ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Impossibility Results on Stability of Phylogenetic Consensus Methods

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:
Delucchi, E., Hoessly, L., Paolini, G. (2020). Impossibility Results on Stability of Phylogenetic Consensus Methods. SYSTEMATIC BIOLOGY, 69(3), 557-565 [10.1093/sysbio/syz071].

Availability:
This version is available at: https://hdl.handle.net/11585/943295 since: 2023-09-28

Published:
DOI: http://doi.org/10.1093/sysbio/syz071

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

12 October 2024

This is the final peer-reviewed accepted manuscript of:

Emanuele Delucchi, Linard Hoessly, Giovanni Paolini, Impossibility Results on Stability of Phylogenetic Consensus Methods, *Systematic Biology*, Volume 69, Issue 3, May 2020, Pages 557–565

The final published version is available online at ***https://dx.doi.org/10.1093/sysbio/syz071***

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/)

When citing, please refer to the published version.

IMPOSSIBILITY RESULTS ON STABILITY OF PHYLOGENETIC CONSENSUS METHODS

EMANUELE DELUCCHI, LINARD HOESSLY, AND GIOVANNI PAOLINI

Abstract. We answer two questions raised by Bryant, Francis and Steel in their work on consensus methods in phylogenetics. Consensus methods apply to every practical instance where it is desired to aggregate a set of given phylogenetic trees (say, gene evolution trees) into a resulting, "consensus" tree (say, a species tree). Various stability criteria have been explored in this context, seeking to model desirable consistency properties of consensus methods as the experimental data are updated (e.g., more taxa, or more trees, are mapped). However, such stability conditions can be incompatible with some basic regularity properties that are widely accepted to be essential in any meaningful consensus method. Here, we prove that such an incompatibility does arise in the case of extension stability on binary trees and in the case of associative stability. Our methods combine general theoretical considerations with the use of computer programs tailored to the given stability requirements.

Context. The problem of merging the information carried by a set of phylogenetic trees into a resultant ("consensus") tree is standard and well-studied. For instance, this problem arises as one tries to combine many gene trees in order to reconstruct a common species phylogeny, or when aggregating a set of estimates resulting from the application of different clustering algorithms to the same genomic data set. More generally, consensus methods have wide applications in biology [9] as well as in other sciences, for example, social choice theory [3].

This variety of applications has motivated a general axiomatic study of consensus methods. In biology, the field was pioneered by McMorris and collaborators, see [8] for a survey. Here, one of the research threads is the study of "stability conditions" for consensus functions, which encode the requirement that a consensus method should be consistent under "restriction" of all input trees to a subset of taxa, see for example [4]. For example, when computed on the branching structures induced on a specific subset of the taxa, the consensus method should output the branching structure induced by the consensus tree computed from the full data. A main question is whether such stability conditions are compatible with "Pareto-type" properties, where one requires that if some partial feature is shared among all trees we want to aggregate, then this feature should be present in the consensus tree as well.

Motivation and aim. Our paper is motivated by two questions asked in a recent work of Bryant, Francis and Steel [6], who followed up on [19]. In their paper, they carry out a detailed feasibility analysis of stability conditions that express "future-proofing" of phylogenetic trees, that is, consistency of consensus methods with respect to increase of experimental evidence. (For example, an increase of the set

of taxa or an increase of the size of the set of trees from which the consensus is to be drawn.) Let us explain intuitively the four properties of consensus methods on which, following [6], we will focus (for precise definitions see Section 1).

> **Regularity** properties ensure that the output does not depend on the naming of taxa nor on the order of the trees. Moreover, if all input trees are equal, then the consensus tree should also be equal to the input trees.
> **Extension stability** requires that, if the input data is updated by including a new taxon in each tree, the branching structure among the "original" taxa is preserved in the updated consensus tree.
> **Associative stability** allows, among other things, to reduce the computation of the consensus tree to a series of consensus problems between pairs of trees.

As is usual, these properties are considered together with a Pareto-type property which, again following [6], we take to be **Pareto on rooted triples**. This means that if all input trees display the same nontrivial branching order when restricted to a specific triple of taxa, then the consensus tree must display the same branching order when restricted to the same taxa.

Bryant, Francis, and Steel conclude by stating two main open questions about the existence of consensus methods [6, Concluding comments]. The first question asks whether there exist regular consensus methods that are extension stable when the input data are restricted to binary trees. The second question asks whether there exist regular consensus methods that are Pareto on rooted triples and associatively stable. We answer both questions in the negative.

Our proof consists in a reduction to a problem of integer linear programming (ILP), which we then solve. ILP has been used in the past in order to compute consensus trees, see [11]; however, ours seems to be the first application of ILP to a (non)existence proof in phylogenetics.

1. Background

1.1. **Phylogenetic trees.** Our setup mostly follows [18], and in particular we restrict our attention to *rooted phylogenetic trees*. We fix a set (say, of taxa) X and write $\mathsf{RP}(X)$ for the set of rooted phylogenetic trees on the leaf set X. A *cluster* of a tree is any set of leaves that consists of all descendants of a particular vertex of the given tree. The set of clusters of a tree forms a hierarchy (We call *hierarchy* any family of subsets of a given set such that any two elements in the family intersect trivially, that is, their intersection is either empty or equal to one of the two sets). For every hierarchy on a set X that contains X itself and all singleton sets, but does not contain the empty set, there is a unique phylogenetic tree whose clusters form the given hierarchy. In particular, two trees have the same associated hierarchy if and only if they are equivalent.

