On the ideals of equivariant tree models

Jan Draisma · Jochen Kuttler

Abstract   We introduce equivariant tree models in algebraic statistics, which unify and generalise existing tree models such as the general Markov model, the strand symmetric model, and group-based models such as the Jukes–Cantor and Kimura models. We focus on the ideals of such models. We show how the ideals for general trees can be determined from the ideals for stars. A corollary of theoretical importance is that the ideal for a general tree is generated by the ideals of its flattenings at vertices. The main novelty is that our results yield generators of the full ideal rather than an ideal which only defines the model set-theoretically.

1 Set-up and theorems

In phylogenetics, tree models have been introduced to describe the evolution of a number of species from a distant common ancestor. Given suitably aligned strings of nucleotides of n species alive today, one assumes that the individual positions in

J. Draisma (✉)
Department of Mathematics and Computer Science, Technische Universiteit Eindhoven,
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
e-mail: j.draisma@tue.nl

J. Draisma
Centrum Wiskunde and Informatica, Amsterdam, The Netherlands

J. Kuttler
Department of Mathematical and Statistical Sciences, University of Alberta,
632 Central Academic Building, Edmonton, AB T6G 2G1, Canada
e-mail: jochen.kuttler@ualberta.ca

these strings have evolved independently and according to the same statistical process. Counting and averaging thus yields an empirical probability distribution on the set $\{A, C, G, T\}^n$. On the other hand, any hypothetical evolutionary tree whose leaves correspond to the n species gives rise to a parameterised family of probability distributions on $\{A, C, G, T\}^n$; see Sect. 2 for details. Here the parameters consist of an initial distribution and transition matrices along the edges of the tree. The challenge is to test whether the tree fits the data, that is, whether the empirical distribution lies in the family. One way to do this is the use of *phylogenetic identities*, equations that vanish identically on the family. This paper is concerned with constructing such identities for general trees from identities for smaller trees. The set-up below unifies and generalises existing tree models in algebraic statistics, while allowing for a clean and elegant treatment with methods from classical invariant theory. For more information on algebraic statistics and its applications see [7,15] and the many references there.

Remark 1.1 The term *phylogenetic invariants* is commonly used for phylogenetic identities. To avoid confusion with the invariants in classical invariant theory, we do not use this term.

First, recall that a tree T is a connected, undirected graph without circuits; all our trees are finite and have at least two vertices. The *valency* of a vertex of T is the number of edges containing it. A vertex of T is called a leaf if it has valency 1, and an internal vertex otherwise; if p, q are vertices, we write $p \sim q$ if there is an edge connecting them. We write vertex(T), leaf(T), internal(T) for the sets of vertices, leaves, and internal vertices of T, respectively. Stars are trees of diameter at most 2, and a centre of a star is a vertex at distance 1 to all other vertices—so if the star has more than 2 vertices, then its centre is unique. A *subtree* of a tree is a connected induced subgraph, and a *substar* is a subtree which itself is a star. So the map that sends a vertex p of T to the induced subgraph on p together with its neighbours is a bijection between vertex(T) and the set of substars of T, except when T consists of a single edge. In all that follows, we work over a ground field K that is algebraically closed and of characteristic zero.

Definition 1.2 A *spaced tree* T is given by the following data: First, a finite undirected tree, also denoted T; second, for every $p \in$ vertex(T) a finite-dimensional vector space V_p; third, a non-degenerate symmetric bilinear form $(\cdot \mid \cdot)_p$ on each V_p; and fourth, for every $p \in$ internal(T) a distinguished basis B_p of V_p which is orthonormal with respect to $(\cdot \mid \cdot)_p$.

The space V_p at a leaf p may also be given a distinguished basis B_p, orthonormal with respect to $(\cdot \mid \cdot)_p$, in which case p is called a *based leaf*. An internal vertex of T will also be called based. Any subtree of the underlying tree of T is regarded as a spaced tree with the data that it inherits from T.

Note that there is some redundancy in this definition: given the distinguished basis B_p at a based vertex p one could *define* $(\cdot \mid \cdot)_p$ by the requirement that B_p be orthonormal. We will leave out the subscript p from the bilinear form when it is obvious from the context. In many applications in algebraic statistics, symmetry is imposed on the algebraic model. This notion is captured well by the following notion of a G-spaced tree. Fix, once and for all, a finite group G.

Definition 1.3 A *G-spaced tree* (or *G-tree*, for short) is a spaced tree T in which the space V_p at every vertex p is a G-module, on which $(\cdot \mid \cdot)_p$ is G-invariant, and in which B_p is G-stable whenever p is a based vertex. Any subtree of the underlying tree of T is regarded as a G-spaced tree with the data that it inherits from T.

The objects that parameterise probability distributions in the algebraic model are what we propose to call representations of spaced trees.

Definition 1.4 Let T be a spaced tree. A *representation* of T is a collection of tensors $(A_{qp})_{q \sim p} \in V_q \otimes V_p$ along the edges of T with $A_{pq} = A_{qp}^t$, where $\cdot^t$ is the natural isomorphism $V_p \otimes V_q \rightarrow V_q \otimes V_p$. The space of all representations of T is denoted rep(T). A representation of T gives rise to a representation of any subtree of T.

For a G-spaced tree T, a *G-representation* or *equivariant representation* of T is a representation $(A_{qp})_{p \sim q}$ where each A_{qp} is a G-invariant element in the G-module $V_q \otimes V_p$. The space of such representations is denoted rep$_G(T)$. A G-representation of T gives rise to a G-representation of any subtree of T.

Using the bilinear form $(\cdot \mid \cdot)_p$ we may identify V_p with its dual V_p^*, and hence $V_q \otimes V_p$ with $V_q \otimes V_p^* \cong \operatorname{Hom}(V_p, V_q)$. Thus viewing A_{qp} as a linear map $V_p \rightarrow V_q$ and, similarly, A_{pq} as a linear map $V_q \rightarrow V_p$, the condition $A_{pq} = A_{qp}^t$ translates into $(A_{qp}v \mid w)_q = (v \mid A_{pq}w)_p$ for all $v \in V_p$ and $w \in V_q$. Put yet differently, if p and q are both based, then this says that the matrix of A_{pq} relative to the bases B_p and B_q is the transpose of the matrix of A_{qp}. In the applications to statistics, the spaces $\operatorname{Hom}(V_p, V_q)$ or the space of $|B_q| \times |B_p|$-matrices are perhaps more natural to work with than $V_q \otimes V_p$, as the elements of a representation correspond to *transition matrices* (Sect. 2). However, there are good reasons to work with $V_q \otimes V_p$; for instance, the correct action of $\operatorname{GL}(V_p) \times \operatorname{GL}(V_q)$ on the edge parameters turns out to be the natural action on $V_q \otimes V_p$ rather than that on $\operatorname{Hom}(V_q, V_p)$; see Lemma 4.1 and the proof of Proposition 2.3. Finally we note that if T is a G-tree, then by the invariance of the bilinear form $(\cdot \mid \cdot)_p$ the identifications above still make sense when passing to G-invariant elements: $(V_q \otimes V_p)^G = \operatorname{Hom}_G(V_p, V_q)$, etc.

Notice the slight discrepancy between our notion of representations and the notion in quiver representation theory, where—apart from the fact that the underlying graph is directed—the spaces V_p form part of the data comprising a representation.

A *T-tensor* is any element of $\bigotimes_{p \in \operatorname{leaf}(T)} V_p$, which space we will denote by $L(T)$ throughout the text. T-tensors correspond to marginal probability distributions in statistics; see Sect. 2. An important operation on spaced trees, representations, and T-tensors is $*$, defined as follows. Given k spaced trees $T_1, \dots, T_k$ whose vertex sets share a common based leaf q with common space V_q and common basis B_q but which trees are otherwise disjoint, we construct a new spaced tree $*_i T_i$ obtained by gluing the T_i along q, while the space at a vertex p of $*_i T_i$ coming from T_i is just the space attached to it in T_i, with the same distinguished bilinear form, and the same basis if p is based. Given representations $A_i \in \operatorname{rep}(T_i)$ for $i = 1, \dots, k$, we write $*_i A_i$ for the representation of $*_i T_i$ built up from the A_i. Now let Ψ_i be a T_i-tensor, for all i. Then we obtain a T-tensor by tensoring as follows:

$$*_i \Psi_i := \sum_{b \in B_q} \otimes_i (b \mid \Psi_i),$$

where we abuse the notation $(\cdot \mid \cdot)$ for the natural contraction

$$V_q \times \bigotimes_{p \in \text{leaf}(T_i)} V_p \to \bigotimes_{p \in \text{leaf}(T_i) \backslash \{q\}} V_p$$

determined by the bilinear form $(\cdot \mid \cdot)_q$. Notice that this $*$ operator is *not* a binary operator extended to several factors; nevertheless, when convenient, we will write $T_1 * \cdots * T_k$ for $*_i T_i$ and $\Psi_1 * \cdots * \Psi_i$ for $*_i \Psi_i$.

Now we come to a fundamental procedure that associates to any representation of a spaced tree T a T-tensor. Let $A \in \text{rep}(T)$. We proceed inductively. First, if T has a single edge pq, then $\Psi_T(A) := A_{qp}$, regarded as an element of $L(T) = V_q \otimes V_p$. If T has more than one edge, then let q be any internal vertex of T. We can then write $T = *_{p \sim q} T_p$, where T_p is the *branch of T around q containing p*, constructed by taking the connected component of $T - q$ (the graph obtained from T by removing q and all edges attached to q) containing p, and reattaching q to p.

The representation A induces representations A_p of the T_p, and by induction $\Psi_{T_p}(A_p)$ has been defined. We now set

$$\Psi_T(A) := *_{p \sim q} \Psi_{T_p}(A_p).$$

A straightforward proof by induction shows that this is independent of the choice of q and that this formula is also valid if q is actually a leaf. Now we can define the key objects of this paper.

Definition 1.5 Let T be a spaced tree. The *general Markov model* associated to T is the algebraic variety

$$\text{CV}_{\text{GM}}(T) := \overline{\{\Psi_T(A) \mid A \in \text{rep}(T)\}} \subseteq L(T),$$

where the closure is taken in the Zariski topology.

Similarly, for a G-spaced tree T, the *equivariant model* associated to T is the algebraic variety

$$\text{CV}_{\text{EM}}(T) := \overline{\{\Psi_T(A) \mid A \in \text{rep}_G(T)\}} \subseteq L(T).$$

Notice that a priori both the individual tensors $\Psi_T(A)$ and the varieties $\text{CV}_{\text{GM}}(T)$, $\text{CV}_{\text{EM}}(T)$ depend on the bases B_q at internal vertices q. This is only natural, as in applications these bases have an intrinsic meaning (Sect. 2). However, more can be said about this dependency (Lemma 4.1).

To streamline our discussion, we will consider CV_{GM} as the special case of CV_{EM} where G is trivial. An important goal in algebraic statistics is finding the ideal of all polynomials on the space $L(T)$ that vanish on $\text{CV}_{\text{EM}}(T)$. Our first result is a procedure for constructing these ideals from the ideals for substars of T.

Theorem 1.6 *For any G-spaced tree T, the ideal $I(\mathrm{CV}_{\mathrm{EM}}(T))$ can be expressed in the ideals $I(\mathrm{CV}_{\mathrm{EM}}(S))$ where S runs over the G-spaced substars of T with at least three leaves. In particular, for any spaced tree, the ideal of $I(\mathrm{CV}_{\mathrm{GM}}(T))$ can be expressed in the ideals $I(\mathrm{CV}_{\mathrm{GM}}(S))$ where S runs over the spaced substars of T with at least three leaves.*

This theorem is admittedly formulated somewhat vaguely. However, its proof in Sect. 4 gives rise to the explicit, recursive Algorithm 4.6 for determining $I(\mathrm{CV}_{\mathrm{EM}}(T))$ from the ideals $I(\mathrm{CV}_{\mathrm{EM}}(S))$; this justifies the present formulation.

