexing by ordered-tripartitions, we conclude that

$$[\hat{\delta}^{n-1}\pi]_{\mu} = \frac{1}{2} \left( 1 + \omega^{j_1 + j_2 + \ldots + j_n} + (\omega^2)^{j_1 + j_2 + \ldots + j_n} \right)$$

Now suppose $|\mu_1| + 2|\mu_2| = 0 \pmod{3}$, then

$$[\hat{\delta}^{n-1}\pi]_{\mu} = \frac{1}{2} \left( 1 + 1 + 1 \right) = 1.$$

If $|\mu_1| + 2|\mu_2| = 1 \pmod{3}$, then

$$[\hat{\delta}^{n-1}\pi]_{\mu} = \frac{1}{2} \left( 1 + \omega + \omega^2 \right) = 0,$$

and if $|\mu_1| + 2|\mu_2| = 2 \pmod{3}$, then

$$[\hat{\delta}^{n-1}\pi]_{\mu} = \frac{1}{2} \left( 1 + \omega^2 + \omega \right) = 0.$$

Thus we have found a basis where all the elements of the initial star-tree tensor are zero unless the tripartition $\mu$ satisfies $|\mu_1| + 2|\mu_2| = 0 \pmod{3}$. Crucially, this statement also holds for the phylogenetic tensor $\hat{P}$ because in this basis the rate matrices of this model are diagonal:

$$[\hat{P}]_{\mu} = [\hat{P}]_{\mu_0 \mu_1 \mu_2},$$

$$= e^{-\lambda} \exp \left( \frac{1}{2} \sum_{\emptyset \neq \phi \subseteq [n-1]} \left[ \alpha_\phi K_1^{(e)} + \beta_\phi K_2^{(e)} \right]_{\mu_0 \mu_1 \mu_2} \right) \times \frac{1}{2} \left( 1 + \omega^{1|\mu_1|} + \omega^{2|\mu_2|} \right).$$

We deal with this condition on $\mu$ by taking $u = u_0 u_1 u_2$ as an ordered-tripartition of the reduced set $[n-1]$ and setting $\mu = u \cdot \gamma(u)$ (considered as the concatenation of strings) where

$$\gamma(u) = \begin{cases} 0, & \text{if } |u_1| + 2|u_2| = 0, \\ 1, & \text{if } |u_1| + 2|u_2| = 2, \\ 2, & \text{if } |u_1| + 2|u_2| = 1. \end{cases}$$

$$3 - (0|u_0| + 1|u_1| + 2|u_2|) \pmod{3}.$$
it follows that
\[
[K_1^{(e)}]_{u,γ′(u)}^{u′,γ(u)} = ω^{[υ_u|υ_{u′}|υ_{u′}]} = [K_1^{(e)}]_{u,γ′(u)}^{u′,γ(u)},
\]
and similarly
\[
[K_2^{(e)}]_{u,γ′(u)}^{u′,γ(u)} = ω^{[υ_{u′}+e|υ_{u′}+e]},
\]
We make the observation that
\[
[F_1]_{u}^{(e;γ)} := \left[f^{(n-1)}\right]_{u;g;u_1,u_2}^{e;γ} = ω^{[υ_{u_1}+γ|υ_{u_2}]} = [K_1^{(e)}]_{u,γ′(u)}^{u′,γ(u)},
\]
and
\[
[F_2]_{u}^{(e;γ)} := \left[f^{(n-1)}\right]_{u;g;u_1,u_2}^{e;γ} = ω^{[υ_{u_2}+e|υ_{u_1}+γ]} = [K_1^{(e)}]_{u,γ′(u)}^{u′,γ(u)},
\]
where \(F_1\) and \(F_2\) are \(2^{n-1} \times 3^{n-1}\) matrices.

Thus we may write
\[
η_u = \sum_{\emptyset \neq \emptyset^c \subseteq [n-1]} α_ε[F_1]_{u}^{ε} + β_ε[F_2]_{u}^{ε}.
\]

Defining the column vectors \(\vec{α} = \{|α_ε|\}, \vec{β} = \{|β_ε|\}\) and \(\vec{η} = |η_u|\), we can write
\[
\vec{η} = F_1\vec{α} + F_2\vec{β},
\]
and define two \(3^{n-1} \times 2^{n-1}\) matrices \(G_1\) and \(G_2\) as
\[
[G_1]_{e}^{u} := \left[f^{−1(n−1)}\right]_{u;g;u_1,u_2}^{e;γ} \quad [G_2]_{e}^{u} := \left[f^{−1(n−1)}\right]_{u;g;u_1,u_2}^{e;γ},
\]
where
\[
f^{-1} = \begin{pmatrix}
1 & 1 & 1 \\
1 & ω & ω^2 \\
1 & ω^2 & ω
\end{pmatrix},
\]
with \(ff^{-1} = 1\).

Considering that
\[
\sum_{v}[f^{−1(n−1)}]_{u}^{v}[f^{−1(n−1)}]_{v}^{w} = δ_{uvw},
\]
for all ordered-tripartitions \(u, w\) of \([n−1]\), we have the matrix products
\[
G_1F_1 = 1, \quad G_1F_2 = 0,
\]
\[
G_2F_1 = 1, \quad G_2F_2 = 0.
\]

Thus the second part of the inversion for this model is
\[
\vec{α} = G_1\vec{η}, \quad \vec{β} = G_2\vec{η}.
\]

Together with (7), these equations give a one-one map between pattern probabilities and edge weights for the group-based model with \(G = \mathbb{Z}_3\).