We say a tree $T' \in \mathsf{RP}(X)$ *refines* a tree $T \in \mathsf{RP}(X)$, and write $T \preceq T'$, if the hierarchy of T is contained in that of T' (this means that every cluster of T is also a cluster of T'). This defines a partially ordered set $(\mathsf{RP}(X), \preceq)$ whose maximal elements are given by the binary trees and whose unique minimal element is the "star" tree, where every leaf is adjacent to the root (the hierarchy of the star tree consists only of the singletons and X itself). Given a tree $T \in \mathsf{RP}(X)$ and a subset

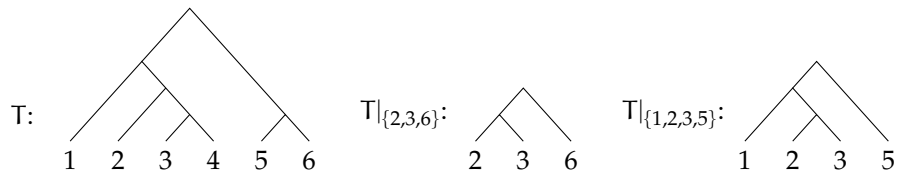

FIGURE 1. A tree T on the set of taxa $X = \{1,2,3,4,5,6\}$ and its restrictions $T|_Y$ for $Y = \{2,3,6\}$ and $Y = \{1,2,3,5\}$.

$Y \subset X$, the restriction of T to Y is the tree $T|_Y \in RP(Y)$ obtained by restricting T to the leaves in Y (see Figure 1). If $|Y| = 3$ and $T|_Y$ is binary, we say that the *rooted triple* $T|_Y$ is displayed by T. Notice that any rooted phylogenetic tree is fully determined by its set of rooted triples.

1.2. **Consensus functions and consensus methods.**

Definition 1.1. Let X be a finite set of taxa and $k \in \mathbb{N}$ a natural number. A *profile* of trees is an element $(T_1, \cdots, T_k) \in RP(X)^k$.

(1) A *k-consensus function on* X is a function
$$\varphi_X^k : RP(X)^k \to RP(X).$$
(2) A *consensus function on* X is a function
$$\varphi_X : \cup_{k \geqslant 1} RP(X)^k \to RP(X).$$
(3) A *k-consensus method* is a function that, for every set Y of taxa, associates with any profile $(T_1, \cdots, T_k) \in RP(Y)^k$ a tree $\varphi_Y^k(T_1, \cdots, T_k) \in RP(Y)$. We consider such a method as a set of functions φ_Y^k, one for every Y, and denote it simply by φ^k.
(4) A *consensus method* is a function that, for every set of taxa Y and any $k \in \mathbb{N}$, associates with any profile $(T_1, \cdots, T_k) \in RP(Y)^k$ a tree in $\varphi_Y(T_1, \cdots, T_k) \in RP(Y)$. We consider such a method as a set of k-consensus methods φ^k, one for every k, and denote it by φ.

We will at times need to consider (k-)consensus methods where the only allowed sets of taxa are the subsets of a given finite set X: in this case, we will speak of a *(k-)consensus method on* X.

To summarize the terminology set-up: "functions" are rules that apply only to a fixed set of taxa, while "methods" do not have this restriction. The length of an input profile is not restricted, unless a prefix is added (as in "k-consensus").

Remark 1.2 (Related definitions in the literature). Our use of the term "consensus method" conforms to [5, 6], whereas the term "consensus function on X" matches [18]. However, the terminology is not completely consistent throughout the literature: in particular, we remark on some instances where the objects we introduced in Definition 1.1 appear under different names. In Day and McMorris' book [8], our k-consensus functions on X and consensus functions on X are called consensus rules and complete consensus rules, respectively. In addition, the "supertrees" treated in [19] are analogues of consensus methods where one does not require the leaf sets of all trees in a profile to coincide.

1.3. **Axiomatic requirements for consensus methods.** We now recall some general axioms for consensus methods. We follow [6] and call a consensus method φ **regular** if it satisfies the following three axioms:

(1) **Unanimity.** The value of φ on any profile consisting of a k-fold repetition of a single tree T is T itself.
(2) **Anonymity.** Changing the order of the trees in a profile does not affect the value of φ on it.
(3) **Neutrality.** Changing the labels on the leaves of the trees in a profile simply relabels the leaves of the consensus tree in the same way.

Furthermore we say that a consensus method φ on a set X of taxa is **Pareto on rooted triples** if the following condition is satisfied for all $Y \subseteq X$ and all trees $T_1, \ldots, T_k \in RP(X)$, $T' \in RP(Y)$:

$$\text{if } T' \preceq T_i|_Y \text{ for all } i = 1, \ldots k, \text{ then } T' \preceq \varphi_X^k(T_1, \ldots, T_k)|_Y.$$

An equivalent rephrasing in more colloquial terminology is that any rooted triple that is displayed by a set of trees must be displayed by their consensus tree as well.

1.4. **Examples.** In the following we mention a few consensus methods that have appeared in the literature, witnessing the existence of methods that do satisfy several combinations of the above-mentioned properties. In order to describe them precisely, we will make use of the characterization of trees by means of hierarchies of clusters, see above. For more examples of consensus methods we refer to [5, 8, 18]. An overview is given in [5, Figure 2].

A first class of consensus methods determines the hierarchy of clusters of the consensus tree based on the frequency of appearance of those clusters in the input trees. **Majority rule**, which is probably the most widely used consensus method in practice, returns the tree determined by the hierarchy of all clusters that appear in more than half of the input trees. **Strict consensus** returns the tree given by the clusters that appear in every input tree, while **loose consensus** returns the tree defined by the set of all clusters that appear in at least one input tree and are compatible with the other input trees.