We now present a variant of Theorem 1.6 which is perhaps less useful for actual computations, but which is of fundamental theoretical interest. This variant uses a second important operation on spaced trees and leaf tensors, namely, *flattening*. Fix any vertex q in a spaced tree T, and define an equivalence relation on $\mathrm{leaf}(T) \cup \{q\}$ by $p \cong r$ if and only if either $p = q = r$ or $p, r \neq q$ lie in the same connected component of $T - q$. Construct a spaced star $\flat_q T$ as follows: First, the vertex set is the set $\mathrm{leaf}(T) \cup \{q\}/\cong$ of equivalence classes, and the class of q is attached to all other classes by an edge. To the class C we attach the space $V_C := \bigotimes_{p \in C} V_p$ equipped with the bilinear form inherited from the V_p, and if all $p \in C$ are based, then C is based with the tensor product of the bases B_p. This new spaced tree $\flat_q T$ is called the *flattening* of T at q. Note that we allow q to be a leaf of T, in which case $\flat_q T$ has a single edge. The space $L(T)$ of T-tensors is naturally identified with the space $L(\flat_q T)$ of $\flat_q T$-tensors, and expanding the definition of Ψ_T at q one readily finds that

$$\mathrm{CV}_{\mathrm{EM}}(T) \subseteq \mathrm{CV}_{\mathrm{EM}}(\flat_q T) \quad \text{for all } q.$$

Our second main result shows that this characterises $\mathrm{CV}_{\mathrm{EM}}(T)$.

Theorem 1.7 *For any G-spaced tree T we have*

$$I(\mathrm{CV}_{\mathrm{EM}}(T)) = \sum_{q \in \mathrm{vertex}(T)} I(\mathrm{CV}_{\mathrm{EM}}(\flat_q T)).$$

Remark 1.8 (1) If T has more than one edge, then it suffices to let q run over internal(T).
(2) To avoid confusion we stress that $\flat_q T$ is *not* a substar of T, unless T itself is a star with centre q, in which case $\flat_q T \cong T$.

Many special cases of our main results are known in the literature. In particular [1] contains set-theoretic versions of our theorems for the general Markov model, and poses Theorem 1.7 for the general Markov model as Conjecture 5. In [16] the ideals of equivariant models with G abelian and all V_p equal to the regular representation KG are determined, following ideas from [8]. An important observation that makes this feasible is that these varieties are toric (Sect. 5). Some more specific references to the literature may be found in Sect. 2, which explains the relevance of spaced trees and their representations to statistics. After that, in Sect. 3 we prove a key tool on multiplying varieties of matrices, which we then use in Sect. 4 to prove our main results. Finally, Sect. 5 contains a result on toricness of certain abelian equivariant models.

2 Relevance to statistics

In the applications of our results to algebraic statistics, the spaced tree T that we start with only has based vertices. Indeed, the bases B_p have some physical meaning. In phylogenetics, for instance, they are usually all equal to $\{A, C, G, T\}$, the building bricks for DNA. Furthermore, an internal vertex r is singled out as root, and the base field is $K := \mathbb{C} \supseteq \mathbb{R}$. An element of V_p which on the basis B_p has non-negative real coefficients that add up to 1 is regarded as a probability distribution on B_p; together they form the *probability simplex* $\Delta(V_p) \subseteq V_p$. A representation of T is called *stochastic* if all maps $A_{qp} : V_p \to V_q$ directed away from r satisfy $A_{qp}\Delta(V_p) \subseteq \Delta(V_q)$, which amounts to saying that the entries of A_{qp}, regarded as a matrix relative to the bases B_q and B_p, are real and non-negative and that A_{qp} has all column sums equal to 1. A root distribution $\pi \in \Delta(V_r)$ and a stochastic representation A of T determine a probability distribution on $\prod_{p \in \mathrm{vertex}(T)} B_p$ and, by taking marginals, a distribution on $\prod_{p \in \mathrm{leaf}(T)} B_p$, which can be thought of as an element $\Phi_T(A, \pi)$ of $\Delta(L(T))$. Write $T = T_1 * \cdots * T_k$ at r and let $A_1, \ldots, A_k$ be the induced representations on the T_i. Then the distribution is

$$\Phi_T(A, \pi) = \sum_{b \in B_r} (b \mid \pi) \left(b \mid \Psi_{T_1}(A_1)\right) \otimes \cdots \otimes \left(b \mid \Psi_{T_k}(A_k)\right),$$

which equals $\Psi_T(A')$, where $A' \in \mathrm{rep}(T)$ is the (non-stochastic) representation obtained from A by composing a single A_{pr} leading away from the root with the diagonal linear map $V_r \to V_r$ determined by $b \mapsto \pi(b)b$. We define the set

$$\mathrm{CV}_{\mathrm{GM}}(T, r) := \{\Phi_T(A, \pi) \mid \pi \in \Delta(V_r) \text{ and } A \in \mathrm{rep}(T) \text{ stochastic}\}.$$

A natural equivariant analogue of this for a G-tree T is

$$\mathrm{CV}_{\mathrm{EM}}(T, r) := \{\Phi_T(A, \pi) \mid \pi \in \Delta(V_r) \ G\text{-invariant and } A \in \mathrm{rep}_G(T) \text{ stochastic}\},$$

but as the following examples from phylogenetics show it also makes sense to allow for arbitrary root distributions rather than G-invariant ones; see below how to handle these.

Example 2.1 In all models below, the B_p are all equal to $\{A, C, G, T\}$ and are all equipped with the same permutation action of some G. Recall that the nucleotides fall into two classes of bases, according to their chemical structure: the purines Adenine and Guanine and the pyrimidines Cytosine and Thymine. This explains some of the choices in the following models. All of them are equivariant models in our sense. The labels of these models are those used in [15].

(1) In the *Jukes–Cantor model JC69* $G = \mathrm{Sym}(\{A, C, G, T\})$ (or the alternating group, which has exactly the same equivariant maps $V_p \to V_p$). One assumes a G-invariant root distribution—which in this case means that it is uniform.

(2) In the *Kimura model K*80 *G* is the dihedral group generated by (A, C, G, T) and (A, G). It is the group of symmetries of the following square.

$$
\begin{array}{ccc}
A & - & C \\
| & & | \\
T & - & G
\end{array}
$$

Again, the root distribution is taken *G*-invariant, which means uniform.

(3) In the *Kimura model K*81 *G* is the Klein 4-group and the root distribution is *G*-invariant (uniform).

(4) In the *strand-symmetric model CS*05 *G* generated by the transpositions (A, G) and (C, T) and the root distribution is *G*-invariant.

(5) In the *HKY*85*model G* is as in the strand-symmetric model, but one allows for non-*G*-invariant root distributions.

(6) In the *Felsenstein model F81 G* is the full symmetric (or alternating) group, and the root distribution arbitrary.

The ideals of all these models were determined in [4,16]. Moreover, [3] gives local equations at biologically meaningful points.

Remark 2.2 A similar construction of tree models appears in [2]. There the spaces at all vertices are required to be the same space W, and the tensors at the edges are allowed to vary in some fixed subspace $\widehat{W}$ of $W \otimes W$ consisting of *symmetric* tensors. Otherwise the construction of the model is the same. We should mention that not all models obtained in this manner fit within our framework. For instance, the model where $\widehat{W}$ is the entire space of symmetric tensors cannot be characterised as the set of *G*-invariant tensors in $W \otimes W$ for some group G acting on W. Our present approach does not apply to this setting.

Similar to the observations in [1], as a consequence of the construction of $\mathrm{CV_{EM}}(T)$, it is a closed cone (i.e. invariant under scalar multiplication in $L(T)$) and therefore uniquely defines a projective variety in $\mathbb{P}(L(T))$, denoted $\mathbb{P}(\mathrm{CV_{EM}}(T))$, and defined by the same ideal as $\mathrm{CV_{EM}}(T)$. Notice that because the elements of $\mathrm{CV_{EM}}(T, r)$ have coordinate sum equal to 1, $\mathrm{CV_{EM}}(T, r)$ actually maps injectively into $\mathbb{P}(L(T))$. The following proposition justifies our quest for the ideal $I(\mathrm{CV_{EM}}(T))$: it contains all homogeneous polynomials vanishing on the statistically meaningful set $\mathrm{CV_{EM}}(T, r)$.

Proposition 2.3 *Provided that all V_p are non-zero, the image of the set $\mathrm{CV_{EM}}(T, r)$ in $\mathbb{P}(L(T))$ is Zariski dense in the variety $\mathbb{P}(\mathrm{CV_{EM}}(T))$.*

Proof First, the set of root distributions on B_r is clearly Zariski-dense in the set of all (complex) $\pi \in V_r$ with $\sum_{b \in B_r} \pi(b) = 1$. Similarly, for a single edge pq pointing away from r, the stochastic matrices in $\mathrm{Hom}_G(V_p, V_q)$ are Zariski dense in the complex matrices in $\mathrm{Hom}_G(V_p, V_q)$ with column sums 1. This follows from an explicit parameterisation of such equivariant stochastic matrices A: for every $b \in B_p/G$ the b-th column of A varies in a certain (scaled probability) simplex of dimension $|B_q/G_b| - 1$, where G_b is the stabiliser of b in G. This simplex is dense in the subset of $V_q^{G_b}$ where the sum of the coordinates is 1.

Next we claim that for A in an open dense subset of $\mathrm{rep}_G(T)$ we can write $\Psi_T(A)$ as $s\Phi_T(A', \pi)$ for some $A' \in \mathrm{rep}_G(T)$ having column sums 1 and some π with $\sum_b \pi(b) = 1$. To see this, first fix a vertex q and take for every vertex $p \sim q$ a copy of the torus $(K^*)^{B_q/G}$, considered as the diagonal subgroup of $\mathrm{GL}(V_q)$ acting by multiplication by a scalar on the span of each G-orbit on B_q and hence centralising G on V_q. The copy for p acts on $\mathrm{Hom}(V_p, V_q)$ by $g(A) := gA$ and on $\mathrm{Hom}(V_q, V_p)$ by $g(A) := Ag$. Note that this latter action is *not* the natural one on $\mathrm{Hom}(V_q, V_p)$, in which g would be replaced by its inverse, but that this action *is* the natural one on $V_q \otimes V_p$; see also Lemma 4.1 and the remarks preceding it. In particular, this action maps representations to representations. A straightforward computation shows that the subtorus

$$H_q := \left\{ (c_{p,b})_{b \in B_q/G, p \sim q} \in \prod_{p \sim q} (K^*)^{B_q/G} \mid \prod_{p \sim q} c_{p,b} = 1 \text{ for all } b \in B_q \right\}$$

leaves Ψ_T invariant. Now let A be any G-representation of T such that all column sums of all A_{qp} directed away from r are non-zero; this is an open dense condition on A. Given any non-root vertex q, by acting with H_q we can achieve that the A_{qp} leading away from r have column sums 1, while the map A_{qp}, where p is the parent of q relative to r, may not. If we do this for all non-root vertices in a bottom-up manner, and finally also for r, then we achieve that all A_{qp} leading away from r have column sums 1, except for a single A_{pr}; note that we have not altered $\Psi_T(A)$ in this process. Denote the column sums of A_{pr} by $(\sigma_b)_{b \in B_r}$. Dividing column b of A_{pr} by σ_b gives a representation A' all of whose matrices leading away from r have column sums 1. Also, for A in an open dense subset, $\sum_b \sigma_b =: s$ is non-zero, and dividing σ by s gives a π adding up to 1 such that $s\Phi(A', \pi) = \Psi(A)$. This proves the claim, and hence the proposition.   □

As we saw in the examples above, one may want to allow arbitrary root distributions, which are not necessarily G-invariant. More generally, one might want to allow the root distribution to vary in a certain self-dual submodule of V_r, and this would require only minor changes in the discussion that follows—but here we concentrate on the situation where all elements of (the probability simplex in) V_r are allowed. We define the set

$$\mathrm{CV}_{\mathrm{EM}}(T, V_r) := \left\{ \Phi_T(A, \pi) \mid A \in \mathrm{rep}_G(T) \text{ stochastic and } \pi \in \Delta(V_r) \right\}.$$

One can elegantly describe $\mathrm{CV}_{\mathrm{EM}}(T, V_r)$ as follows. Let T' be the spaced tree obtained from T by connecting a new vertex r' to the root r of T and attaching to r' the G-module $V_{r'} := V_r$, endowed with the same bilinear form. Then $L(T') = V_{r'} \otimes L(T)$, and since $V_{r'} \cong V_{r'}^*$ we may think of $\mathrm{CV}_{\mathrm{EM}}(T')$ as a subset of $\mathrm{Hom}(V_{r'}, L(T))$. Since it also consists of G-fixed points, and since this identification is G-equivariant, it is a subset of $\mathrm{Hom}_G(V_{r'}, L(T))$.