**Inversion of the K3ST model**

We now consider the K3ST model [36] which occurs as the group-based model with
\[
G = \mathbb{Z}_2 \times \mathbb{Z}_2 = \{(0,0), (0,1), (1,0), (1,1)\} \oplus \mathrm{mod}\ 2
\]
\[
\cong \{(12)(34), (13)(24)\}.
\]

In this model a generic rate matrix is given by
\[
Q = −(α + β + γ) 1 + αK_{01} + βK_{10} + γK_{11},
\]
where
\[
K_{01} = 1 \otimes K = \begin{pmatrix}
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix},
\]
\[
K_{10} = K \otimes 1 = \begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0
\end{pmatrix},
\]
\[
K_{11} = K \otimes K = \begin{pmatrix}
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0
\end{pmatrix}.
\]

We already know that the \(2 \times 2\) Hadamard matrix \(h\) diagonalizes \(K\), so we see immediately that \(H = h \otimes h\) diagonalizes this model:
\[
\hat{K}_{01} := HK_{01}H^{-1} = 1 \otimes hKH^{-1} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0
\end{pmatrix},
\]
\[
\hat{K}_{10} := HK_{10}H^{-1} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0
\end{pmatrix},
\]
\[
\hat{K}_{11} := HK_{11}H^{-1} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0
\end{pmatrix}.
\]

Of course \(H\) is the character table of \(\mathbb{Z}_2 \times \mathbb{Z}_2\) and the permutation matrices (8), together with \(K_{00} := 1\), give the regular representation \(ρ_{reg} \cong id \oplus id \oplus id \oplus sgn \oplus sgn \oplus id \oplus sgn \otimes sgn\), where we recall the basic result that the tensor product of two irreducible representations of a group \(G\) gives an irreducible representation of \(G \times G\).
Simplifying notation, for this model we index tensors with indices given as pairs: \(i, j = 00, 01, 10, 11 \in \mathbb{Z}_2 \times \mathbb{Z}_2\); and we express the individual parts using lower case Roman characters. For example, we write \(i := ab = 01\), with \(a = 0\) and \(b = 1\). This gives matrix elements:

\[
\begin{align*}
\mathbf{K}_{01}^{cd_{ab}} &= \delta_{ac} \delta_{bd}(1)^{-b}, \\
\mathbf{K}_{11}^{cd_{ab}} &= \delta_{ac} \delta_{bd}(1)^{a+b};
\end{align*}
\]

and more complicated tensor products such as

\[
\begin{align*}
\mathbf{K}_{01} \otimes \mathbf{K}_{10} \otimes 1^{c_{1}c_{2}c_{3}d_{1}d_{2}d_{3}}_{a_{1}b_{1}a_{2}b_{2}a_{3}b_{3}} &= \delta_{a_{1}c_{1}} \delta_{b_{1}d_{1}} \delta_{a_{2}c_{2}} \delta_{b_{2}d_{2}} \delta_{a_{3}c_{3}} \delta_{b_{3}d_{3}} (1)^{b_{1}+b_{2}}.
\end{align*}
\]

Again we interpret strings such as \(\nu \equiv \alpha_{1}\alpha_{2}\ldots\alpha_{n}\) and \(\nu \equiv \beta_{1}\beta_{2}\ldots\beta_{n}\) as ordered-bipartitions \(\nu = \alpha_{0}1\uparrow_{1}1\) and \(\nu = \nu_{0}1\uparrow_{1}\) of the set \([n]\). We can then write matrix elements of tensor products as

\[
\begin{align*}
\mathbf{K}_{01}^{\nu_{1}\nu_{2}} &= \delta_{\mu_{1}} \delta_{\nu_v}(1)^{v_{1}+v_{2}}, \\
\mathbf{K}_{10}^{\nu_{1}\nu_{2}} &= \delta_{\mu_{1}} \delta_{\nu_v}(1)^{v_{1}+v_{2}}, \\
\mathbf{K}_{11}^{\nu_{1}\nu_{2}} &= \delta_{\mu_{1}} \delta_{\nu_v}(1)^{v_{1}+v_{2}}.
\end{align*}
\]

Taking the stationary distribution \(\pi = \frac{1}{2} (1,1,1,1)^{T}\) as initial distribution, the zero edge-length star-tree distribution is given by

\[
\begin{align*}
\tilde{\pi}^{-1} &= \frac{1}{4} \delta_{i_{1}i_{2}} \delta_{i_{1}i_{3}} \ldots \delta_{i_{1}i_{n}},
\end{align*}
\]

which in the finer index representation is

\[
\begin{align*}
\tilde{\pi}^{-1} &= \delta_{i_{1}j_{1}} \delta_{i_{2}j_{2}} \ldots \delta_{i_{n}j_{n}},
\end{align*}
\]