A second type of examples is of recursive nature. The idea is to associate to each profile of trees $\mathcal{P} = (T_1, \cdots, T_k) \in RP(X)$ a partition $\Pi(\mathcal{P})$ of X whose blocks will form the maximal clusters of the returned consensus tree. (A *block* of a partition is simply one of its elements, that is, a subset of X.) Then, for every block B of $\Pi(\mathcal{P})$, one computes the partition $\Pi(\mathcal{P}|_B)$ associated with the profile restricted to B, and so forth recursively. The union of the blocks of all partitions is then the hierarchy of the consensus tree. **Adams consensus** [1] defines $\Pi(\mathcal{P})$ as the set of nonempty intersections of the maximal clusters of the trees in $\mathcal{P}$. In **Aho consensus** [2] (which is called local consensus in [5]) the partition $\Pi(\mathcal{P})$ is the set of the connected components of the graph with vertex set X and where a pair of vertices $\{a, b\} \subseteq X$ is joined by an edge if there is some $c \in X$ for which the rooted triple defined by the hierarchy $\{\{a, b\}, \{a\}, \{b\}, \{c\}\}$ is displayed by all trees in $\mathcal{P}$.

All these consensus methods are regular. Aho consensus and Adams consensus are Pareto on rooted triples. As we will discuss later on, strict consensus is

associatively stable (see Definition 2.6), whereas majority consensus, loose consensus, Aho consensus, and Adams consensus are not [6].

2. Results and methodology

2.1. **Extension stability.** This section focuses on consensus methods that satisfy the following condition, defined in [19] and meant to encode the fact that a consensus should behave consistently with respect to passing to subsets of taxa.

Definition 2.1. Fix a positive integer k, and let $\{\varphi^k_Y\}_{\emptyset \neq Y \subseteq X}$ be a k-consensus method on a set X of taxa. This k-consensus method on X is called **extension stable** if, for all nonempty subsets $Y \subseteq X' \subseteq X$ and every profile $(T_1, \ldots, T_k) \in \mathsf{RP}(X')$,

$$\varphi^k_Y(T_1|_Y, \ldots, T_k|_Y) \preceq \varphi^k_{X'}(T_1, \ldots, T_k)|_Y.$$

A k-consensus method φ^k is extension stable if, for all X, the k-consensus method $\{\varphi^k_Y\}_{\emptyset \neq Y \subseteq X}$ on X is extension stable. A consensus method φ is extension stable if for all $k \geqslant 1$ the k-consensus method φ^k is extension stable.

One of the main results of [6] is that *no regular* 2*-consensus method is extension stable*. In the same paper, the feasibility of different relaxations of extension stability was discussed. The first question left open in [6] is about extension stability under the restriction of the domain of consensus methods to binary trees. More precisely, fix a set X of taxa and let $\mathsf{RBP}(X) \subseteq \mathsf{RP}(X)$ denote the subset of all rooted binary phylogenetic trees – that is, the phylogenetic trees on X where every internal vertex has exactly two children.

Definition 2.2. A k-consensus method φ^k_* on X (resp. k-consensus method φ^k, consensus method φ) is *extension stable on binary trees* if the method obtained by restricting each φ^k_* (resp. φ^k, φ) to the set of binary phylogenetic trees is extension stable (in the sense of Definition 2.1, replacing $\mathsf{RP}(X')$ with $\mathsf{RBP}(X')$).

Question 2.3 ([6])**.** Is there a regular consensus method that is extension stable on binary trees?

Theorem 2.4. *There is no regular extension stable* k*-consensus method among profiles of binary trees on more than* 4 *taxa, for any even profile size* k*. In particular, there is no regular and extension stable consensus method on profiles of binary trees.*

The gist of the proof is a verification by means of a computer program that there is no extension stable 2-consensus method on sets of 5 taxa. We will give the details of the computation in the Appendix. The sufficiency of this verification depends on the following, easily checked fact.

Remark 2.5. For every positive even integer k, any k-consensus function φ^k_X on a set X of taxa induces a 2-consensus function φ^2_X on X by setting

$$\varphi^2_X(T_1, T_2) := \varphi^k_X(\underbrace{T_1, \ldots, T_1}_{k/2 \text{ times}}, \underbrace{T_2, \ldots, T_2}_{k/2 \text{ times}}).$$

Regularity and extension stability of φ^k_X are inherited by φ^2_X.

2.2. **Associative stability.**

Definition 2.6 ([6])**.** Let φ denote a consensus method on a set of taxa X. We say that φ is **associatively stable** if the following equality is satisfied for all $T_1, \ldots, T_k \in \mathsf{RP}(X)$:

$$\varphi_X^k(T_1, \ldots, T_k) = \varphi_X^2(\varphi_X^{k-1}(T_1, \ldots, T_{k-1}), T_k).$$

In [6] it is noted that Adams consensus is associatively stable when restricted to trees of height 2, but not for trees of height 4, and that Aho consensus is not associatively stable even for trees of height 2. (The *height* of a rooted tree is the maximum distance between the root and any leaf.) On the other hand, associative stability is satisfied by some elementary methods such as strict consensus, which, however, fails to be Pareto on rooted triples. This motivates the following.

Question 2.7 ([6])**.** Is there a regular consensus method that satisfies associative stability and is Pareto on rooted triples?

Theorem 2.8. *There exists no regular, associatively stable consensus method that is Pareto on rooted triples.*

As was already remarked in [6], if φ is a regular and associatively stable consensus method on a set of taxa X, then φ_X^2 is a commutative, idempotent and associative binary operation on $\mathsf{RP}(X)$. Thus it is enough to prove that such a binary operation does not exist.