Proposition 2.4 *The image of* $\mathrm{CV}_{\mathrm{EM}}(T, V_r)$ *in* $\mathbb{P}(L(T))$ *is a Zariski dense subset in the projective variety associated to the cone*

$$\overline{\mathrm{CV}_{\mathrm{EM}}(T')V_{r'}},$$

where we regard $\mathrm{CV}_{\mathrm{EM}}(T')$ *as a subset of* $\mathrm{Hom}_G(V_{r'}, L(T))$. *The ideal of this cone can be determined from the ideal of* $\mathrm{CV}_{\mathrm{EM}}(T')$.

Proof We have

$$\Phi_T(A, \pi) = \Psi_{T'}(A')\pi,$$

where A' is obtained from A by putting the identity I_{V_r} along the edge rr'. This shows that $\Phi_T(A, \pi)$ is contained in $\mathrm{CV}_{\mathrm{EM}}(T')V_{r'}$. For the converse we reason as before: for A' in an open dense subset of $\mathrm{rep}_G(T')$ we can write $\Psi_{T'}(A')$ as $\Psi_{T'}(A)$ where all A_{qp} directed away from r have column sums equal to 1 except possibly for $A_{r'r}$. We have

$$
\begin{aligned}
\Psi_{T'}(A)\pi &= \sum_{b \in B_r} (A_{r'r}b|\pi)(b \mid \Psi_{T_1}(A_1)) \otimes \cdots \otimes (b \mid \Psi_{T_m}(A_m)) \\
&= \sum_{b \in B_r} (b|A_{rr'}\pi)(b \mid \Psi_{T_1}(A_1)) \otimes \cdots \otimes (b \mid \Psi_{T_m}(A_m)) \\
&= \Phi_T(A, A_{rr'}\pi) = s\Phi_T(A, \pi'),
\end{aligned}
$$

where s is taken such that $\pi' := s^{-1}A_{rr'}\pi$ has $\sum_b \pi'(b) = 1$.

The proof of the last statement is deferred to the end of Sect. 4. $\qquad\square$

3 Multiplying varieties of matrices

In this section, we derive a key tool that will be used in the proofs of our results. As before let K be an algebraically closed field of characteristic 0, and let t be a natural number. For $\mathbf{k}, \mathbf{l} \in \mathbb{N}^t$ let $M_{\mathbf{k},\mathbf{l}}$ denote the space $M_{k_1,l_1} \times \cdots \times M_{k_t,l_t}$, where $M_{k,l}$ is the space of $k \times l$-matrices over K. To formulate and prove our results in their full strength, it is convenient to use some notions from the language of schemes, for which we refer to [11]. The main point here is that we do not require ideals to be radical.

Recall that if X is an affine variety, then a closed subscheme S of X is given by an ideal I of the ring $K[X]$ of regular functions on X: the underlying closed subset of X is the set of zeros of I, and the K-algebra associated to S is $K[X]/I$. We write $I(S)$ for the ideal of the subscheme S. If X is a variety on which some group Γ acts, then S is called a Γ-subscheme if and only if $I(S) \subseteq K[X]$ is Γ-stable. Finally, if $f : X \to Y$ is a map between varieties, inducing the pull back homomorphism $f^\sharp : K[Y] \to K[X]$, and if $S \subseteq X$ is a subscheme, then the *image scheme* of S is defined as the scheme theoretic closure of $f(S)$, i.e. the subscheme of Y defined by the ideal $(f^\sharp)^{-1}(I(S))$. By slight abuse of notation it is usually denoted as $f(S)$. It is clear that if f is Γ-equivariant for some group Γ acting on X and Y, then the image of a Γ-subscheme is again a Γ-subscheme. Also notice that if S is a subvariety, i.e., if $I(S)$ is radical, then so is $f(S)$—it is precisely the Zariski closure of the set-theoretic image of S under f.

We now specialise to multiplying schemes of matrices. We write $\mu \colon M_{\mathbf{k},\mathbf{l}} \times M_{\mathbf{l},\mathbf{m}} \to M_{\mathbf{k},\mathbf{m}}$ for the multiplication and $\mu^{\sharp}$ for the co-multiplication, $\mu^{\sharp}(f)(A, B) = f(A \cdot B)$. Given two subschemes $V \subseteq M_{\mathbf{k},\mathbf{l}}$ and $W \subseteq M_{\mathbf{l},\mathbf{m}}$, put

$$V \cdot W := \mu(V, W).$$

If V, W are subvarieties this is just

$$\overline{\{AB \mid A \in V, B \in W\}},$$

where $AB := (A_1 B_1, \ldots, A_t B_t)$. In general, the underlying topological space is still the closure of the set $\mu(V, W)$, but the ideal is $(\mu^{\sharp})^{-1}(I(V \times W))$.

The operation $\cdot$ is associative in the sense that, given a third subscheme $U \subseteq M_{\mathbf{j},\mathbf{k}}$, one has

$$U \cdot (V \cdot W) = (U \cdot V) \cdot W,$$

we therefore simply write $U \cdot V \cdot W$.

Let $\mathrm{GL}_{\mathbf{l}} = \mathrm{GL}_{l_1} \times \mathrm{GL}_{l_2} \times \cdots \times \mathrm{GL}_{l_t} \subseteq M_{\mathbf{l}}$. Frequently the subschemes we are interested in will be invariant by left- or right-multiplication by $\mathrm{GL}_{\mathbf{l}}$. In this context it is worth mentioning that a subscheme $V \subseteq M_{\mathbf{k},\mathbf{l}}$ is a $\mathrm{GL}_{\mathbf{l}}$-subscheme, i.e., stable by right-multiplication, if and only if $V \cdot M_{\mathbf{l},\mathbf{l}} = V$. This follows from the fact that for any K-algebra A and any nonconstant $f \in K[x_1, \ldots, x_m]$, $\mathrm{Spec}\, A[x_1, \ldots, x_m]_f$ is dense in $\mathrm{Spec}\, A[x_1, \ldots, x_m]$, and thus $V \cdot \mathrm{GL}_{\mathbf{l}}$ is (scheme theoretically) dense in $V \cdot M_{\mathbf{l}}$.

Theorem 3.1 *Let* $\mathbf{k}, \mathbf{l}, \mathbf{m} \in \mathbb{N}^t$ *and let* V *and* W *be subschemes of* $M_{\mathbf{k},\mathbf{l}}$ *and* $M_{\mathbf{l},\mathbf{m}}$, *respectively. Then we have*

$$I(V \cdot M_{\mathbf{l},\mathbf{l}} \cdot W) = I(V \cdot M_{\mathbf{l},\mathbf{m}}) + I(M_{\mathbf{k},\mathbf{l}} \cdot W). \tag{1}$$

In the following proof we will use the First Fundamental Theorem in invariant theory, which describes the invariant polynomials of GL_n on a direct sum of copies of K^n and the dual space $(K^n)^*$. This theorem is due to Weyl [17]. Another tool from representation theory is the Reynolds operator: when a reductive group acts rationally on a vector space, then the Reynolds operator ρ is the projection onto the invariant vectors with kernel the direct sum of all non-trivial irreducible submodules. Modern treatments on invariant theory are [5,10,12].

Proof The inclusion $\supseteq$ is obvious. To prove the opposite inclusion we first replace V by $V \cdot M_{\mathbf{l},\mathbf{l}}$ and W by $M_{\mathbf{l},\mathbf{l}} \cdot W$—this clearly keeps invariant both sides of (1), and moreover turns V and W into $\mathrm{GL}_{\mathbf{l}}$-subschemes, where $\mathrm{GL}_{\mathbf{l}} := \mathrm{GL}_{l_1} \times \cdots \times \mathrm{GL}_{l_t}$, which acts on $M_{\mathbf{k},\mathbf{l}} \times M_{\mathbf{l},\mathbf{m}}$ by

$$(g_1, \ldots, g_t)((A_1, \ldots, A_t), (B_1, \ldots, B_t)) = ((A_1 g_1^{-1}, \ldots, A_t g_t^{-1}), (g_1 B_1, \ldots, g_t B_t))$$

Let $f \in I(V \cdot W)$, which now equals the left-hand side of (1). Define $h \in K[M_{\mathbf{k},\mathbf{l}} \times M_{\mathbf{l},\mathbf{m}}]$ by $h = \mu^{\sharp}(f)$, so that $h(A, B) = f(AB)$. Then h is invariant with respect to the action of $GL_{\mathbf{l}}$, and moreover h is in the ideal of $V \times W$. This latter fact implies that

$$h \in I(V \times M_{\mathbf{l},\mathbf{m}}) + I(M_{\mathbf{k},\mathbf{l}} \times W);$$

split $h = h_1 + h_2$ accordingly. Applying the Reynolds operator $\rho : K[M_{\mathbf{k},\mathbf{l}} \times M_{\mathbf{l},\mathbf{m}}] \to K[M_{\mathbf{k},\mathbf{l}} \times M_{\mathbf{l},\mathbf{m}}]^{GL_{\mathbf{l}}}$ yields $h = \rho(h_1) + \rho(h_2)$. By $GL_{\mathbf{l}}$-invariance of V and W, $\rho(h_1)$ and $\rho(h_2)$ still are elements of $I(V \times M_{\mathbf{l},\mathbf{m}})$ and $I(M_{\mathbf{k},\mathbf{l}} \times W)$, respectively. Furthermore, $\rho(h_1)$ and $\rho(h_2)$ lie in

$$K[M_{\mathbf{k},\mathbf{l}} \times M_{\mathbf{l},\mathbf{m}}]^{GL_{\mathbf{l}}} = K[M_{k_1,l_1} \times M_{l_1,m_1}]^{GL_{l_1}} \otimes \cdots \otimes K[M_{k_t,l_t} \times M_{l_t,m_t}]^{GL_{l_t}}.$$

By the First Fundamental Theorem for GL_{l_i} applied to k_i covectors and m_i vectors the pullback of multiplication $M_{k_i,l_i} \times M_{l_i,m_i} \to M_{k_i,m_i}$ is a surjective homomorphism

$$K[M_{k_i,m_i}] \to K[M_{k_i,l_i} \times M_{l_i,m_i}]^{GL_{l_i}}$$

for all $i = 1, \ldots, t$. Hence the pullback of multiplication $M_{\mathbf{k},\mathbf{l}} \times M_{\mathbf{l},\mathbf{m}} \to M_{\mathbf{k},\mathbf{m}}$ is a surjective homomorphism

$$K[M_{\mathbf{k},\mathbf{m}}] \to K[M_{\mathbf{k},\mathbf{l}} \times M_{\mathbf{l},\mathbf{m}}]^{GL_{\mathbf{l}}};$$

let $\rho(h_1), \rho(h_2)$ lift under this surjection to $\bar{h}_1, \bar{h}_2 \in K[M_{\mathbf{k},\mathbf{m}}]$, respectively. Note that $\bar{h}_1, \bar{h}_2$ are not unique if $l_i < \min\{k_i, m_i\}$ for some i, but this is irrelevant here. We now have $\bar{h}_1 \in I(V \cdot M_{\mathbf{l},\mathbf{m}}), \bar{h}_2 \in I(M_{\mathbf{k},\mathbf{l}} \cdot W)$. Moreover, restricted to the image $M_{\mathbf{k},\mathbf{l}} \cdot M_{\mathbf{l},\mathbf{m}}$ we have $f = \bar{h}_1 + \bar{h}_2$, since

$$f(AB) = h(A, B) = \rho(h_1)(A, B) + \rho(h_2)(A, B) = \bar{h}_1(AB) + \bar{h}_2(AB)$$

for all $A \in M_{\mathbf{k},\mathbf{l}}, B \in M_{\mathbf{l},\mathbf{m}}$. Hence $f' := f - (\bar{h}_1 + \bar{h}_2)$ vanishes on $M_{\mathbf{k},\mathbf{l}} \cdot M_{\mathbf{l},\mathbf{m}}$, which contains both $V \cdot M_{\mathbf{l},\mathbf{m}}$ and $M_{\mathbf{k},\mathbf{l}} \cdot W$. But then f' lies in both $I(V \cdot M_{\mathbf{l},\mathbf{m}})$ and $I(M_{\mathbf{k},\mathbf{l}} \cdot W)$, and hence $f \in I(V \cdot M_{\mathbf{l},\mathbf{m}}) + I(M_{\mathbf{k},\mathbf{l}} \cdot W)$ as claimed. □

We will use Theorem 3.1 to describe the ideal of $V \cdot W$ explicitly from the ideals of V and W. Before we can do this we need one more tool. Suppose $\alpha : X \times K^n \to Y$ is a morphism where X and Y are affine varieties over K, and suppose S is a closed subscheme of Y defined by an ideal I. Then there exists a uniquely determined subscheme S' of X such that $\alpha(S' \times K^n) \subseteq S$ and such that S' is maximal with this property. Scheme-theoretically, S' is equal to $S' = \bigcap_{v \in K^n} i_v^{-1}(\alpha^{-1}(S))$ where for $v \in K^n$, i_v is the inclusion $X \to X \times \{v\} \subseteq X \times K^n$. The ideal of S' is determined as follows: Let I' be the ideal in $K[X]$ generated by all functions of the form $f'_v = \alpha^{\sharp}(f)(\cdot, v) = f(\alpha(\cdot, v))$ where $f \in I$ and $v \in K^n$ is an arbitrary (closed) point. In other words, $f'_v(x) = f(\alpha(x, v))$ for $x \in X$.