Recall that elements of the Hadamard matrix can be written as \([h]_{ab} = (-1)^{a+b}\), where \(a, b \in \mathbb{Z}_2\) and we allow multiplication \(\times\) by extending to the ring of integers \(\mathbb{Z}\). In the transformed basis, we have

\[
\begin{align*}
\mathbf{K}_{01}^{\nu_{1}\nu_{2}} &= \delta_{\mu_{1}} \delta_{\nu_v}(1)^{v_{1}+v_{2}}, \\
&= \frac{1}{4} \sum \delta_{i_{1}j_{1}} \delta_{i_{2}j_{2}} \ldots \delta_{i_{n}j_{n}},
\end{align*}
\]

We recall (5), so under this model we can express a generic phylogenetic tensor as

\[
P = e^{-\lambda} \exp \left( \sum_{\emptyset \neq e \subseteq [n-1]} \alpha_{e} K_{01}^{e} + \beta_{e} K_{10}^{e} + \gamma_{e} K_{11}^{e} \right) \cdot \delta^{n-1} \pi.
\]

To exclude the vanishing components we define, for all ordered bipartitions \(u = u_{0}1\uparrow_{1}\) and \(v = v_{0}1\uparrow_{1}\) of the set \([n]\), we define

\[
\gamma(u) = \begin{cases} 
0, & \text{if } |u_{1}| \text{ is even}, \\
1, & \text{if } |u_{1}| \text{ is odd};
\end{cases}
\]

and interpret \(u \cdot \gamma(u)\) as the string \(u \cdot \gamma(u) = a_{1}a_{2} \ldots a_{n-1}y(u)\). Then, for each pair \(u, v\) of ordered-bipartitions of \([n-1]\), we define

\[
\eta_{u,v} := \sum_{\emptyset \neq e \subseteq [n-1]} \alpha_{e} K_{01}^{e} + \beta_{e} K_{10}^{e} + \gamma_{e} K_{11}^{e} \bigg|_{u \cdot \gamma(u), v \cdot \gamma(v)}.
\]

This gives the inversion

\[
P_{u,v} = e^{-\lambda} \exp \left( \eta_{u,v} \right), \quad \eta_{u,v} = \lambda + \ln \left( \mathcal{P}_{u,v} \right).
\]

Consider the \(2^{n} \times 2^{n-1}\) rectangular matrices \(F_{01}, F_{10}\) and \(F_{11}\) with components

\[
\begin{align*}
[F_{01}]_{u,v} &= [K_{01}^{e}]_{u,v} = (-1)^{u_{1}+v_{1}}, \\
[F_{10}]_{u,v} &= [K_{10}^{e}]_{u,v} = (-1)^{v_{1}}, \\
[F_{11}]_{u,v} &= [K_{11}^{e}]_{u,v} = (-1)^{u_{1}+v_{1}};
\end{align*}
\]

where \(e \subseteq [n-1]\) and \(u = u_{0}1\uparrow_{1}\) and \(v = v_{0}1\uparrow_{1}\) are ordered-bipartitions of \([n-1]\). If we define the column vector \(\vec{\eta} := (\eta_{u,v})\) indexed by pairs of ordered-bipartitions and the column vectors \(\vec{\alpha} := (\alpha_{e})\) and \(\vec{\gamma} := (\gamma_{e})\) indexed by subsets of \([n-1]\), we then have the matrix equation

\[
\vec{\eta} = F_{01} \vec{\alpha} + F_{10} \vec{\beta} + F_{11} \vec{\gamma}.
\]

Writing \(H^{(n)} = H^{(n-1)} \otimes H\) with \(H^{(1)} = H\), we note that

\[
\begin{align*}
[F_{01}]_{u,v} &= [H^{(n-1)}]_{u,v}^{e_{u}e_{v}}, \\
[F_{10}]_{u,v} &= [H^{(n-1)}]_{u,v}^{e_{v}}, \\
[F_{11}]_{u,v} &= [H^{(n-1)}]_{u,v}^{e_{u}e_{v}};
\end{align*}
\]
and define the $2^n \times 2^n$ rectangular matrices $G_{01}, G_{10}$ and $G_{11}$ as
\[
G_{01}^{\mu, \nu} = \left[ H^{-1(n-1)} \right]_{\mu, \nu}^{\alpha, \beta}, \quad G_{10}^{\mu, \nu} = \left[ H^{-1(n-1)} \right]_{\mu, \nu}^{\alpha, \beta}.
\]
Noting that
\[
\sum_{\lambda, \omega} \left[ H^{-1(n-1)} \right]_{\lambda, \nu}^{\mu, \lambda} \left[ H^n - 1 \right]_{\lambda, \omega}^{\nu, \omega} = \delta_{\lambda \nu} \delta_{\omega, \nu},
\]
for all $\mu, \nu, \omega, \chi, \sigma$ ordered-bipartitions of $[n-1]$, we then have the matrix identities
\[
G_{01}F_{01} = 1, \quad G_{10}F_{10} = 1, \quad G_{11}F_{11} = 1,
\]
and
\[
G_{01}F_{10} = 0 = G_{01}F_{11} = G_{10}F_{11} = G_{10}F_{01} = G_{11}F_{01} = G_{11}F_{10}.
\]
Writing
\[
\bar{a} = G_{01} \tilde{a}, \quad \bar{b} = G_{10} \tilde{b}, \quad \bar{c} = G_{11} \tilde{c},
\]
completes the inversion for the K3ST model.