Lemma 2.9. *There exists no regular, associative* 2*-consensus function which is Pareto on rooted triples for any set of* 5 *or more taxa.*

Remark 2.10. The only regular and associatively stable consensus method on 3 and 4 taxa, which is Pareto on rooted triples, is Adams consensus. This is discussed in the Appendix.

The proof of Lemma 2.9 rests on a computational check of the case of 5 taxa (see Appendix). From there, the full generality follows via the following lemma.

Lemma 2.11. *Fix a positive integer* k *and a set of taxa* X*. Every regular* k*-consensus function on* X *which is Pareto on rooted triples and associatively stable induces a* k*-consensus function on every subset of* X *which is also regular, Pareto on rooted triples, and associatively stable.*

Proof. Fix a subset $Y \subseteq X$ and an enumeration $x_1, \ldots, x_l$ of the set $X \setminus Y$. Given any tree $T \in \mathsf{RP}(Y)$, define a tree $T^X \in \mathsf{RP}(X)$ as in Figure 2. Notice that, for all $T_1, T_2 \in \mathsf{RP}(Y)$,

$$T_1^X = T_2^X \text{ if and only if } T_1 = T_2. \tag{1}$$

Now, given a consensus method φ on X we can define a consensus method ψ_Y on Y by setting, for every positive integer k,

$$\psi_Y^k(T_1, \ldots, T_k) := \varphi_X^k(T_1^X, \ldots, T_k^X)|_Y.$$

We immediately observe that regularity of φ implies regularity of ψ. If φ_X^k is Pareto on rooted triples, then

$$\varphi_X^k(T_1^X, \ldots, T_k^X) = (\psi_Y^k(T_1, \ldots, T_k))^X.$$

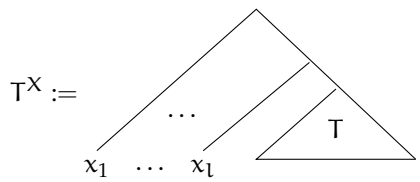

FIGURE 2.

If in addition φ is associatively stable, we can use this equation in order to write, for every $T_1, \ldots, T_k \in RP(Y)$,

$$\begin{aligned}\psi^2_Y(\psi^{k-1}_Y(T_1, \ldots, T_{k-1}), T_k)^X &= \varphi^2_X(\psi^{k-1}_Y(T_1, \ldots, T_{k-1})^X, T_k^X) \\ &= \varphi^2_X(\varphi^{k-1}_X(T_1^X, \ldots, T_{k-1}^X), T_k^X) \\ &= \varphi^k_X(T_1^X, \ldots, T_k^X) = \psi^k_Y(T_1, \ldots, T_k)^X.\end{aligned}$$

In view of Equation (1), this proves associative stability of ψ. □

3. Concluding discussion

We have answered the two main questions left open in [6], about extension stability and associative stability of consensus methods on phylogenetic trees. On the one hand, we have proved that, under widely accepted regularity requirements, there cannot exist any consensus method that is stable under addition of taxa, even when the input trees are required to be binary (Theorem 2.4). We thus strengthen the result of Bryant, Francis and Steel. The meaning of this theorem is that, no matter which method is used in order to extract a consensus from a profile of binary trees, the branching structure in the consensus tree is not guaranteed to hold once the set of available taxa is enlarged – even if the "augmented" input trees agree with the original profile when restricted to the previously available taxa. Our other main result, Theorem 2.8, states that there is no associatively stable consensus method that satisfies some common regularity and Pareto-type properties. This means that, when enlarging the set of trees from which consensus is extracted, it may not be enough to compute the consensus between the new trees and the "old" consensus tree, and thus one is forced to carry out the computation anew, starting from the complete profile of trees. In fact, as pointed out in [6], there do exist consensus methods that satisfy associative stability: such methods however fail to simultaneously possess both basic regularity and Pareto properties. In this light, our result can be interpreted by saying that those basic properties are intrinsically complex – and, in particular, the substantial computational advantage that is granted by associative stability is "too much to hope for".

We note that our considerations about rooted phylogenetic trees have implications for other types of data structures. First, our impossibility results apply also to consensus among **unrooted trees**, see for example [14, Section 3.17] for the relevance of this case. Given a profile of rooted trees, one can consider a profile of unrooted trees obtained from the former by appending a "special leaf" to the

root of all trees in the rooted profile. From any regular (and stable) consensus tree among the unrooted trees we can then extract a regular (and stable) consensus tree for the rooted profile by deleting the special leaf. Thus, no regular and (extension- or associatively) stable consensus method can exist for unrooted trees on more than 6 leaves. Moreover, **supertree consensus** is a generalization of tree consensus, hence any claim of non-existence of tree-consensus functions implies in particular the same claim for supertree consensus.

As was also remarked in [6], such negative results are valuable inasmuch as they uncover the intrinsic limitations of certain approaches, thus helping direct future research towards feasible paths. We suggest four such possible directions of further research.

1. Prompted by discussions with researchers at the Swiss Institute of Bioinformatics, we propose to consider single splits as the basic features of phylogenetic trees, instead of rooted triples. A split in a (rooted) tree is a bipartition of the taxa that is obtained by deleting an edge of the tree.
   Q. Are there regular consensus methods on phylogenetic trees that are Pareto-optimal with respect to single splits and satisfy extension stability (with respect to single splits) or associative stability?
2. Another way forward may involve relaxing the constraint on the data structures outputted by consensus methods. The minimum hybridization network associated with a profile of trees is a rooted phylogenetic network that "minimally deviates from being a tree" among those displaying all trees in the profile. We refer to [14, Section 11.5] for a precise description, and only mention here that the deviation from a true tree form is measured by the number h of reticulations of the resulting network. Trees have no reticulation, that is, $h = 0$.