Since $K[X \times K^n] = K[X] \otimes_K K[x_1, x_2, \ldots, x_n]$, for any $f \in K[Y]$, we may write $\alpha^\sharp(f)$ uniquely as

$$\alpha^\sharp(f) = \sum_i h_i \otimes m_i \tag{2}$$

where the m_i are some linearly independent monomials in $K[x_1, x_2, \ldots, x_n]$, and $h_i \in K[X]$. It is now elementary to check that I' is generated by all $h_i \in K[X]$ that appear in such an expression (2) as f runs through I. Indeed, the ideal generated by f'_v where v ranges over K^n is precisely the ideal generated by all $h_1, h_2, \ldots, h_s$. This is easily seen by picking s points $v_1, v_2, \ldots, v_s$ in K^n such that $\det[m_i(v_j)] \neq 0$, which is possible as K is infinite and the m_i are supposed to be linearly independent. This observation is important because it shows how to compute a finite list of generators for I' out of finitely many generators for I: if $f_1, f_2, \ldots, f_m$ generate I, then the (finite) collection of all h_is appearing in one of the $\alpha^\sharp(f_j)$s generates I'.

We will apply this construction to the case where $X = M_{\mathbf{k},\mathbf{l}}$, $K^n = M_{\mathbf{l},\mathbf{m}}$ and $\alpha = \mu$ equal to matrix multiplication. Then for $V \subseteq M_{\mathbf{k},\mathbf{m}}$, the ideal $I(V')$ is generated by all functions on $M_{\mathbf{k},\mathbf{l}}$ of the form $f(xB)$ where $B \in M_{\mathbf{l},\mathbf{m}}$ is arbitrary, and $f \in I(V)$. In fact, thinking of the entries of B as variables, we may expand $f(xB)$ as a polynomial in the entries of B; the coefficients are then the required elements of the ideal of V'. See also Example 5.4.

The following corollary is crucial for explicit computations; it is a slight generalisation of [1, Lemma 12], which gives the same equations for $V \cdot M_{\mathbf{l},\mathbf{m}}$. Although there the result is stated only for subvarieties, their proof should also go through essentially unchanged.

Corollary 3.2 *Let* $\mathbf{k}, \mathbf{l}, \mathbf{m} \in \mathbb{N}^t$ *and let* V *be a subscheme of* $M_{\mathbf{k},\mathbf{l}}$ *with* $V = V \cdot M_{\mathbf{l},\mathbf{l}}$. *Define the scheme* $R_{\mathbf{l}} \subseteq M_{\mathbf{k},\mathbf{m}}$ *by the ideal generated by all* $(l_i + 1)$*-minors of the* i*-th component, for all* $i = 1, \ldots, t$. *Then we have*

$$I(V \cdot M_{\mathbf{l},\mathbf{m}}) = I(V') + I(R_{\mathbf{l}}) \tag{3}$$

where V' *is defined as in the preceding paragraph as the unique maximal subscheme of* $M_{\mathbf{k},\mathbf{m}}$ *such that* $V' \cdot M_{\mathbf{m},\mathbf{l}} \subseteq V$.

It is well known that $M_{\mathbf{k},\mathbf{l}} \cdot M_{\mathbf{l},\mathbf{m}} = R_{\mathbf{l}}$ as schemes, and in particular that the ideal of $R_{\mathbf{l}}$ is radical.

Proof The inclusion $\supseteq$ follows from $V \cdot M_{\mathbf{l},\mathbf{m}} \subseteq V' \cap R_{\mathbf{l}}$: First, $V \cdot M_{\mathbf{l},\mathbf{m}} \subseteq R_{\mathbf{l}} = M_{\mathbf{k},\mathbf{l}} \cdot M_{\mathbf{l},\mathbf{m}}$ is clear. Second, $V \cdot M_{\mathbf{l},\mathbf{l}} = V$ implies that $V \cdot M_{\mathbf{l},\mathbf{m}} \cdot M_{\mathbf{m},\mathbf{l}} \subseteq V$, i.e. $V \cdot M_{\mathbf{l},\mathbf{m}} \subseteq V'$.

For the opposite inclusion, set $W := M_{\mathbf{m},\mathbf{l}} \cdot M_{\mathbf{l},\mathbf{m}}$, and apply Theorem 3.1 with $(\mathbf{k}, \mathbf{l}, \mathbf{m}, V, W)$ replaced by $(\mathbf{k}, \mathbf{m}, \mathbf{m}, V', W)$. Indeed, $V' \cdot M_{\mathbf{m},\mathbf{l}} \subseteq V$ by definition of V', so that the left-hand side of (3) is contained in $I(V' \cdot M_{\mathbf{m},\mathbf{l}} \cdot M_{\mathbf{l},\mathbf{m}})$, which is the left-hand side of (1) with (V, W) replaced by (V', W). With this substitution the right-hand side of (1) reads

$$I(V' \cdot M_{\mathbf{m},\mathbf{m}}) + I(M_{\mathbf{k},\mathbf{m}} \cdot W)$$

which, as $M_{\mathbf{k},\mathbf{m}} \cdot W = R_{\mathbf{l}}$, equals the right-hand side of (3).   □

The corollary is the reason why we had to use subschemes instead of subvarieties: in general, V' is not a variety even if V is, so the ideal $\sqrt{I(V')}$ of functions vanishing on the closed points of V' may be larger than $I(V')$, and hard to compute. However, the corollary shows that to compute $I(V \cdot M_{l,m})$ only the ideal $I(V')$ is needed, and for this ideal generators can be found as described above.

We will apply Theorem 3.1 and its corollary in the following setting: Let V be a representation of G, Ω the set of all irreducible characters, and for $\omega \in \Omega$ denote by M_ω a fixed irreducible representation of type ω. Then $V \cong \bigoplus_{\omega \in \Omega} V[\omega]$, where $V[\omega]$ is the sum of all submodules of V isomorphic to M_ω. Moreover, $V[\omega] \cong M_\omega \otimes \mathrm{Hom}_G(M_\omega, V) \cong M_\omega \otimes K^{m(\omega,V)}$ with $m(\omega, V)$ the multiplicity of M_ω in V. In particular, if W is another representation of G, then the space of equivariant maps $V \to W$ is

$$\mathrm{Hom}_G(V, W) \cong \bigoplus_{\omega \in \Omega} \mathrm{Hom}\left(K^{m(\omega,V)}, K^{m(\omega,W)}\right). \tag{4}$$

The varieties we are interested in are subvarieties of $\mathrm{Hom}_G(V, W)$ stable by multiplication with $\mathrm{End}_G(V)$ or $\mathrm{End}_G(W)$ where V, W are some representations of G. So let U, V, W be three representations of G, and suppose $S \subseteq \mathrm{Hom}_G(U, V)$ and $T \subseteq \mathrm{Hom}_G(V, W)$ are subvarieties or subschemes. To apply Theorem 3.1, we may identify $\mathrm{Hom}_G(V, W)$ with $M_{k,l}$, $\mathrm{Hom}_G(U, V)$ with $M_{l,m}$, and $\mathrm{End}_G(V)$ with $M_{l,l}$, by putting $l_i = m(\omega_i, V)$, $k_i = m(\omega_i, W)$, and $m_i = m(\omega_i, U)$, respectively, where $\Omega = \{\omega_1, \omega_2, \ldots, \omega_t\}$.

With these identifications in place, the ideal of $T \cdot \mathrm{End}_G(V) \cdot S$ is equal to $I(T \cdot M_{l,m}) + I(M_{k,l} \cdot S)$. Similarly, if T is stable by right-multiplication with $\mathrm{End}_G(V, V)$, then the ideal of $T \cdot \mathrm{Hom}_G(U, V)$ may be computed using Corollary 3.2. We will see several applications of this in the next section.

4 Proofs of the main results

Before proving our main results, we investigate how Ψ_T and $\mathrm{CV}_{\mathrm{EM}}(T)$ behave under base changes. Thus let T be a G-tree and write GL_T for the product $\prod_{p \in \mathrm{vertex}(T)} \mathrm{GL}(V_p)^G$. On the one hand, this group acts on $\mathrm{rep}_G(T)$ by the action of $\mathrm{GL}(V_q)^G \times \mathrm{GL}(V_p)^G$ on $(V_q \otimes V_p)^G$. We already encountered a special case of this action in the proof of Proposition 2.3. On the other hand, given $h \in \mathrm{GL}_T$ one can define a new G-spaced tree hT as follows: the underlying tree of hT is the same as that of T and the space V_p at each vertex is also the same as that of T, but the bilinear form $(\cdot|\cdot)_p'$ is determined by

$$(h_p u | h_p v)_p' := (u|v)_p \quad \text{for } u, v \in V_p,$$

where the latter bilinear form is the one assigned to p in T. Finally, a vertex p is based in hT if and only if it is based in T, and then the basis associated to it in hT is $B_p' := hB_p$, where B_p is the distinguished basis of V_p in T. A representation A

of T gives a representation of hT, also denoted A, by simply taking the same tensors $A_{qp} \in V_q \otimes V_p$ along the edges.

Lemma 4.1 *In the setting above we have $\Psi_{hT}(A) = h\Psi_T(h^{-1}A)$.*

Proof If T has only two vertices $p \sim q$, then this just the obvious equality $A_{qp} = (h_q, h_p)(h_q^{-1}, h_p^{-1})A_{qp}$. If T has more than two vertices, we pick any internal vertex q of T and split $T = *_i T_i$ and $A = *_i A_i$ at q. Assuming the result for all T_i we find

$$
\begin{aligned}
\Psi_{hT}(A) &= \sum_{b \in B'_q} \otimes_i \left(b | \Psi_{hT_i}(A_i) \right)' \\
&= \sum_{b \in B'_q} \otimes_i \left(b | h\Psi_{T_i}(h^{-1}A_i) \right)' \\
&= \sum_{b \in B_q} \otimes_i \left(h_q b | h\Psi_{T_i}(h^{-1}A_i) \right)' \\
&= h \sum_{b \in B_q} \otimes_i \left(b | \Psi_{T_i}(h^{-1}A) \right) \\
&= h\Psi_{hT} \left(h^{-1}A \right).
\end{aligned}
$$

$\square$

In particular, this lemma implies that $\mathrm{CV}_{\mathrm{EM}}(T) = \mathrm{CV}_{\mathrm{EM}}(hT)$. For $G = \{1\}$ we note that if T' is any spaced tree with the same underlying tree as T and the same G-modules V_p at the vertices, but different bilinear forms and different (orthonormal) bases, then there exists an $h \in \mathrm{GL}_T$ with $hT = T'$. In this sense the variety $\mathrm{CV}_{\mathrm{GM}}(T)$ does not depend on the chosen bases and forms, as long as they are compatible.