**Inversion of the $\mathbb{Z}_r$ model**

We now consider the group based model for $\mathbb{Z}_r = \{0, 1, 2, \ldots, (r-1)\} + (\bmod r) \cong \langle \sigma : \sigma^r = e \rangle$. For this model the generic rate matrix has the form
\[
Q = -\lambda 1 + \sum_{i=1}^{r} \alpha^i K_{\sigma^i},
\]
where $\lambda = \sum_{i=1}^{r} \alpha^i$ and
\[
K_{\sigma^i} = \begin{pmatrix}
0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 1 \\
\end{pmatrix},
\]
so that $K_{\sigma^i} = K^i_{\sigma}$.

Defining $\omega = e^\frac{2\pi i}{r}$, we have $\omega^r = 1$ and $1 + \omega + \omega^2 + \ldots + \omega^{r-1} = 0$ and $[f^j]^j = \omega^j$ where $i, j = 0, 1, 2, \ldots, r-1$. Of course, $f$ is the character table of $\mathbb{Z}_r$ and $[f^{-1}]^j = \frac{1}{r} \omega^{-j}$.

**Lemma 1.**

\[
\sum_{\nu} \left[ \mu \otimes f \otimes \ldots \otimes f \right]_{\nu}^{\mu, \nu} \left[ f^{-1} \otimes f^{-1} \otimes \ldots \otimes f^{-1} \right]_{\nu}^{\mu, \nu} = \delta_{\mu \nu},
\]
where $\mu, \nu, \mu'$ are ordered-$r$-partitions of the set $[n]$ defined by the strings $i_1 i_2 \ldots i_n$, $j_1 j_2 \ldots j_n$ and $k_1 k_2 \ldots k_n$, respectively.

**Proof.** The result is obvious by the definition of tensor product. However, explicitly we have
\[
\sum_{\nu} \left[ f \otimes f \otimes \ldots \otimes f \right]_{\nu}^{\mu, \nu} \left[ f^{-1} \otimes f^{-1} \otimes \ldots \otimes f^{-1} \right]_{\nu}^{\mu, \nu} = \frac{1}{r} \sum_{\nu} \left[ \sum_{i_1, i_2, \ldots, i_n} \omega^{i_1 i_2 + \ldots + i_n} \right] \delta_{i_1 j_1} \delta_{i_2 j_2} \ldots \delta_{i_n j_n}
\]
which clearly equals $1$ if $i_\ell - k_\ell = 0$ for all $\ell$, and, by repeatedly applying $1 + \omega + \omega^2 + \ldots + \omega^{r-1} = 0$, equals 0 otherwise.

Thus the change of basis $K_{\sigma^i} \mapsto \hat{K}_{\sigma^i} = fK_{\sigma^i}f^{-1}$ will give diagonal matrices $\hat{K}_{\sigma^i}$. Additionally,

**Lemma 2.** In the diagonal basis, the matrices $\hat{K}_{\sigma^i} := fK_{\sigma^i}f^{-1}$ have matrix elements given by $[\hat{K}_{\sigma^i}]_{ij} = \omega^i \delta_{ij}$.

**Proof.** Consider the matrix elements $[K_{\sigma^i}]_{ij} = \delta_{i \sigma^i(j)}$. Thus
\[
[fK_{\sigma^i}f^{-1}]_{ij} = \sum_{k, l} \omega^{k \sigma^i(j) - l} \omega^{l - j} = \sum_{l} \omega^{l (i + k) - l} = \omega^i \sum_{l} \omega^{l (i - j)} = \omega^i \delta_{ij},
\]
where we have used $\omega^{l (i + k) - l} = \omega^l$.

Now
\[
[\delta^{\mu - 1} \pi]_{i_1 i_2 \ldots i_n} = \frac{1}{r} \delta_{i_1 i_2} \delta_{i_1 i_3} \ldots \delta_{i_1 i_n},
\]
and
\[
[\delta^{\mu - 1} \pi]_{i_1 i_2 \ldots i_n} = \frac{1}{r} \sum_{\nu} \delta_{i_1 i_2 + \ldots + i_n \nu} \omega^{i_1 i_2 + \ldots + i_n \nu}
\]
and
\[
[\delta^{\mu - 1} \pi]_{i_1 i_2 \ldots i_n} = \frac{1}{r} \sum_{\nu} \delta_{i_1 i_2 + \ldots + i_n \nu} \omega^{i_1 i_2 + \ldots + i_n \nu}
\]
for $i_1 + i_2 + \ldots + i_n = 0 \bmod r$, otherwise.
Translating this result using the ordered-$r$-partitions for indices, we have

**Lemma 3.** In the diagonal basis, the uniform initial distribution on the star tree has components

\[
\begin{bmatrix}
\delta^{n-1}_\mu
\end{bmatrix}
= \begin{cases}
1 & \text{if } 0|\mu_0|+1|\mu_1|+2|\mu_2|+\ldots+(r-1)|\mu_{r-1}| = 0(\text{mod } r), \\
0, & \text{otherwise},
\end{cases}
\]

where \( \mu = \mu_0;\mu_1;\mu_2;\ldots;\mu_{r-1} \) is an ordered-$r$-partition of the set \([n]\).