   Now notice that a hybridization network displays every triple that is displayed by any input tree. By contracting some edges on this network we can decrease reticulation, coming closer to a tree-network, at the cost of having to "take a stance" by resolving some rooted triples. Our results show that certain regularity, Pareto and stability conditions cannot be satisfied by a tree (that is, a non-reticulated network). It is natural to ask how reticulated consensus networks need to be, in order to behave in a stable fashion.
   Q. Find a bound on the required number of reticulations of a regular, extension stable consensus network (for example, as a function of the number of leaves of the trees of the input profile).
3. Since performing inference often leads to a collection of different phylogenetic trees for the set of taxa under investigation, consensus methods are mostly applied to aggregate the given data in this context [14]. Hence both the extension of consensus methods towards reasonable choices of random variables as well as their axiomatic limitations are an interesting avenue for future research. We have only considered deterministic consensus methods (cf. Definition 1.1), and thus our impossibility results do not apply to probabilistic consensus methods such as greedy consensus,

see for example [9]. In particular we leave it as an interesting question to determine whether there exist probabilistic consensus methods satisfying (probabilistic) analogues of Pareto and stability properties.

4. Our methods combine theoretical reductions with explicit computations, and are in principle adaptable to explore other stability conditions that appeared in the literature, for instance in [4]. Even if such conditions are sometimes dependent on each other (for example, our Theorem 2.4 implies incompatibility of condition (I6) in [4] with regularity assumptions, even for binary trees), this approach is in its essence case-by-case. From a systematic point of view, we believe that it would be interesting to shed more light on general conceptual or structural obstructions to the existence of consensus methods with given properties. A possible instance of work in this direction is the following. Recall that an associatively stable regular consensus method is equivalent to a partial order on the set of inputs (for example, phylogenetic trees on n taxa) that is a meet-semilattice [12, Chapter 1, p. 7] and equivariant with respect to the action on the permutation group S_n on the labels of taxa. The existence of such a consensus method that is Pareto on rooted triples is then equivalent to the existence of a (S_n-equivariant) meet-semilattice structure $\vee_?$ on the set of phylogenetic trees such that, for all trees T_1, T_2,

$$F(T_1 \vee_? T_2) \geqslant F(T_1) \vee F(T_2) \tag{2}$$

where $F : RP(X) \to \prod_{A \subseteq [n], |A|=3} RP(A)$ is the canonical function that associates with a tree T the set of rooted triples $(T|_A)_{A \subseteq [n], |A|=3}$, and $\leqslant$ is the natural partial order (resp. $\vee$ the meet operation) in the codomain of F. In fact, if $|A| = 3$, $RP(A)$ has a natural semilattice structure (the only one on four elements that is not a lattice and not a chain) and this defines naturally a "product" semilattice structure on $\prod_{A \subseteq [n], |A|=3} RT(A)$, see for example [17, Chapter 3].

Q. Find obstructions to the existence of an S_n-equivariant meet-semilattice structure on $RP(X)$ that satisfies Equation 2.

Appendix A. Proofs

In this appendix, we describe the computer programs used to prove Theorem 2.4 and Lemma 2.9. Source code is freely available online, see [10]. We will make use of standard terminology from group theory (see for example [16]). When discussing associative stability, we will also make use of standard terminology from the theory of partially ordered sets (see for example [7]).

Given a finite set X, let us denote by $\mathcal{S}(X)$ the symmetric group on X. Then $\mathcal{S}(X)$ acts naturally on the set $RP(X)$, by permuting the labels of the leaves. Given a permutation $\sigma \in \mathcal{S}(X)$ and a tree $T \in RP(X)$, we write $\sigma(T)$ for the tree obtained from T by permuting the labels of the leaves according to σ. In particular, this induces an action of the symmetric group on the set $RBP(X)$ of rooted binary phylogenetic trees. The neutrality axiom can be restated by saying that the k-consensus functions $\varphi_X^k \colon RP(X)^k \to RP(X)$ should be equivariant with respect to the action of $\mathcal{S}(X)$.

A.1. **Extension stability on binary trees.** As was discussed just after the Theorem's statement, in order to prove Theorem 2.4 it is enough to check that there is no extension stable 2-consensus method among profiles of binary trees on a set X of 5 taxa. Since there is only a finite number of consensus methods on X, it is possible (at least in principle) to check every such consensus method by means of a computer program. The number of 2-consensus functions $\varphi_X^2 \colon \mathsf{RBP}(X)^2 \to \mathsf{RP}(X)$ is, however, intractably large already for a set X of cardinality 5. In order to obtain an answer in a reasonable amount of time, we formulate our problem in the context of integer linear programming (see [15]).

Consider the following set $\mathcal{T}$ of triples of phylogenetic trees:

$$\mathcal{T} = \{(T, T_1, T_2) \in \mathsf{RP}(Y) \times \mathsf{RBP}(Y) \times \mathsf{RBP}(Y) \mid Y \subseteq X\}.$$

For every triple $(T, T_1, T_2) \in \mathcal{T}$, we introduce a boolean variable $m_{T,T_1,T_2} \in \{0, 1\}$. Denote by $\mathcal{M}$ the set of all these boolean variables. A 2-consensus method φ on X (restricted to profiles of binary trees) corresponds to the following assignment of the variables in $\mathcal{M}$:

$$m_{T,T_1,T_2} = \begin{cases} 1 & \text{if } \varphi(T_1, T_2) = T; \\ 0 & \text{otherwise.} \end{cases}$$

Conversely, an assignment of the variables in $\mathcal{M}$ yields a 2-consensus method on X (restricted to binary trees), provided that the following linear relations are satisfied:

$$\sum_{T \in \mathsf{RP}(Y)} m_{T,T_1,T_2} = 1 \quad \text{for all } T_1, T_2 \in \mathsf{RBP}(Y), \text{ for all } Y \subseteq X.$$

Our aim is now to enforce all the requirements for our consensus method by means of linear equalities or inequalities involving the variables in $\mathcal{M}$.