Remark 4.2 A stronger basis-independency holds at vertices of valency two. There the operation $*$ boils down to matrix multiplication, or composition of linear maps, and this will enable us to apply Theorem 3.1. Indeed, let U, V, W be vector spaces equipped with non-degenerate symmetric bilinear forms and let $\Psi_1 \in W \otimes V$ and $\Psi_2 \in V \otimes U$ be arbitrary. Let B be any orthonormal basis of V. We claim that the element

$$
\sum_{b \in B} (b|\Psi_1) \otimes (b|\Psi_2) \in W \otimes U
$$

does not depend on B, and under the identification $U \cong U^*$ coincides with the linear map $U \to W$ which is the composition of Ψ_1 and Ψ_2, considered as linear maps under the identifications $W \otimes V = W \otimes V^* = \mathrm{Hom}(V, W)$ and $V \otimes U = V \otimes U^* = \mathrm{Hom}(U, V)$. It suffices to verify this for rank-one tensors $\Psi_1 = w \otimes v$ and $\Psi_2 = v' \otimes u$. The expression above is then

$$
\left[\sum_{b \in B} (b|v)(b|v') \right] w \otimes u.
$$

By the orthonormality of B this reduces to $(v|v')w \otimes u$, as claimed.

Thus, in hindsight, we could have left out the orthonormal bases at vertices of valency 2 in the definition of (G-)spaced trees, and defined the operation $*$ as composition of linear maps. We have not done so to keep the treatment of internal vertices uniform.

Next we observe that the map Ψ_T defined in Sect. 1 behaves well with respect to the group action. Let T be a G-spaced tree. Note that G acts naturally on $\text{rep}(T)$ by its action on each tensor product $V_p \otimes V_q$ with $p \sim q$.

Lemma 4.3 *The map $\Psi_T \colon \text{rep}(T) \to L(T)$ is G-equivariant.*

Proof If T has exactly two vertices the assertion is immediate. Otherwise, let q be an inner vertex of T and split T around q into $T_1, \ldots, T_k$. The T_i are G-trees in a natural way. By induction, we may assume that Ψ_{T_i} is an equivariant map. Then

$$
\begin{aligned}
\Psi_T(gA) &= \sum_{b \in B_q} \left(b \mid \Psi_{T_1}(gA_1)\right) \otimes \cdots \otimes \left(b \mid \Psi_{T_k}(gA_k)\right) \\
&= \sum_{b \in B_q} \left(b \mid g\Psi_{T_1}(A_1)\right) \otimes \cdots \otimes \left(b \mid g\Psi_{T_k}(A_k)\right) \\
&= g \sum_{b \in B_q} \left(g^{-1}b \mid \Psi_{T_1}(A_1)\right) \otimes \cdots \otimes \left(g^{-1}b \mid g\Psi_{T_k}(A_k)\right) \\
&= g\Psi_T(A),
\end{aligned}
$$

where the second equality follows from the G-invariance of $(\cdot \mid \cdot)_q$ and the last equality follows from the fact that g^{-1} permutes B_q. □

This lemma implies that $\text{CV}_{\text{EM}}(T) \subseteq L(T)^G$. In what follows we focus on the ideal of $\text{CV}_{\text{EM}}(T)$ inside $K[L(T)^G]$. To obtain the ideal inside $K[L(T)]$, one just adds linear equations cutting out $L(T)^G$ from $L(T)$.

Remark 4.4 For $k = 2$ the computation in the proof of the lemma can be replaced by the following argument, using the notation of Remark 4.2. If U, V, W are G-modules and V carries a G-invariant symmetric bilinear form, then the unique bilinear map $W \otimes V \times V \otimes U \to W \otimes U$ sending $(w \otimes v, v' \otimes u)$ to $(v|v')w \otimes u$ is G-equivariant. So at vertices of valency 2 it is not crucial that G permutes the basis.

With these preparations, we are now ready to prove our first main result.

Proof of Theorem 1.6 Let T be a G-spaced tree. We recursively express the ideal of $\text{CV}_{\text{EM}}(T)$ into the ideals of $\text{CV}_{\text{EM}}(S)$ for substars S of T with at least three leaves, as follows. First, if T has only two vertices $p \sim q$, then $\text{CV}_{\text{EM}}(T) = (V_p \otimes V_q)^G = L(T)^G$ and we are done. Second, if T is itself a star with at least three leaves, then we are also done. Third, suppose that T contains a vertex q of valency 2, and split T accordingly as $T = T_1 * T_2$, so that

$$
\text{CV}_{\text{EM}}(T) = \overline{\{\Psi_1 * \Psi_2 \mid \Psi_i \in \text{CV}_{\text{EM}}(T_i) \text{ for } i = 1, 2\}}.
$$

Let L_1 be the space $\bigotimes_{p \in \mathrm{leaf}(T_1) \setminus \{q\}} V_p$ and $L_2 = \bigotimes_{p \in \mathrm{leaf}(T_2) \setminus q} V_p$. Of course L_1, L_2 are naturally G-representations. Now the map $(\Psi_1, \Psi_2) \mapsto \Psi_1 * \Psi_2$ from $L(T_1)^G \times L(T_2)^G$ to $L(T)^G$ is just matrix multiplication if we identify $L(T_1)^G$ with $\mathrm{Hom}_G(V_p, L_1)$ and $L(T_2)^G$ with $\mathrm{Hom}_G(L_2, V_p)$; see Remark 4.2.

We want to apply Theorem 3.1. Recall the definition of ω_i and $m(\omega_i, V)$ from Sect. 3. Now define $\mathbf{k}, \mathbf{l}, \mathbf{m}$ by $k_i := m(\omega_i, L_1), l_i := m(\omega_i, V_q)$ and $m_i := m(\omega_i, L_2)$. Then $\mathrm{Hom}_G(V_q, L_1) = M_{\mathbf{k},\mathbf{l}}$, $\mathrm{Hom}_G(L_2, V_q) = M_{\mathbf{l},\mathbf{m}}$, and $\mathrm{End}_G(V_q) = M_{\mathbf{l},\mathbf{l}}$; see Eq. (4). Notice that $V := \mathrm{CV}_{\mathrm{EM}}(T_1)$ (resp. $W := \mathrm{CV}_{\mathrm{EM}}(T_2)$) are stable under right- (resp. left-) multiplication with $M_{\mathbf{l},\mathbf{l}}$, and $\mathrm{CV}_{\mathrm{EM}}(T) = V \cdot W = V \cdot M_{\mathbf{l},\mathbf{l}} \cdot W$. Thus Theorem 3.1 applies and we deduce that

$$I(\mathrm{CV}_{\mathrm{EM}}(T)) = I(\mathrm{CV}_{\mathrm{EM}}(T_1) * \mathrm{CV}_{\mathrm{EM}}(\flat_q T_2)) + I(\mathrm{CV}_{\mathrm{EM}}(\flat_q T_1) * \mathrm{CV}_{\mathrm{EM}}(T_2)),$$

where $\mathrm{CV}_{\mathrm{EM}}(\flat_q T_i) \cong \mathrm{Hom}_G(L_i, V_q) \cong \mathrm{Hom}_G(V_q, L_i) \cong (L_i \otimes V_q)^G$ because G acts preserving the form. Recursively, we may assume that the ideals of $\mathrm{CV}_{\mathrm{EM}}(T_1)$ and $\mathrm{CV}_{\mathrm{EM}}(T_2)$ have been computed. Finally, the two terms on the right-hand side can be expressed into $I(\mathrm{CV}_{\mathrm{EM}}(T_1))$ and $I(\mathrm{CV}_{\mathrm{EM}}(T_2))$ using Corollary 3.2: Following the recipe at the end of Sect. 3, we may compute e.g. $I(\mathrm{CV}_{\mathrm{EM}}(T_1) * \mathrm{CV}_{\mathrm{EM}}(\flat_q T_2)) = I(\mathrm{CV}_{\mathrm{EM}}(T_1) \cdot \mathrm{Hom}_G(L_2, V_q))$. This concludes the case where T contains a vertex of valency 2.

Finally, if T is neither a star nor contains a vertex of valency two, then it contains an edge pr where both p and r are internal vertices of valency at least three. Let T' be the G-tree obtained from T by inserting two vertices q_1 and q_2 between p and r so that $p \sim q_1 \sim q_2 \sim r$, setting $V_{q_1} := V_r$ with the same bilinear form and basis, and $V_{q_2} := V_p$ with the same bilinear form and basis. Note that every G-spaced sub-star of T' with at least three leaves is also a G-spaced substar of T. This is why we inserted two vertices rather than one: what space should we attach to a single vertex between p and q? See below for a comment on this. By the previous construction, we can express the ideal of $\mathrm{CV}_{\mathrm{EM}}(T')$ in the ideals of $\mathrm{CV}_{\mathrm{EM}}(S)$ of all substars S of T', hence of T, with at least three leaves. So we are done if we can show that $\mathrm{CV}_{\mathrm{EM}}(T) = \mathrm{CV}_{\mathrm{EM}}(T')$. But any $A' \in \mathrm{rep}_G(T')$ gives rise to an $A \in \mathrm{rep}_G(T)$ by setting $A_{pr} := A'_{pq_1} A'_{q_1 q_2} A'_{q_2 r}$. Using Remark 4.2 one finds that $\Psi_T(A) = \Psi_{T'}(A')$. Conversely, for any $A \in \mathrm{rep}_G(T)$ we can factorise A_{pr} into $A'_{pq_1} A'_{q_1 q_2} A'_{q_2 r}$ with G-invariant factors. This gives a representation A' of T' with $\Psi_{T'}(A') = \Psi_T(A)$. This concludes the proof of the theorem.

$\square$

Remark 4.5 Note that for $G = 1$, i.e., for the general Markov model, the proof above can be simplified slightly: one does not need the decomposition into isotypic components, and may apply Theorem 3.1 with $t = 1$ directly.

The proof above yields the following algorithm for computing $I(\mathrm{CV}_{\mathrm{EM}}(T))$ from the ideals of substars.

Algorithm 4.6

Input: a G-spaced tree T and finite generating sets of the ideals $I(\mathrm{CV}_{\mathrm{EM}}(S)) \subseteq K[L(S)^G]$ for all substars S in T with at least three leaves.

Output: a finite generating set of the ideal $I(\mathrm{CV}_{\mathrm{EM}}(T)) \subseteq K[L(T)^G]$.

Procedure:

(1) If T contains only two vertices $p \sim q$, then return the empty set and quit.

(2) If T is a star with at least three leaves, then a finite generating set of $I(\mathrm{CV}_{\mathrm{EM}}(T))$ is part of the input; return this set and quit.

(3) If T has a vertex of valency 2, then choose such a vertex q and split $T = T_1 * T_2$ at q. Apply this algorithm to T_1 and to T_2 (with the ideals of their substars) to find finite generating sets F_1, F_2 of the ideals of $\mathrm{CV}_{\mathrm{EM}}(T_1)$ and $\mathrm{CV}_{\mathrm{EM}}(T_2)$, respectively. Let L_1, L_2 be as in the proof of theorem 1.6, identify $L(T)^G = \mathrm{Hom}_G(L_1, L_2)$ with $M_{\mathbf{k},\mathbf{m}}$ as in that proof and in Equation (4), and write $\mathbf{l}$ for the tuple of multiplicities $m(\omega_i, V_q)$. Identify $L(T_1)$ with $M_{\mathbf{k},\mathbf{l}}$ and $L(T_2)$ with $M_{\mathbf{l},\mathbf{m}}$, so that $F_1 \subseteq K[M_{\mathbf{k},\mathbf{l}}]$ and $F_2 \subseteq K[M_{\mathbf{l},\mathbf{m}}]$. Write Ψ for an element in $M_{\mathbf{k},\mathbf{m}}$ whose coordinates are variables.

 (a) Let F' denote the collection of all $(l_i + 1) \times (l_i + 1)$-minors of the ith component of Ψ for all $i = 1, \ldots, t$.

 (b) For an element $\Psi_0 \in M_{\mathbf{m},\mathbf{l}}$ with new variables as coordinates, expand $f(\Psi \cdot \Psi_0)$ for each $f \in F_1$, and take all coefficients of monomials in Ψ_0, which are polynomials in Ψ. Collect these polynomials in F_1'.