Again recall that for this model a generic phylogenetic tensor can be written as

\[
P = e^{-\lambda} \exp \left( \sum_{\emptyset \neq e \subseteq [n-1], \ell \in [r-1]} \alpha^{(e)}_\mu K^{(e)}_{\sigma^e} \right) \delta^{n-1}_\pi,
\]

where \( \pi = \frac{1}{r}(1,1,\ldots,1)^T \). In the diagonal basis \( \hat{P} := f^{(n)} \cdot P \) and as a consequence of Lemma 3 \( \hat{P} \) will have many vanishing components. To avoid this we take \( u = u_0;u_1;u_2;\ldots;u_{r-1} \) as an ordered-$r$-partition of \([n-1]\) and set

\[
y(u) = r - (0|u_0|+1|u_1|+2|u_2|+\ldots+(r-1)|u_{r-1}|) \pmod{r}.
\]

If we define \( \mathcal{P}_u := [\hat{P}]_{u,y(u)} \) and

\[
\eta_u := \left[ \sum_{\emptyset \neq e \subseteq [n-1], \ell \in [r-1]} \alpha^{(e)}u_{\mu}K^{(e)}_{\sigma^e} \right]_{u,y(u)}
\]

we then have the first part of the inversion for the \( \mathbb{Z}_r \) model:

\[
\mathcal{P}_u = e^{-\lambda} \exp(\eta_u),
\]

\[
\eta_u = \ln(\mathcal{P}_u) + \lambda.
\]

For each \( i \in [r-1] \), we define the column vectors \( \vec{a}_i := [v_i]_{\emptyset \neq e \subseteq [n-1]} \) and, for each \( \emptyset \neq e \subseteq [n-1] \) and \( u \) an ordered-$(r-1)$-partition of \([n-1]\), we define the rectangular \( r^{n-1} \times 2^{n-1} \) matrices

\[
[F_i]_{\mu}^e := \left[K^{(e)}_{\mu} \right]_{u,y(u)}\quad (F_i)_{\mu}^e := \left[K^{(e)}_{\mu} \right]_{u,y(u)},
\]

so we have the vector equation

\[
\eta = F_1\vec{a}_1 + F_2\vec{a}_2 + \ldots + F_{r-1}\vec{a}_{r-1}.
\]

We claim that

**Lemma 4.**

\[
[F_1]_{\mu}^e = \left[f^{(n-1)}_{\mu} \right]_{\nu,\bar{\nu}(\emptyset;\bar{\mu})},
\]

\[
[F_2]_{\mu}^e = \left[f^{(n-1)}_{\mu} \right]_{\nu,\bar{\nu}(\emptyset;\bar{\mu})},
\]

\[
\ldots
\]

\[
[F_{r-1}]_{\mu}^e = \left[f^{(n-1)}_{\mu} \right]_{\nu,\bar{\nu}(\emptyset;\bar{\mu})}.
\]

**Proof.** We recall that \( [\hat{K}_{\sigma^e}]_{\mu} = \omega^{\mu e_0} \delta_{ij} \), so, for \( \mu = \mu_0;\mu_1;\mu_2;\ldots;\mu_{r-1} \) an ordered-$r$-partition of \([n]\), and \( e \) a subset of \([n-1]\) we have

\[
[\hat{K}_{\sigma^e}]_{\mu} = \omega^{(0|u_0|e_0)+1|u_1|e_0|+\ldots+(r-1)|u_{r-1}|e_0|},
\]

and

\[
[\hat{K}_{\sigma^e}]_{\mu} = \omega^{(0|u_0|e_0)+1|u_1|e_0|+\ldots+(r-1)|u_{r-1}|e_0|},
\]

because \( e \subseteq [n-1] \). On the other hand \( [f^{(n-1)}]^e_\mu = \omega^{e_0} \), so

\[
[f^{(n-1)}]_{\mu} = \omega^{e_0(0|u_0|e_0)+1|u_1|e_0|+\ldots+(r-1)|u_{r-1}|e_0|},
\]

where \( e \) appears in the \( e \)-th position.

Define, for \( i \in [r-1] \), the rectangular \( 2^{n-1} \times r^{n-1} \) matrices

\[
[G_1]_{e_0} = \left[f^{(n-1)}_{\mu} \right]_{\nu,\bar{\nu}(\emptyset;\bar{\mu})},
\]

\[
[G_2]_{e_0} = \left[f^{(n-1)}_{\mu} \right]_{\nu,\bar{\nu}(\emptyset;\bar{\mu})},
\]

\[
\ldots
\]

\[
[G_{r-1}]_{e_0} = \left[f^{(n-1)}_{\mu} \right]_{\nu,\bar{\nu}(\emptyset;\bar{\mu})}.
\]

Of course \( G_iF_j = \delta_{ij}I \), so we now have the second part of the inversion:

\[
\vec{a}_i = G_i\eta.
\]

Inversion of any abelian group-based model

**Lemma 5.** Any (finitely generated) abelian group \( G \) is isomorphic to a direct product of cyclic groups of prime-power order, i.e. \( G \cong \mathbb{Z}_{p_1} \times \mathbb{Z}_{p_2} \times \ldots \times \mathbb{Z}_{p_t} \) where each \( r_i = p_i^{n_i} \) where \( p_i \) is prime and \( n_i \) is a positive integer.
Lemma 6. The group-based model arising from the G is defined only up to group isomorphisms of G.