(1) Unanimity is equivalent to the following set of direct assignments:

$$m_{T,T,T} = 1 \quad \text{for all } T \in \mathsf{RBP}(Y), \text{ for all } Y \subseteq X.$$

(2) Anonymity is enforced as follows:

$$m_{T,T_1,T_2} = m_{T,T_2,T_1} \quad \text{for all } (T, T_1, T_2) \in \mathcal{T}.$$

(3) Neutrality is given by:

$$m_{T,T_1,T_2} = m_{\sigma(T),\,\sigma(T_1),\,\sigma(T_2)} \quad \text{for all } (T, T_1, T_2) \in \mathcal{T}, \text{ for all } \sigma \in \mathcal{S}(X).$$

(4) Extension stability is slightly more complicated to encode. Consider any triple $(T, T_1, T_2) \in \mathcal{T}$, and let Y be the set of leaves of T. For every subset $Z \subsetneq Y$, and for every tree $T' \in \mathsf{RP}(Z)$ such that $T' \not\preceq T|_Z$, we require that

$$m_{T,T_1,T_2} + m_{T',\,T_1|_Z,\,T_2|_Z} \leqslant 1.$$

The reason is that, if $m_{T,T_1,T_2} = m_{T',\,T_1|_Z,\,T_2|_Z} = 1$, then

$$\varphi(T_1|_Z, T_2|_Z) = T' \not\preceq T|_Z = \varphi(T_1, T_2)|_Z$$

which violates extension stability. Conversely, a violation of extension stability translates into having $m_{T,T_1,T_2} = m_{T',\,T_1|_Z,\,T_2|_Z} = 1$ for some trees T, T', T_1, T_2 with $T' \not\preceq T|_Z$.

Remark A.1. In a practical implementation, the equalities given in (1), (2), and (3) can be used to greatly reduce the number of variables involved in the model. Indeed, instead of using one boolean variable m_{T,T_1,T_2} for every triple $(T, T_1, T_2) \in \mathcal{T}$, we use one variable for every orbit $[T, T_1, T_2] \in \mathcal{T} / (\mathcal{S}(X) \times \mathbb{Z}_2)$. Here the action of $\mathcal{S}(X) \times \mathbb{Z}_2$ on $\mathcal{T}$ is as follows: a permutation $\sigma \in \mathcal{S}(X)$ maps (T, T_1, T_2) to $(\sigma(T), \sigma(T_1), \sigma(T_2))$; the generator of $\mathbb{Z}_2$ maps (T, T_1, T_2) to (T, T_2, T_1). From a practical point of view, this means that we choose a representative for each orbit, and rewrite all the linear constraints in terms of variables m_{T,T_1,T_2} where (T, T_1, T_2) is the representative of its orbit. In addition, thanks to (1), we can remove all the variables of the form $m_{T,T',T'}$ for $T \neq T'$, because their value must be 0. The optimizations described here are essential, in order to make the number of variables tractable.

After optimizations, we obtain a model consisting of 11,688 boolean variables. We use the solver Gurobi [13] to check that there exists no assignment of the variables that satisfies all the previous constraints. Our program runs in approximately 4 minutes on a laptop with an Intel Core i7 processor (8×2.80 GHz) and 16 GB of RAM. This running time includes both the computation of the model and the proof of infeasibility.

A.2. **Associative stability.** Again as discussed after the Theorem's statement, in order to prove Theorem 2.8 it is sufficient to show that there exists no regular associative 2-consensus function on a set X of 5 taxa which is Pareto on rooted triples. Just as in the case of extension stability, we formulate our problem in the context of integer linear programming.

The set of triples that we need to consider is simply $\mathcal{T} = \mathsf{RP}(X)^3$ in this case. For every triple $(T, T_1, T_2) \in \mathcal{T}$, we introduce a boolean variable $m_{T,T_1,T_2} \in \{0, 1\}$ with the same meaning as in the previous section. Again, denote by $\mathcal{M}$ the set of all these boolean variables.

As before, we need to express the fact that to every input (T_1, T_2) corresponds a unique output T. We ensure this by requiring the following linear relations to be satisfied:

$$\sum_{T \in \mathsf{RP}(X)} m_{T,T_1,T_2} = 1 \quad \text{for all } T_1, T_2 \in \mathsf{RP}(X).$$

Assignments of the variables in $\mathcal{M}$ satisfying the previous relations are in one-to-one correspondence with 2-consensus functions $\varphi^2_X \colon \mathsf{RP}(X)^2 \to \mathsf{RP}(X)$.