 (c) For an element $\Psi_0 \in M_{\mathbf{k},\mathbf{l}}$ with new variables as coordinates, expand $f(\Psi_0 \cdot \Psi)$ for each $f \in F_2$, and take all coefficients of monomials in Ψ_0, which are polynomials in Ψ. Collect these polynomials in F_2'.

 Return $F' \cup F_1' \cup F_2'$ and quit.

(4) Take an edge pr in T connecting two vertices of valency at least three. Construct a G-tree T' by inserting two new vertices q_1, q_2 such that $p \sim q_1 \sim q_2 \sim r$ and setting $V_{q_1} := V_r$ with the same basis and bilinear form and $V_{q_2} := V_p$ with the same basis and bilinear form. Run this algorithm on T', return the same output as for T', and quit.

Although the tree grows in the last step, it is easy to see that this algorithm terminates: after inserting vertices, in the call with T' the tree is broken into two trees, each of which have strictly less substars with at least three vertices. This algorithm is partly carried out in Example 5.4.

Remark 4.7 In the last step of both the proof and the algorithm we could also have inserted a single vertex q between p and r, with V_q equal to the G-module having multiplicities $m(\omega_i, V_q) = \min\{m(\omega_i, V_p), m(\omega_i, V_r)\}$ for all i, so that all G-equivariant maps $V_r \to V_p$ factorise through V_q. One can show that this V_q carries a G-invariant, non-degenerate symmetric bilinear form since V_p and V_r do. This set-up would have raised two minor problems. First, the object T' thus constructed is strictly speaking not a G-tree, as V_q may not have an orthogonal basis permuted by G. But as we saw in Remarks 4.2 and 4.4 this is not really a problem: we can still apply Theorem 3.1 at q to split T' into smaller trees. Second, the G-spaced stars S_p', S_r' with centres p and r in T' are not equal to the G-spaced stars S_p, S_r around p and r in T. Hence after expressing $I(\mathrm{CV}_{\mathrm{EM}}(T'))$ in the ideals $I(\mathrm{CV}_{\mathrm{EM}}(S))$ for all stars S with at least three leaves in T', we still need to express the ideals of $\mathrm{CV}_{\mathrm{EM}}(S_p')$ and $\mathrm{CV}_{\mathrm{EM}}(S_r')$ in $\mathrm{CV}_{\mathrm{EM}}(S_p)$ and $\mathrm{CV}_{\mathrm{EM}}(S_r)$, respectively, to prove the theorem. The following lemma does just that. This would give a slight variant of the algorithm above.

Lemma 4.8 *Let S', S be G-spaced stars with the same underlying star having q as centre and $p_1, \ldots, p_k$ as leaves. Suppose that both stars have the same space V_q with the same basis B_q and that we are given G-equivariant injections $\tau_i : V'_{p_i} \to V_{p_i}$ for $i = 1, \ldots, k$, where V'_{p_i} and V_{p_i} are attached to p_i in S' and S, respectively. Denote by τ the induced injection $L(S') \to L(S)$. Then*

$$I(\mathrm{CV}_{\mathrm{EM}}(S')) = \tau^{\sharp} I(\mathrm{CV}_{\mathrm{EM}}(S)).$$

In particular, a finite generating set for $I(\mathrm{CV}_{\mathrm{EM}}(S))$ gives a finite generating set for $I(\mathrm{CV}_{\mathrm{EM}}(S'))$ under pull-back by τ.

Proof For the inclusion $\supseteq$ note that any G-representation $A = (A_{p_i,q})_i$ of S gives rise to a representation $A' = (\tau_i A_{p_i,q})_i$ of S' satisfying $\tau \Psi_S(A) = \Psi_{S'}(A')$. Hence τ maps $\mathrm{CV}_{\mathrm{EM}}(S)$ into $\mathrm{CV}_{\mathrm{EM}}(S')$.

For the inclusion $\subseteq$ note that, as V_{p_i} is a completely reducible G-module, there exist G-equivariant surjections $\pi_i : V'_{p_i} \to V_{p_i}$ with $\pi_i \tau_i = \mathrm{id}_{V_{p_i}}$. Now the induced projection $\pi : L(S') \to L(S)$ maps $\mathrm{CV}_{\mathrm{EM}}(S')$ into $\mathrm{CV}_{\mathrm{EM}}(S)$, and if $f \in I(\mathrm{CV}_{\mathrm{EM}}(S))$, then $f = \tau^{\sharp}(\pi^{\sharp} f)$, where $\pi^{\sharp} f$ lies in $I(\mathrm{CV}_{\mathrm{EM}}(S'))$.   □

Now we prove our second main result.

Proof of Theorem 1.7 Recall the statement of the theorem: for any G-tree T we have

$$I(\mathrm{CV}_{\mathrm{EM}}(T)) = \sum_{q \in \mathrm{vertex}(T)} I(\mathrm{CV}_{\mathrm{EM}}(\flat_q T)). \tag{$*$}$$

We proceed by induction. First, the statement is a tautology for a star T. Next, suppose that T has an inner vertex q of valency 2 and split $T = T_1 * T_2$ at q. By induction we may assume that the theorem holds for T_i. The proof of Theorem 1.6 shows that

$$I(\mathrm{CV}_{\mathrm{EM}}(T)) = I(\mathrm{CV}_{\mathrm{EM}}(T_1) * \mathrm{CV}_{\mathrm{EM}}(\flat_q T_2)) + I(\mathrm{CV}_{\mathrm{EM}}(\flat_q T_1) * \mathrm{CV}_{\mathrm{EM}}(T_2)),$$

so it suffices to prove that each of these terms is contained in the right-hand side of ($*$); we do so for the first term. We use the notation $L_1, L_2, \mathbf{k}, \mathbf{l}, \mathbf{m}$ from the proof of Theorem 1.6. By Corollary 3.2 the ideal of $\mathrm{CV}_{\mathrm{EM}}(T_1) * \mathrm{CV}_{\mathrm{EM}}(\flat_q T_2)$ is spanned by polynomials in $\Psi \in L(T) = \mathrm{Hom}_G(L_2, L_1)$ of the following two forms:

(1) for all $i = 1, \ldots, t$ the $(l_i + 1)$-minors of the i-th component of Ψ, regarded as an element of $M_{\mathbf{k},\mathbf{m}}$; and
(2) all polynomials of the form $\Psi \mapsto f(\Psi \Psi_0)$, where $f \in I(\mathrm{CV}_{\mathrm{EM}}(T_1))$ and Ψ_0 is some element of $\mathrm{Hom}_G(V_p, L_2)$.

The first type of elements are clearly equations for $\mathrm{CV}_{\mathrm{GM}}(\flat_q T)$, so we need only worry about the second type of equations. By induction we may assume that f is an equation for $\mathrm{CV}_{\mathrm{GM}}(\flat_r T_1)$ for some vertex r of T_1. But then $\Psi \mapsto f(\Psi \Psi_0)$ vanishes on $\mathrm{CV}_{\mathrm{GM}}(\flat_r T)$, and we are done.

Finally, if T is not a star and does not contain a vertex of valency 2, then we proceed as in the proof of Theorem 1.6. We choose an edge pr in T where both p and r have valency at least three, and insert vertices q_1, q_2 with $p \sim q_1 \sim q_2 \sim r$ and $V_{q_1} := V_r$ and $V_{q_2} := V_p$ to obtain a new G-spaced tree T'. We claim that both sides of (*) remain unchanged in replacing T by T'. For the left-hand side this was proved in the proof of Theorem 1.6. The right-hand side gains two terms, namely, $I(\mathrm{CV}_{\mathrm{EM}}(\flat_{q_1} T'))$ and $I(\mathrm{CV}_{\mathrm{EM}}(\flat_{q_2} T'))$. However, the definition of flattening readily implies that $\mathrm{CV}_{\mathrm{EM}}(\flat_p T) \subseteq \mathrm{CV}_{\mathrm{EM}}(\flat_{q_2} T')$, so that $I(\mathrm{CV}_{\mathrm{EM}}(\flat_{q_2} T')) \subseteq I(\mathrm{CV}_{\mathrm{EM}}(\flat_p T))$. Similarly, we find $I(\mathrm{CV}_{\mathrm{EM}}(\flat_{q_1} T')) \subseteq I(\mathrm{CV}_{\mathrm{EM}}(\flat_r T))$, and hence the two extra terms on the right-hand side of (*) are redundant. Now the theorem for T follows from that for T', which in turn follows by induction as in the previous case. □

Finally, we prove the last statement of Proposition 2.4 which says that $\overline{\mathrm{CV}_{\mathrm{EM}}(T') V_{r'}}$ can be computed from the ideal of $\mathrm{CV}_{\mathrm{EM}}(T')$.

Proof of Proposition 2.4 We will apply Corollary 3.2 where $\mathrm{CV}_{\mathrm{EM}}(T')$ will play the role of V and $V_{r'}$ will play the role of $M_{\mathbf{l},\mathbf{m}}$. To this end we proceed as in the proof of Theorem 1.6. First set $L := \bigotimes_{p \in \mathrm{leaf}(T') \setminus \{r'\}} V_p$ and let $\omega_1, \ldots, \omega_t$ be the characters of G. Next define $\mathbf{k}, \mathbf{l}, \mathbf{m}$ as follows: $k_i := m(\omega_i, L)$, $l_i := m(\omega_i, V_{r'})$, and $m_i := m(\omega_i, KG)$ with KG the left regular representation of G. View $\mathrm{CV}_{\mathrm{EM}}(T')$ as a subvariety of $\mathrm{Hom}_G(V_{r'}, L)$, which can be thought of as $M_{\mathbf{k},\mathbf{l}}$, and view $V_{r'}$ as the space $\mathrm{Hom}_G(KG, V_{r'})$ under the isomorphism $v \mapsto (g \mapsto gv)$, which can be thought of as $M_{\mathbf{l},\mathbf{m}}$. It is easy to see that $\mathrm{CV}_{\mathrm{EM}}(T')$ is closed under composition with $\mathrm{Hom}_G(V_{r'}, V_{r'})$, which is $M_{\mathbf{l},\mathbf{l}}$, so that Corollary 3.2 applies. We conclude that $I(\mathrm{CV}_{\mathrm{EM}}(T') V_{r'})$, regarded as a subset of $M_{\mathbf{k},\mathbf{m}}$ is generated by the rank-$(l_i + 1)$-minors of the i-th block for $i = 1, \ldots, t$ and the polynomials $L \to K$ of the form $\Psi \mapsto f(\Psi \Psi_0)$, where Ψ is regarded as a G-homomorphism $KG \to L$, f runs over $I(\mathrm{CV}_{\mathrm{EM}}(T'))$ and Ψ_0 over all elements of $\mathrm{Hom}_G(V_r, KG)$, which is $M_{\mathbf{l},\mathbf{m}}$.

Of course, like in Algorithm 4.6, this can be made into a finite set of generators by taking the entries of Ψ_0 to be variables, taking f in a finite generating set of $I(\mathrm{CV}_{\mathrm{EM}}(T'))$, expanding, and taking the coefficients of the monomials in Ψ_0. □

We have now reduced the ideals of our equivariant models to those for stars, and argued their relevance for statistical applications. The main missing ingredients for successful applications are equations for star models. These are very hard to come by: [9] posed several conjectures concerning these for the general Markov model, and special cases of these conjectures were proved in [1,13,14]. For certain important equivariant models equations were found in [4,16]. Roughly speaking, the less symmetry one imposes on the model, the harder it is to find equations. The following proposition offers some explanation for this.

Proposition 4.9 *Suppose that T is a star with (based) centre r. Let $b_1, \ldots, b_s$ be representatives of the G-orbits on B_r and denote by G_i the stabiliser of b_i in G. Let C_i denote the cone of pure tensors in $\bigotimes_{p \in \mathrm{leaf}(T)}(V_p^{G_i}) \subseteq L(T)$, and denote by ρ the Reynolds operator for G. Then*

$$\mathrm{CV_{EM}}(T) = \overline{\rho(C_1)} + \cdots + \overline{\rho(C_s)},$$

where the addition corresponds to taking the join of these varieties.