Proof: A generic rate matrix for the group-based model arising from G is given by

\[ Q = -\lambda 1 + \sum_{\sigma \in G} \alpha^\sigma K_\sigma. \]

Under a group isomorphism \( \phi : G \to G' \), we have \( \phi(\sigma_i \sigma_j) = \phi(\sigma_i)\phi(\sigma_j) \).

Recall (2), so that the matrix elements \([K_\sigma]^j_i\) is set via the action \( \sigma_i \mapsto \sigma_i \), so \( \sigma_i = \sigma \).

If we consider the regular representation of \( G' \) we then have \([K_{\phi(\sigma)}]^j_i\) defined by \( \phi(\sigma_i) \mapsto \phi(\phi(\sigma_i)) \).

Now \( \phi(\sigma_i)\phi(\sigma_j) = \phi(\sigma_i \sigma_j) = \phi(\sigma_j \sigma_i) \) and, because \( \phi \) is a group isomorphism, this occurs if and only if \( \sigma_i = \sigma_j \). Thus \([K_{\phi(\sigma)}]^j_i = [K_\sigma]^j_i\) for all \( i \) and \( j \).

This means that we can restrict attention to a single representative in the isomorphism class of \( G \). Of course, for this purpose we choose the representative guaranteed by Lemma 5.

Thus, for any abelian group \( G \) with generators \( \sigma_1, \sigma_2, \ldots, \sigma_q \) the corresponding group-based model has rate generators given by

\[ L_\sigma = -1 + K_{\sigma_1}^{m_1} \otimes K_{\sigma_2}^{m_2} \otimes \cdots \otimes K_{\sigma_q}^{m_q}, \]

for all \( e \neq \sigma = (\sigma_1^{m_1}, \sigma_2^{m_2}, \ldots, \sigma_q^{m_q}) \in G \), where \( K_\sigma \) is the permutation matrix representing the generator \( \sigma_i \in \mathbb{Z}_{r_i} \).

The character table \( f \) of \( G \) is simply the tensor product of the individual character tables of the \( \mathbb{Z}_{r_i} \):

\[ f = f_1 \otimes f_2 \otimes \cdots \otimes f_q. \]

In the diagonal basis we have matrix elements

\[ \left[ \tilde{K}_{f_i} K_{f_k}^{-1} \right]^j_i = (\omega_k)^{i_j} \delta_{ij}, \]

where \( \omega_k \) is a \( k \)-th root of unity. Thus

\[ \left[ \tilde{K}_{f_1} K_{f_2} \otimes \cdots \otimes K_{f_q} \right]_{i_1 \cdots i_q}^{j_1 \cdots j_q} = \delta_{i_1 j_1} \delta_{i_2 j_2} \cdots \delta_{i_q j_q} (\omega_k)^{i_1 m_1 + \cdots + i_q m_q}. \]

We write phylogenetic tensors for this model in the form

\[ P_{i_1 j_1 \cdots i_s j_s \cdots i_{s-q} j_{s-q} \cdots i_q j_q}, \]

where \( 0 \leq i_j \leq r_i \) for all \( 0 \leq s \leq q \). We simplify notation by writing each group of indices as \( \mu^{(s)} := i_1 i_2 \cdots i_m \) where \( \mu^{(s)} \) is an ordered-\( r_s \)-partition of \( \{n\} \).

Lemma 7. In the diagonal basis, the uniform initial distribution on the star tree has components

\[ \left[ \delta_n^{-1} \Pi \right]_{\mu^{(1)} \mu^{(2)} \cdots \mu^{(q)}} = \begin{cases} 1, & \text{if } 0 | \mu_0^{(i)} | + 1 | \mu_1^{(i)} | + \cdots + (r_i-1) | \mu_{r_i-1}^{(i)} | = 0 \forall i; \\ 0, & \text{otherwise}. \end{cases} \]

A generic phylogenetic tensor for this model can be expressed as

\[ P = e^{-\lambda} \exp \left( \sum_{\mu^{(i)} \in [n]-1} \alpha_{\mu^{(i)}} \phi_{\mu^{(i)}} \otimes \phi_{\mu^{(i)}} \otimes \cdots \otimes \phi_{\mu^{(i)}} \right) \delta^{n-1} \Pi, \]

where \( \Pi \) is the uniform distribution on \( \sum_{i=1}^{q} r_i \) states, i.e.

\[ \Pi = \left( \sum_{i=1}^{q} r_i \right)^{-1} (1, 1, \ldots, 1)^T. \]

In the diagonal basis \( \bar{P} = (f_1 \otimes f_2 \otimes \cdots \otimes f_q)^{(n)} \cdot P \), and, as a consequence of the previous lemma, \( P \) has many vanishing components. To avoid these, for each \( i \in \{q\} \) we take \( u_i^{(1)} = u_0^{(1)} u_1^{(1)} u_2^{(1)} \cdots u_{r_i-1}^{(1)} \) as an ordered-\( r_i \)-partition of \( [n-1] \) and set

\[ \gamma_i(u_i^{(1)}) = r_i - \left( 0 | u_0^{(1)} | + 1 | u_1^{(1)} | + 2 | u_2^{(1)} | + \cdots + (r_i-1) | u_{r_i-1}^{(1)} | \right) \text{ (mod } r_i \). \]