Unanimity, anonymity, and associative stability of φ^2_X, are equivalent to $\mathsf{RP}(X)$ being endowed with a partial order relation $\leqslant$ (not to be confused with the previously defined $\preceq$), such that every pair of trees $T_1, T_2 \in \mathsf{RP}(X)$ has a unique greatest lower bound, given precisely by the tree $\varphi^2_X(T_1, T_2)$, see [12, Chapter 1, p. 7]. Notice that, in particular, $T_1 \leqslant T_2$ if and only if $\varphi^2_X(T_1, T_2) = T_1$. The validity of the latter expression is represented by the value of the variable m_{T_1,T_1,T_2}; this leads us to introduce new variables

$$p_{T_1,T_2} := m_{T_1,T_1,T_2} \quad \text{for } T_1, T_2 \in \mathsf{RP}(X),$$

with the following meaning:

$$p_{T_1,T_2} = \begin{cases} 1 & \text{if } T_1 \leqslant T_2; \\ 0 & \text{otherwise.} \end{cases}$$

Notice that the variables p_{T_1,T_2} are simply aliases for some variables in $\mathcal{M}$.

We are now ready to translate all requirements for our consensus function into linear constraints.

(1) Reflexive property of the partial order $\leqslant$:

$$p_{T,T} = 1 \quad \text{for all } T \in RP(X).$$

Notice that this set of assignments is equivalent to unanimity.

(2) Antisymmetric property of the partial order $\leqslant$:

$$p_{T_1,T_2} + p_{T_2,T_1} \leqslant 1 \quad \text{for all } T_1, T_2 \in RP(X) \text{ with } T_1 \neq T_2.$$

(3) Transitive property of the partial order $\leqslant$:

$$p_{T_1,T_3} \geqslant p_{T_1,T_2} + p_{T_2,T_3} - 1 \quad \text{for all } T_1, T_2, T_3 \in RP(X).$$

(4) The tree $\varphi^2_X(T_1, T_2)$ must be a lower bound of T_1 and T_2:

$$p_{T,T_1} \geqslant m_{T,T_1,T_2} \text{ and } p_{T,T_2} \geqslant m_{T,T_1,T_2} \quad \text{for all } T, T_1, T_2 \in RP(X).$$

(5) The tree $\varphi^2_X(T_1, T_2)$ must be greater than every lower bound of T_1 and T_2:

$$m_{T,T_1,T_2} + p_{T',T_1} + p_{T',T_2} \leqslant p_{T',T} + 2 \quad \text{for all } T, T', T_1, T_2 \in RP(X).$$

Indeed, the only way to violate this constraint is to set $m_{T,T_1,T_2} = 1$ (that is, $\varphi^2_X(T_1, T_2) = T$), $p_{T',T_1} = p_{T',T_2} = 1$ (that is, T' is a lower bound of T_1 and T_2), and $p_{T',T} = 0$ (that is, $T' \nleqslant T$).

(6) Neutrality:

$$m_{T,T_1,T_2} = m_{\sigma(T),\,\sigma(T_1),\,\sigma(T_2)} \quad \text{for all } T, T_1, T_2 \in RP(X).$$

(7) Pareto property on rooted triples:

$$m_{T,T_1,T_2} = 0 \quad \text{if } T_1|_Y = T_2|_Y \not\preceq T|_Y \text{ for some } Y \subseteq X \text{ with } |Y| = 3.$$

Remark A.2. As for extension stability, in our actual implementation we significantly reduce the number of variables. First, we only use one variable m_{T,T_1,T_2} for every orbit with respect to the action of $\mathcal{S}(X) \times \mathbb{Z}_2$, as described in Remark A.1. We also discard the variables of the form $m_{T,T',T'}$ with $T \neq T'$, because their value must be 0. A further observation is that $T \not< \sigma(T)$ for every $T \in RP(X)$ and $\sigma \in \mathcal{S}(X)$: in fact, if $T < \sigma(T)$, then using neutrality we obtain the contradiction

$$T < \sigma(T) < \sigma(\sigma(T)) < \cdots < \sigma^k(T) = T,$$

where k is the order of σ. Therefore we can discard all the variables m_{T,T_1,T_2} where T is in the same $\mathcal{S}(X)$-orbit of T_1 (respectively T_2) and $T \neq T_1$ (respectively $T \neq T_2$), because their value must be 0 (recall that, if $m_{T,T_1,T_2} = 1$, then $T \leqslant T_1$ and $T \leqslant T_2$). Finally, we can discard in advance all the variables appearing in (7), because their value is 0.

With the optimizations described above, we end up with a model having 15,878 boolean variables. As for extension stability, we use Gurobi [13] to check that there is no assignment of the variables in $\mathcal{M}$ for which all the previous constraints are satisfied. Our program runs in approximately 8 minutes on a laptop with an Intel Core i7 processor (8×2.80 GHz) and 16 GB of RAM.

Appendix B. Consensus methods on small sets of taxa

For every set X of at most 4 taxa there is a unique regular associative consensus function on X that is Pareto on rooted triples, namely Adams consensus. This can be checked using a variant of the program described in the Appendix. The corresponding partial order relation $\leqslant$ discussed earlier is represented in Figures 3 and 4, for $X = \{1, 2, 3\}$ and $X = \{1, 2, 3, 4\}$ respectively.

Notice that Adams consensus on a set X of (at most) 4 taxa also satisfies extension stability, not only on binary trees, see [6]. However, it is not the only consensus method on X which satisfies extension stability on binary trees.

Acknowledgements

Emanuele Delucchi and Linard Hoessly were supported by the Swiss National Science Foundation Professorship grant PP00P2_179110/1.

References

[1] Edward N. Adams. N-trees as nestings: Complexity, similarity, and consensus. *J. Classif.*, 3(2):299–317, Sep 1986.

[2] Alfred V. Aho, Yehoshua Sagiv, Thomas G. Szymanski, and Jeffrey D. Ullman. Inferring a tree from lowest common ancestors with an application to the optimization of relational expressions. *SIAM J. Comput.*, 10(3):405–421, 1981.