Proof Elements in an open dense subset of $\mathrm{CV_{EM}}(T)$ look like

$$\sum_{b \in B_r} \bigotimes_{p \in \mathrm{leaf}(T)} A_{pr}b = \sum_{i=1}^s \sum_{g \in G/G_i} \bigotimes_{p \in \mathrm{leaf}(T)} A_{pr}gb_i$$

$$= \sum_{i=1}^s \sum_{g \in G/G_i} g \left(\bigotimes_{p \in \mathrm{leaf}(T)} A_{pr}b_i \right)$$

$$= \sum_{i=1}^s |G/G_i| \rho \left(\bigotimes_{p \in \mathrm{leaf}(T)} A_{pr}b_i \right)$$

$$= \sum_{i=1}^s \rho \left(|G/G_i| \bigotimes_{p \in \mathrm{leaf}(T)} v_{i,p} \right),$$

where $v_{ip} = A_{pr}b_i \in V_p^{G_i}$; the latter element clearly lies in the join $\overline{\rho(C_1)} + \cdots + \overline{\rho(C_s)}$. This argument can be reversed to show the opposite inclusion. □

This is particularly interesting in the case when for all internal vertices q, B_q is a single G-orbit. As usual, we may assume that T is a star, and then $\mathrm{CV_{EM}}(T)$ is simply $\overline{\rho(C)}$ where C is the set of pure tensors in $\bigotimes_{p \in \mathrm{leaf}(T)}(V_p^H)$ with $H = G_b$ the stabiliser of some element $b \in B_q$. ρ being a linear projection now means that the ideal may be computed by elimination theory, at least in principle. This applies to (1), (2), and (6) in Example 2.1.

In the following section we record some further observations for abelian groups G.

5 Abelian groups and toricness

In this section we collect some results on the equivariant model for an abelian group G. The fact that all irreducible representations of G are one-dimensional makes G-equivariant models somewhat easier to analyse than general equivariant models. Recall that an element v in a G-representation is called a weight vector if it is a common eigenvector of all elements of G; in that case the function $\lambda : G \to K^*$ determined by $gv = \lambda(g)v$ is a character of G. We also say that G scales v by λ. The following results are a slight generalisation of results in [8,16].

Proposition 5.1 *Suppose that T is a star with (based) centre r, that G is abelian and that B_r is a single G-orbit. Then $\mathrm{CV_{EM}}(T)$ is a toric variety. More specifically, there exist tori S_p in $\mathrm{GL}(V_p)$ for $p \in \mathrm{leaf}(T)$, diagonalised by certain bases of G-weight vectors such that $\prod_{p \in \mathrm{leaf}(T)} S_p$ stabilises $\mathrm{CV_{EM}}(T)$ with a dense orbit.*

Proof Fix $b \in B_r$. A typical element of $CV_{EM}(T)$ looks like

$$\sum_{g \in G/G_b} g \left(\bigotimes_{p \in \text{leaf}(T)} v_p \right)$$

with $v_p \in V_p^{G_b} =: V_p'$. As G is abelian, V_p' is a G-module. Choose any basis of V_p' that diagonalises G and let S_p be the associated torus in $GL(V_p')$, regarded as a torus in $GL(V_p)$ acting trivially on a G-stable complement of V_p' in V_p. Set $S := \prod_{p \in \text{leaf}(T)} S_p$. Then $(s_p)_p \in S$ sends the element above to

$$\sum_{g \in G/G_b} g \left(\bigotimes_p s_p v_p \right),$$

which again lies in $CV_{EM}(T)$. Moreover, as each S_p has a dense orbit on V_p', S has a dense orbit on $CV_{EM}(T)$. ∎

Remark 5.2 If G is abelian and has k orbits on V_p, then $CV_{EM}(T)$ for a star T is a join of k toric varieties. This fact is exploited in [4] for the strand-symmetric model, where $k = 2$.

Theorem 5.3 *Suppose that G is an abelian group and that T is a G-tree in which G has a single orbit on all B_p with $p \in \text{internal}(T)$. Then $CV_{EM}(T)$ is a toric variety. More precisely, there exists a torus S acting linearly on $L(T)$, stabilising $CV_{EM}(T)$ with a dense orbit, and diagonalised by a tensor product of G-weight bases of the V_p, $p \in \text{leaf}(T)$.*

There is a subtlety here: unlike in Proposition 5.1 such a torus can in general *not* be found in $\prod_{p \in \text{leaf}(T)} Z_{GL(V_p)} G$.

Proof We proceed by induction. First, if T is the single edge pq, then $CV_{EM}(T) = (V_p \otimes V_q)^G = \bigoplus_{\lambda + \mu = 0} V_p[\lambda] \otimes V_q[\mu]$ where (λ, μ) ranges over pairs of characters of G. Let S be a maximal torus in $GL((V_p \otimes V_q)^G)$ which is diagonal with respect to a basis of pure tensors $v_\lambda \otimes v_{-\lambda}$ with $v_\lambda \in V_p$ and $v_{-\lambda} \in V_q$ weight vectors of weights $\lambda, -\lambda$. View S as a torus in $GL(V_p \otimes V_q)$ acting trivially on all $V_p[\lambda] \otimes V_q[\mu]$ with $\lambda + \mu \neq 0$. This S has the properties claimed in the theorem.

Second, if T is a star, then the proposition above does the trick. Third, if T is neither a star nor an edge, then let $q \in \text{internal}(T)$ be any internal vertex of valency two. As in the proof of Theorem 1.6 we may add such a vertex, if necessary, without changing $CV_{EM}(T)$—and in fact, if q is inserted between the internal vertices p, r, then, as B_p, B_r are G-orbits, $V_q := KG$ is sufficiently large. Write $T = T_1 * T_2$ at q, let $L_i := \bigoplus_{p \in \text{leaf}(T_i) \setminus q} V_p$, and let S_1, S_2 be the tori whose existence is claimed by the theorem for the G-trees T_i. In particular, S_i is diagonalised by the tensor product

of a G-weight basis C_i of L_i and a G-weight basis D_i of V_q (such a basis is always orthogonal and may be chosen orthonormal). But since V_q is a permutation module with a single orbit, every weight occurs at most once in V_q, so that (after scaling) $D_1 = D_2 =: D$. Now we let S_i act on L_i as follows: let $c_i \in C_i$ have G-weight λ. If $-\lambda$ is not a weight in V_q, then $S_i c_i := c_i$. If, on the other hand, $-\lambda$ is a weight in V_q, then it is the weight of a unique $d \in D$, and we let S_i scale c_i by the character with which it scales $d \otimes c_i \in L(T_i)$. Now $S := S_1 \times S_2$ acts on $L_1 \otimes L_2 = L(T)$, and is diagonalised by the tensor product of C_1 and C_2.

Finally we verify that S stabilises $\mathrm{CV}_{\mathrm{EM}}(T)$ with a dense orbit. To see this, observe that the map $(V_q \otimes L_1)^G \times (V_q \otimes L_2)^G \to (L_1 \otimes L_2)^G$ given by $(\Psi_1, \Psi_2) \mapsto \sum_{b \in B_q} (\Psi_1 \mid b) \otimes (\Psi_2 \mid b)$ is $S_1 \times S_2$-equivariant: it sends $(d_1 \otimes c_1, d_2 \otimes c_2)$, where d_i and c_i have opposite weight, to $(d_1 \mid d_2)(c_1 \otimes c_2)$, which scales with the same S-characters by definition of the action of S. Since S_i has a dense orbit on $\mathrm{CV}_{\mathrm{EM}}(T_i)$, S has a dense orbit on $\mathrm{CV}_{\mathrm{EM}}(T)$.   □

Theorem 5.3 reduces the computation of the ideals of certain equivariant models to the combinatorics of toric varieties (where we do not require toric varieties to be normal). However, this combinatorics can be very intricate, and it requires great ingenuity to find explicit generators as in [16]. We conclude with an example.

Example 5.4 First let T be a star with centre q and four leaves $p_1, \ldots, p_4$. Let $G = \mathbb{Z}_2 = \{1, x\}$ and $V_p = KG$ for all vertices $p \in T$, with basis G and induced form. For this G and V_p, a star with four leaves is the smallest G-star for which $\mathrm{CV}_{\mathrm{EM}}(T) \subsetneq L(T)^G$.

We are free to choose any basis on $L(T)$, so we pick the orthonormal product basis of the basis of KG diagonalising G. Let us denote this basis by (t, s) where $t = 1 + x$ spans the trivial, and $s = 1 - x$ the sign representation in KG. We will label the basis of $L(T)$ given by the pure tensors $b_1 \otimes b_2 \otimes b_3 \otimes b_4$ ($b_i \in \{t, s\}$) as follows: for a subset $I \subseteq \{1, 2, 3, 4\}$ let $b_I = b_1 \otimes b_2 \otimes b_3 \otimes b_4$ where $b_i = s$ if $i \in I$, and $b_i = t$ otherwise. Then $L(T)^G$ is spanned by all b_I such that $|I|$ is even.

As $\mathrm{Hom}_G(KG, KG) \cong K^2$, an equivariant representation of T is specified by 8 parameters y_i, x_i ($i = 1, 2, 3, 4$) where x_i is dual to s and y_i is dual to t in KG, and $A_{p_i q} = (x_i s + y_i t) \otimes 1 + (y_i t - x_i s) \otimes x$ (where p_1, p_2, p_3, p_4 are the leaves). By Proposition 4.9, $\mathrm{CV}_{\mathrm{EM}}(T) = \rho(P)$ where P is the variety of pure tensors in $L(T)$. Specifically, if $A = (x_i, y_i)_{i=1}^4$ is a representation, then

$$\Psi_T(A) = 2\rho \left(\bigotimes_i (x_i s + y_i t) \right).$$

Let the variables dual to the basis $(b_I)_{I \subseteq \{1,2,3,4\}}$ be denoted by x_I. Then $x_I(\Psi_T(A)) = 0$ if $|I|$ is odd, and

$$x_I(\Psi_T(A)) = \left(\prod_{i \in I} x_i \right) \left(\prod_{i \notin I} y_i \right),$$

Fig. 1 The spaced tree T
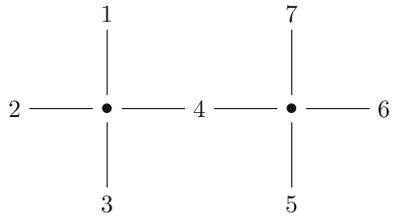
if $|I|$ is even. Thinking of $\mathrm{CV}_{\mathrm{EM}}(T)$ as a subvariety of $L(T)^G$, for determining the ideal we need to consider only functions in the x_I with $|I|$ even. There are some obvious relations, namely, if I, J are two subsets of $\{1, 2, 3, 4\}$ with even number of elements and with complements I^c, J^c, respectively, then

$$f_{I,J} := x_I x_{I^c} - x_J x_{J^c}$$

is in the ideal of $\mathrm{CV}_{\mathrm{EM}}(T)$. Note that it is enough to consider $f_{I,J}$ where every subset appears once as I or J, because $f_{I,J} + f_{J,J'} = f_{I,J'}$, and if $|I| = |J| = 2$, then $f_{I,J}$ is nonzero only if $I \cap J$ contains one element. Altogether it follows that the ideal generated by the $f_{I,J}$'s is already generated by

$$f_{\emptyset,\{1,2\}}, \; f_{\emptyset,\{1,3\}}, \; f_{\emptyset,\{1,4\}}.$$

One can show that $I(\mathrm{CV}_{\mathrm{EM}}(T))$ is generated by the $f_{I,J}$'s. Indeed, by [6, Theorem 3.1] this boils down to showing that these $f_{I,J}$'s correspond to a Markov basis for the module of $\mathbb{Z}$-linear relations among the 8 vectors $(\mathbf{a}, \mathbf{b}) \in \{0, 1\}^4 \times \{0, 1\}^4 \subseteq \mathbb{N}^4 \times \mathbb{N}^4$ where $|\mathbf{a}| := \sum_i a_i$ is even and $\mathbf{b} = \mathbf{1} - \mathbf{a}$ with $\mathbf{1} = (1, 1, 1, 1)$. We omit the combinatorial details here.

To illustrate Theorem 1.6, we now consider a G-tree T obtained by gluing together two stars as above at one common leaf (Fig. 1). Notice that by Remark 4.7 and Lemma 4.8 $\mathrm{CV}_{\mathrm{EM}}(T) = \mathrm{CV}_{\mathrm{EM}}(T')$ where T' is the tree T with vertex 4 removed and the centres of the two stars in T joined by an edge.