We then define

\[ P_{\mu^{(1)} \mu^{(2)} \cdots \mu^{(q)}} := \bar{P}_{\mu^{(1)} \gamma_1(u_0^{(1)}) \gamma_2(u_1^{(2)}) \gamma_3(u_2^{(3)}) \cdots \gamma_q(u_q^{(q)})}, \]

and

\[ \eta_{\mu^{(1)} \mu^{(2)} \cdots \mu^{(q)}} := \left[ \sum_{\mu^{(i)} \in [n]-1} \alpha_{\mu^{(i)}} \phi_{\mu^{(i)}} \otimes \cdots \otimes \phi_{\mu^{(i)}} \right]_{\mu^{(1)} \gamma_1(u_0^{(1)}) \mu^{(2)} \gamma_2(u_1^{(2)}) \cdots \gamma_q(u_q^{(q)})}, \]

so that we have the first part of the inversion

\[ P_{\mu^{(1)} \mu^{(2)} \cdots \mu^{(q)}} = e^{-\lambda} \exp \left( \eta_{\mu^{(1)} \mu^{(2)} \cdots \mu^{(q)}} \right), \]

\[ \eta_{\mu^{(1)} \mu^{(2)} \cdots \mu^{(q)}} = \lambda + \ln \left( P_{\mu^{(1)} \mu^{(2)} \cdots \mu^{(q)}} \right). \]
We define the column vectors \( \hat{\alpha}_{s_{1}^{2}...s_{q}} \) and \( \hat{\eta} \) := \( \{(u_{1},u_{2})...,(u_{n})\} \) where \( u_{i} \) is an ordered-\( r_{i} \)-partition of \([n-1] \), and we define the \( (r_{1}r_{2}...r_{q})^{n-1} \times 2^{n-1} \) matrices

\[
F_{s_{1}^{2}...s_{q}}u_{1}u_{2}...u_{q} := \left[ K_{1}^{(e)} \right]_{u_{1}u_{2}}(u_{1}) \left[ K_{2}^{(e)} \right]_{u_{2}u_{3}}(u_{2}) ... \left[ K_{q}^{(e)} \right]_{u_{q}u_{1}}(u_{q})
\]

\[
\hat{\eta} := \sum_{s_{1}^{2}...s_{q}:1 \leq n \leq n-1} F_{s_{1}^{2}...s_{q}}\hat{\alpha}_{s_{1}^{2}...s_{q}}.
\]

If we define the \( 2^{n-1} \times (r_{1}r_{2}...r_{q})^{n-1} \) matrices

\[
G_{s_{1}^{2}...s_{q}}u_{1}u_{2}...u_{q} := \left[ f_{1}^{(n-1)} \right]_{u_{1}}^{(s_{1})} \left[ f_{2}^{(n-1)} \right]_{u_{2}}^{(s_{2})} ... \left[ f_{q}^{(n-1)} \right]_{u_{q}}^{(s_{q})}
\]

where in each term \( e \) appears in the \( s_{i}^{th} \) position, we have the orthogonality relations

\[
G_{s_{1}^{2}...s_{q}}\delta_{s_{1}^{i}}\delta_{s_{2}^{j}}...\delta_{s_{q}^{k}} = \delta_{s_{1}^{i}}\delta_{s_{2}^{j}}...\delta_{s_{q}^{k}} \ 1.
\]

This gives us the second part of the inversion of any group-based model:

\[
\hat{\alpha}_{s_{1}^{2}...s_{q}} = G_{s_{1}^{2}...s_{q}}\hat{\eta}.
\]

**Conclusion**

In this article we have given an alternative derivation of the inversion of group-based phylogenetic models. Primarily our method relies on the remarkable intertwining relation between branching events and Markov evolution (4), and the resulting simplified expression of phylogenetic tensors given in (5). From there we took a representation theoretic approach concentrating on the structure of tensor indices.

**Competing interests**

The authors declare that they have no competing interests.

**Authors’ contributions**

JGS conducted much of the theoretical work presented, and was primarily responsible for drafting of the paper. PDJ contributed to the theoretical work and assisted in editing of the paper. BRH contributed to the theoretical work and contributed substantially to the writing of the paper. All authors read and approved the final manuscript.

**Authors’ information**

Jeremy G. Sumner is a ARC Research Fellow, Peter D. Jarvis is a Alexander von Humboldt Fellow and Barbara R. Holland is a ARC Future Fellow.

**Acknowledgements**

This research was supported by Australian Research Council Discovery Grants DP0877447 (JGS and PDJ) and FT101000381 (BRH).