[3] Kenneth J. Arrow. A difficulty in the concept of social welfare. *J. Polit. Econ.*, 58(4):328–346, 1950.

[4] Jean-Pierre Barthélémy, Fred R. McMorris, and Robert C. Powers. Stability conditions for consensus functions defined on n-trees. *Math. Comput. Modelling*, 22(1):79–87, 1995.

[5] David Bryant. A classification of consensus methods for phylogenetics. In *Bioconsensus*, volume 61 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 163–183. Amer. Math. Soc., Providence, RI, 2003.

[6] David Bryant, Andrew Francis, and Mike Steel. Can we "future-proof" consensus trees? *Syst. Biol.*, 66(4):611–619, 2017.

[7] Brian A. Davey and Hilary A. Priestley. *Introduction to lattices and order*. Cambridge university press, 2002.

[8] William H. E. Day and Fred R. McMorris. *Axiomatic Consensus Theory in Group Choice and Biomathematics*. Frontiers in Applied Mathematics. Society for Industrial and Applied Mathematics, 2003.

[9] James H. Degnan, Michael DeGiorgio, David Bryant, and Noah A. Rosenberg. Properties of consensus methods for inferring species trees from gene trees. *Syst. Biol.*, 58(1):35–54, 2009.

[10] Emanuele Delucchi, Linard Hoessly, and Giovanni Paolini. Phylogenetic consensus. `https://github.com/giove91/phylogenetic-consensus`, 2018. GitHub repository.

[11] Jianrong Dong, David Fernández-Baca, and Fred R McMorris. Constructing majority-rule supertrees. *Algorithms for Molecular Biology*, 5(1):2, 2010.

[12] George Grätzer. *General lattice theory*. Birkhäuser Verlag, Basel-Stuttgart, 1978. Lehrbücher und Monographien aus dem Gebiete der Exakten Wissenschaften, Mathematische Reihe, Band 52.

[13] Gurobi Optimization LLC. Gurobi optimizer reference manual, 2018.

[14] Daniel H. Huson, Regula Rupp, and Celine Scornavacca. *Phylogenetic networks*. Cambridge University Press, 2010.

[15] Christos H. Papadimitriou and Kenneth Steiglitz. *Combinatorial optimization: algorithms and complexity*. Courier Corporation, 1998.

[16] Joseph J. Rotman. *An introduction to the theory of groups*, volume 148. Springer, 2012.
[17] Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.
[18] Mike Steel. *Phylogeny: Discrete and Random Processes in Evolution*. CBMS-NSF Regional Conference Series on Mathematics. Society for Industrial and Applied Mathematics, 2016.
[19] Mike Steel and Joel D. Velasco. Axiomatic opportunities and obstacles for inferring a species tree from gene trees. *Syst. Biol.*, 63(5):772–778, 2014.

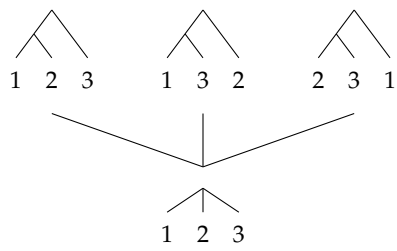

FIGURE 3. Partial order $\leqslant$ associated with Adams consensus, for $X = \{1, 2, 3\}$. The consensus tree of T_1 and T_2 is the highest common descendant of T_1 and T_2.

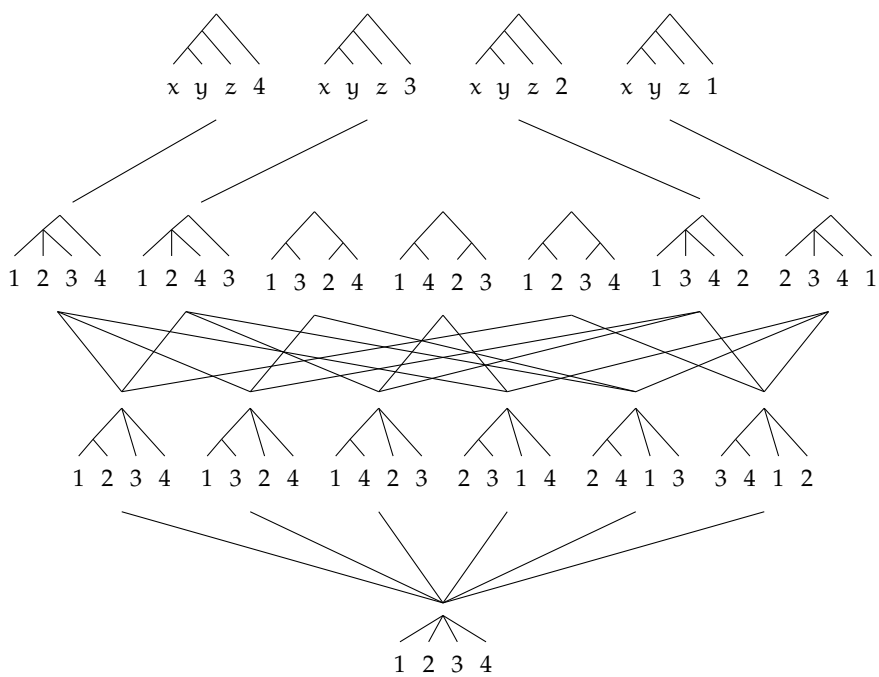

FIGURE 4. Partial order $\leqslant$ associated with Adams consensus, for $X = \{1, 2, 3, 4\}$. Each of the four elements on the top represents three different trees, obtained by choosing the values of x, y, z in all possible ways.