Algorithm 4.6 first identifies a vertex of valency 2; here vertex 4. We then write $T = T_1 * T_2$, with T_1 the left and T_2 the right star with four leaves each. By the above we know the ideals of $\mathrm{CV}_{\mathrm{EM}}(T_i)$. The content of Theorem 1.6 in this situation is that

$$I(\mathrm{CV}_{\mathrm{EM}}(T)) = I\left(\mathrm{CV}_{\mathrm{EM}}(T_1) * L(T_2)^G\right) + I\left(L(T_1)^G * \mathrm{CV}_{\mathrm{EM}}(T_2)\right).$$

Because of the symmetry of the problem, we only consider the first summand. We keep the notation introduced above with respect to T_i: the variables on $L(T_1)^G$ will be x_I ($I \subseteq \{1, 2, 3, 4\}$, $|I|$ even) and those on $L(T_2)^G$ will be y_I ($I \subseteq \{4, 5, 6, 7\}$, $|I|$ even). Finally the variables on $L(T)^G$ will be z_J where $J \subseteq \{1, 2, 3, 5, 6, 7\}$ has an even number of elements; J corresponds to the basis vector $b_J = b_1 \otimes b_2 \otimes b_3 \otimes b_5 \otimes b_6 \otimes b_7$ where $b_i = s$ if $i \in J$ and $b_i = t$ otherwise. We also adopt the convention that $x_I, y_I, z_I = 0$ if $|I|$ is odd.

The ideal of $CV_{EM}(T_1) * L(T_2)^G$ is generated by $I' = I(CV_{EM}(T_1)')$, where the accent refers to the notation of Corollary 3.2 and the discussion preceding it, and certain 2×2-minors, since both representations of $\mathbb{Z}_2$ occur with multiplicity one in V_4. These minors are of the form

$$z_{I_1 \cup I_2} z_{I'_1 \cup I'_2} - z_{I_1 \cup I'_2} z_{I'_1 \cup I_2}$$

where $I_1, I'_1 \subseteq \{1, 2, 3\}$, $I_2, I'_2 \subseteq \{5, 6, 7\}$ are all distinct and either all even or all odd. Next we show how to find generators of I'. The space $L(T_2)^G$ is isomorphic to $\mathrm{Hom}_G(V_4, \bigotimes_{p \in \mathrm{leaf}(T_2) \setminus \{4\}}) \cong M_{\mathbf{m},\mathbf{l}}$ with $\mathbf{m} = (4, 4)$ and $\mathbf{l} = (1, 1)$ by Eq. (4). Similarly,

$$L(T)^G \cong \mathrm{Hom}_G \left(\bigotimes_{p \in \mathrm{leaf}(T_2) \setminus \{4\}} V_p, \bigotimes_{p \in \mathrm{leaf}(T_1) \setminus \{4\}} V_p \right) \cong M_{\mathbf{k},\mathbf{m}}$$

with $\mathbf{k} = (4, 4)$ and $\mathbf{m}$ as above.

Let $\Psi_0 \in L(T_2)^G$ be arbitrary and let $\Psi \in L(T)^G$. Then $\Psi \cdot \Psi_0 \in \mathrm{Hom}_G(V_4, \bigotimes_{p \in \mathrm{leaf}(T_1) \setminus \{4\}} V_p) = L(T_1)^G$ and a straightforward computation shows that

$$x_I(\Psi \cdot \Psi_0) = \begin{cases} \sum_{J \subseteq \{5,6,7\}} z_{J \cup I}(\Psi) y_J(\Psi_0) & \text{if } 4 \notin I, \text{ and} \\ \sum_{J \subseteq \{5,6,7\}} z_{J \cup I \setminus \{4\}}(\Psi) y_{J \cup \{4\}}(\Psi_0) & \text{if } 4 \in I. \end{cases}$$

To avoid clumsy notation, let us write $\tilde{x}_I$ for the function $x_I(\Psi \cdot \Psi_0)$ in the arguments (Ψ, Ψ_0). Then for example

$$\tilde{x}_\emptyset = z_\emptyset y_\emptyset + z_{\{5,6\}} y_{\{5,6\}} + z_{\{5,7\}} y_{\{5,7\}} + z_{\{6,7\}} y_{\{6,7\}} \tag{5}$$
$$\tilde{x}_{\{1,2,3,4\}} = z_{\{1,2,3,5\}} y_{\{4,5\}} + z_{\{1,2,3,6\}} y_{\{4,6\}} + z_{\{1,2,3,7\}} y_{\{4,7\}}$$
$$+ z_{\{1,2,3,5,6,7\}} y_{\{4,5,6,7\}} \tag{6}$$
$$\tilde{x}_{\{1,2\}} = z_{\{1,2\}} y_\emptyset + z_{\{1,2,5,6\}} y_{\{5,6\}} + z_{\{1,2,5,7\}} y_{\{5,7\}} + z_{\{1,2,6,7\}} y_{\{6,7\}} \tag{7}$$

and finally

$$\tilde{x}_{\{3,4\}} = z_{\{3,5\}} y_{\{4,5\}} + z_{\{3,6\}} y_{\{4,6\}} + z_{\{3,7\}} y_{\{4,7\}} + z_{\{3,5,6,7\}} y_{\{4,5,6,7\}}. \tag{8}$$

I' is then generated by all f'_{Ψ_0} where f is in the ideal of $CV_{EM}(T_1)$. As in Algorithm 4.6, step 33b , this is the same as the ideal generated by all coefficients of monomials in the y_I's. As an example let us consider $f_{\emptyset,\{1,2\}} = x_\emptyset x_{\{1,2,3,4\}} - x_{\{1,2\}} x_{\{3,4\}}$. Using the relations (5)–(8), we get an expression in the z_I's and y_I's. As a function on $M_{\mathbf{k},\mathbf{m}} \times M_{\mathbf{m},\mathbf{l}}$

it is equal to

$$
\begin{aligned}
f_{\emptyset,\{1,2\}}(\Psi\Psi_0) = \; & \tilde{x}_\emptyset \tilde{x}_{\{1,2,3,4\}} - \tilde{x}_{\{1,2\}}\tilde{x}_{\{3,4\}} \\
= \; & (z_\emptyset z_{\{1,2,3,5\}} - z_{\{1,2\}}z_{\{3,5\}})y_\emptyset y_{\{4,5\}} \\
& + (z_\emptyset z_{\{1,2,3,6\}} - z_{\{1,2\}}z_{\{3,6\}})y_\emptyset y_{\{4,6\}} \\
& + (z_\emptyset z_{\{1,2,3,7\}} - z_{\{1,2\}}z_{\{3,7\}})y_\emptyset y_{\{4,7\}} \\
& + (z_\emptyset z_{\{1,2,3,5,6,7\}} - z_{\{1,2\}}z_{\{3,5,6,7\}})y_\emptyset y_{\{4,5,6,7\}} \\
& + (z_{\{5,6\}}z_{\{1,2,3,5\}} - z_{\{1,2,5,6\}}z_{\{3,5\}})y_{\{4,5\}}y_{\{5,6\}} \\
& + (z_{\{5,7\}}z_{\{1,2,3,5\}} - z_{\{1,2,5,7\}}z_{\{3,5\}})y_{\{4,5\}}y_{\{5,7\}} \\
& + (z_{\{6,7\}}z_{\{1,2,3,5\}} - z_{\{1,2,6,7\}}z_{\{3,5\}})y_{\{4,5\}}y_{\{6,7\}} \\
& + (z_{\{5,6\}}z_{\{1,2,3,6\}} - z_{\{1,2,5,6\}}z_{\{3,6\}})y_{\{4,6\}}y_{\{5,6\}} \\
& + (z_{\{5,7\}}z_{\{1,2,3,6\}} - z_{\{1,2,5,7\}}z_{\{3,6\}})y_{\{4,6\}}y_{\{5,7\}} \\
& + (z_{\{6,7\}}z_{\{1,2,3,6\}} - z_{\{1,2,6,7\}}z_{\{3,6\}})y_{\{4,6\}}y_{\{6,7\}} \\
& + (z_{\{5,6\}}z_{\{1,2,3,7\}} - z_{\{1,2,5,6\}}z_{\{3,7\}})y_{\{4,7\}}y_{\{5,6\}} \\
& + (z_{\{5,7\}}z_{\{1,2,3,7\}} - z_{\{1,2,5,7\}}z_{\{3,7\}})y_{\{4,7\}}y_{\{5,7\}} \\
& + (z_{\{6,7\}}z_{\{1,2,3,7\}} - z_{\{1,2,6,7\}}z_{\{3,7\}})y_{\{4,7\}}y_{\{6,7\}} \\
& + (z_{\{5,6\}}z_{\{1,2,3,5,6,7\}} - z_{\{1,2,5,6\}}z_{\{3,5,6,7\}})y_{\{5,6\}}y_{\{4,5,6,7\}} \\
& + (z_{\{5,7\}}z_{\{1,2,3,5,6,7\}} - z_{\{1,2,5,7\}}z_{\{3,5,6,7\}})y_{\{5,7\}}y_{\{4,5,6,7\}} \\
& + (z_{\{6,7\}}z_{\{1,2,3,5,6,7\}} - z_{\{1,2,6,7\}}z_{\{3,5,6,7\}})y_{\{6,7\}}y_{\{4,5,6,7\}}.
\end{aligned}
\tag{9}
$$

Every single coefficient of a monomial in the y_I's then gives a generator for I'. It should be clear how to proceed in principle with the other $f_{I,J}$'s. So Algorithm 4.6 calls itself twice, once for T_1 and once for T_2.

References

1. Allman, E.S., Rhodes, J.A.: Phylogenetic ideals and varieties for the general Markov model. Adv. Appl. Math. **40**(2), 127–148 (2008)
2. Buczyńska, W., Wiśniewski, J.A.: On geometry of binary symmetric models of phylogenetic trees. J. Eur. Math. Soc. **9**(3), 609–635 (2007)
3. Casanellas, M., Fernández-Sánchez, J.: The geometry of the Kimura 3-parameter model. Adv. Appl. Math. (2007, to appear). Preprint available from http://arxiv.org/abs/math/0702834
4. Casanellas, M., Sullivant, S.: The strand symmetric model. In: Algebraic Statistics for Computational Biology. Cambridge University Press, Cambridge (2005)
5. Derksen, H., Kemper, G.: Computational Invariant Theory, vol. 130 of Encyclopaedia of Mathematical Sciences. Springer, Berlin (2002)

6. Diaconis, H., Sturmfels, B.: Algebraic algorithms for sampling from conditional distributions. Ann. Stat. **26**(1), 363–397 (1998)
7. Eriksson, N., Ranestad, K., Sturmfels, B., Sullivant, S.: Phylogenetic algebraic geometry. In: Projective varieties with unexpected properties, pp. 237–255. Walter de Gruyter GmbH & Co. KG (2005)
8. Evans, S.N., Speed, T.P.: Invariants of some probability models used in phylogenetic inference. Ann. Statist. **21**(1), 355–377 (1993)
9. Garcia, L.D., Stillman, M., Sturmfels, B.: Algebraic geometry of Bayesian networks. J. Symb. Comput. **39**(3–4), 331–355 (2005)
10. Goodman, R., Wallach, N.R.: Representations and Invariants of the Classical Groups. Cambridge University Press, London (1998)
11. Hartshorne, R.: Algebraic Geometry. Springer, New York (1977)
12. Kraft, H., Procesi, C.: A Primer in Invariant Theory (unpublished). Text available from http://www.math.unibas.ch/~kraft/Papers/KP-Primer.pdf
13. Landsberg, J.M., Manivel, L.: On the ideals of secant varieties of Segre varieties. Found. Comput. Math. **4**(4), 397–422 (2004)
14. Landsberg, J.M., Weyman, J.: On the ideals and singularities of secant varieties of Segre varieties. Bull. Lond. Math. Soc. **39**(4), 685–697 (2007)
15. Pachter, L., Sturmfels, B., (eds.): Algebraic Statistics for Computational Biology. Cambridge University Press, Cambridge (2005)
16. Sturmfels, B., Sullivant, S.: Toric ideals of phylogenetic invariants. J. Comput. Biol. **12**, 204–228 (2005)
17. Weyl, H.: The Classical Groups, their Invariants and Representations. Princeton University Press, Princeton, NJ (1939)