Received: 15 April 2014 Accepted: 6 November 2014

Published online: 04 December 2014

**References**

1. Felsenstein J. Inferring Phylogenies. Sunderland: Sinauer Associates; 2004.
2. Semple C, Steel M. Phylogenetics. Oxford: Oxford University Press, 2003.
3. Hendy MD, Penny D. A framework for the quantitative study of evolutionary trees. Syst Zool 1989, 38:297–309.
4. Hendy MD: The relationship between simple evolutionary tree models and observable sequence data. Syst Zool 1989, 38:310–321.
5. Hendy MD, Penny D. Spectral analysis of phylogenetic data. J Class 1993, 101:1–20.
6. Steel M, Hendy M, Székely L, Erdos P: Spectral analysis and a closest tree method for genetic sequences. Appl Math Lett 1992, 5:663–67.
7. Székely LA, Erdos P, Steel M, Penny D: A Fourier inversion formula for evolutionary trees. Appl Math Lett 1993, 6:13–16.
8. Hendy MD, Penny D, Steel M: A discrete Fourier analysis for evolutionary trees. Proc Natl Acad Sci 1994, 91:3339–3343.
9. Székely LA, Steel MA, Erdős PL: Fourier calculus on evolutionary trees. Adv Appl Math 1993, 14:200–216.
10. Hendy MD, Charleston MA: Hadamard conjugation: a versatile tool for modelling nucleotide sequence evolution. New Zealand J Biol 1993, 31:231–257.
11. Holland BR, Penny D, Hendy MD: Outgroup misplacement and phylogenetic inaccuracy under a molecular clock – a simulation study. Syst Biol 2003, 52:229–238.
12. Hendy MD: A combinatorial description of the closest tree algorithm for finding evolutionary trees. Discrete Math 1991, 96:51–58.
13. Lento GM, Hickson RE, Chambers GK, Penny D: Use of spectral analysis to test hypotheses on the origin of pinnipeds. Mol Biol Evol 1995, 12:56–56.
14. Huber KT, Watson EE, Hendy MD: An algorithm for constructing local regions in a phylogenetic network. Mol Phylogen Evol 2001, 19:51–58.
15. Huber KT, Langston M, Penny D, Moulton V, Hendy M: Sprotrenet: a package for computing spectra and median networks. Appl Bioinform 2002, 1:2041–2059.
16. Schliep KP: Some applications of statistical phylogenetics. Ph.D. thesis. Massey University, 2009.
17. von Haeseler A, Churchill GA: Network models for sequence evolution. J Mol Evol 1993, 37:277–85.
18. Bryant D: Extending tree models to split networks. In Algebraic Statistics and Computational Biology. Edited by Pachter L, Sturmfels B. Cambridge: Cambridge University Press; 2005:297–368.
19. Bryant D: Hadamard phylogenetic methods and the -taxon process. Bull Math Biol 2009, 71:379–399.
20. Matsen FA, Steel M: Phylogenetic mixtures on a single tree can mimic a tree of another topology. Syst Biol 2007, 56:767–775.
21. Matsen FA, Mossel E, Steel M: Mixed-up trees: the structure of phylogenetic mixtures. Bull Math Biol 2008, 70:1115–1139.
22. Griffiths RC, Matzopoulos P: Ancestral inference from samples of DNA sequences with recombination. J Comput Biol 1996, 3:479–502.
23. Griffiths RC, Matzopoulos P: An ancestral recombination graph. In Progress in Population Genetics and Human Evolution, Volume 87 of IMA volumes in mathematics and its applications. Berlin: Springer Verlag; 1997:257–270.
24. Jin G, Naikule H, Snir S, Tuller T: Maximum likelihood of phylogenetic networks. Bioinformatics 2006, 21:2604–2611.
25. Stimmer K, Moulton V: Likelihood analysis of phylogenetic networks using directed graphical models. Mol Biol Evol 2008, 17:875–881.
26. Strimmer K, Wiuf C, Moulton V. Recombination analysis using directed graphical models. Mol Biol Evol 2001, 18:97–99.
27. Sumner JG, Holland BR, Jarvis PD. The algebra of the general Markov model on trees and networks. Bull Math Biol 2012, 74(4):858–880.
28. Sumner JG, Charleston MA, Jeromin LS, Jarvis PD. Markov invariants, plethysms, and phylogenetics. J Theor Biol 2008, 253:601–615.
29. Sumner JG, Jarvis PD. Markov invariants and the isotropy subgroup of a quartet tree. J Theor Biol 2009, 258:302–310.
30. Holland BR, Jarvis PD, Sumner JG. Low-parameter phylogenetic inference under the general markov model. Syst Biol 2013, 62:78–92.
31. Bashford JD, Jarvis PD, Sumner JG, Steel MA. \( U(1) \times U(1) \times U(1) \) symmetry of the Kimura 3ST model and phylogenetic branching processes. J Phys A Math Gen 2004, 37:L1–L9.
32. Sumner JG, Fernández-Sánchez J, Jarvis PD. Lie Markov models. J Theor Biol 2012, 298:16–31.
33. Sagan BE. The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions. Second Edition. New York: Springer, 2001. Graduate Texts in Mathematics.
34. Evans SN, Speed T. Invariants of some probability models used in phylogenetic inference. Ann Stat 1993, 21:355–377.
35. Sumner JG, Jarvis PD. Entanglement invariants and phylogenetic branching. J Math Biol 2005, 51:18–36.
36. Kimura M. Estimation of evolutionary distances between homologous nucleotide sequences. Proc Natl Acad Sci 1981, 78:1454–1458.

doi:10.1186/s12862-014-0236-6
Cite this article as: Sumner et al. A tensorial approach to the inversion of group-based phylogenetic models. BMC Evolutionary Biology 2014 14:236.