

Biodiversity, Shapley value and phylogenetic trees: some remarks

Hubert Stahn

► **To cite this version:**

Hubert Stahn. Biodiversity, Shapley value and phylogenetic trees: some remarks. Journal of Mathematical Biology, Springer Verlag (Germany), 2020, 80 (3), pp.717-741. 10.1007/s00285-019-01439-z . hal-02417290

HAL Id: hal-02417290

<https://hal-amu.archives-ouvertes.fr/hal-02417290>

Submitted on 5 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Biodiversity, Shapley value and phylogenetic trees: some remarks

Hubert Stahn¹

Abstract

This paper explores the main differences between the Shapley values of a set of taxa introduced by Haake et al. (J Math Biol 56:479–497, 2007. <https://doi.org/10.1007/s00285-007-0126-2>) and Fuchs and Jin (J Math Biol 71:1133–1147, 2015. <https://doi.org/10.1007/s00285-014-0853-0>), the latter having been found identical to the Fair Proportion index (Redding and Mooers in Conserv Biol 20:1670–1678, 2006. <https://doi.org/10.1111/j.1523-1739.2006.00555.x>). In line with Shapley (in: Kuhn, Tucker (eds) Contributions to the theory of games, volume II, annals of mathematics studies 28, Princeton University Press, Princeton, 1953), we identify the cooperative game basis for each of these two classes of phylogenetic games and use them (i) to construct simple formulas for these two Shapley values and (ii) to compare these different approaches. Using the set of weights of a phylogenetic tree as a parameter space, we then discuss the conditions under which these two values coincide and, if they are not the same, revisit Hartmann's (J Math Biol 67:1163–1170, 2013. <https://doi.org/10.1007/s00285-012-0585-y>) convergence result. An example illustrates our main argument. Finally, we compare the species ranking induced by these two values. Considering the Kendall and the Spearman rank correlation coefficient, simulations show that these rankings are strongly correlated. These results are consistent with Wicke and Fischer (J Theor Biol 430:207–214, 2017. <https://doi.org/10.1016/j.jtbi.2017.07.010>), who reach similar conclusions with a different simulation method.

Keywords Biodiversity · Phylogenetic trees · Shapley value · Fair Proportion index

Mathematics Subject Classification 92B99 · 91A12 · 05C05

I wish to thank three anonymous referees for their helpful comments on the early versions of this paper. The usual disclaimer applies concerning the remaining errors. Financial support from the Labex AMSE (ANR-11-IDEX-0001-02) and the ANR GREEN-Econ (ANR-16-CE03-0005) are gratefully acknowledged.

✉ Hubert Stahn
hubert.stahn@univ-amu.fr

¹ Aix-Marseille University (Aix-Marseille School of Economics), CNRS, EHESS & ECM, AMSE, Château Lafarge, 50 Chemin du château Lafarge, 13290 Les Milles, France

1 Introduction

In recent years, biodiversity conservation programs have been placing increasing emphasis on a notion from cooperative game theory: the Shapley value.¹ Shapley (1953), asked a very simple question: assuming that a group of individuals shares a common goal whose outcome is measurable, how can we evaluate the contribution of each individual to this specific objective? Where conservation programs are concerned, this question becomes: how can we evaluate the contribution of a species to overall biodiversity and how can we organize a conservation policy under a limited budget (Weitzman 1998) by targeting specific species?

To answer this question from a game theoretical point of view, Shapley (1953) looks at the outcome that each sub-group can achieve by itself, i.e. the characteristic function, and deduces, under several axioms, a unique imputation rule which specifies the individual contributions. These axioms depict a set of acceptable restrictions on the imputation rule. He assumes that (i) the outcome of the largest group is attributed to the members (efficiency), (ii) two individuals who contribute in the same way to each particular subgroup receive the same reward (symmetry) and (iii) an individual with a zero contribution to each subgroup receives nothing (null-player). He also introduces a more technical, but nevertheless important, assumption (iv) which states that the imputation rule is additive with respect to the characteristic functions. This means that the rewards obtained in a game resulting from the sum of two characteristic functions is the sum of the rewards linked to each of these characteristic functions.

If we now move back to biodiversity conservation problems, this Shapley metric provides an estimate of the contribution of one particular species to the overall phylogenetic diversity of a set of taxa. Its application simply requires the construction of the characteristic function of this game, i.e. the biodiversity index that a subset of taxa achieves by itself. The answer to this question can be found in Faith's (1992) seminal contribution. Each set of taxa has a phylogenetic diversity measure whose definition can be applied recursively in order to obtain the biodiversity index of each subgroup. The Shapley axiomatic does the rest. It provides a measure of the contribution of each species to the global biodiversity index.

This biodiversity measure, which gives rise to several applications, was also studied from a theoretical point of view (Haake et al. 2007; Hartmann 2013; Fuchs and Jin 2015; Wicke and Fischer 2017). However, comparing the early contribution of Haake et al. (2007) with the more recent work of Fuchs and Jin (2015) yields the impression that they are working with two different Shapley values. Wicke and Fischer (2017) provide a discussion of these different definitions. Of course, the first paper considers unrooted trees while the second introduces rooted trees. But this difference is not crucial, since any rooted binary tree can always be transformed into an equivalent unrooted tree by deleting the root and its two incident edges, and re-connecting the

¹ For recent applications see for instance Cadotte et al. (2010), Martyn et al. (2012), Redding and Mazel (2014), Volkmann et al. (2014), Jensen et al. (2016).

two vertices with a new edge. So if the same operator turning a phylogenetic tree into a characteristic function is applied, the uniqueness of the Shapley imputation would induce the same values. This clearly suggests that these two papers do not use the same characteristic function. Notably, they do not use the same definition of the phylogenetic diversity of a subgroup. This issue, evoked by Crozier et al. (2005) and Faith (2006), is related to the inclusion of the root in the computation of this phylogenetic diversity. Actually, Haake et al. (2007) consider the subtree spanned by the sub-group of taxa, while Fuchs and Jin (2015) include the root in this sub-group. This clearly raises several questions. What is the real difference between these two Shapley values? Can we quantify this difference, and under which restrictions do the two values coincide? Since Fuchs and Jin (2015) show that their Shapley value is equivalent to the Fair Proportion index (Redding and Mooers 2006), is the Shapley value introduced by Haake et al. (2007) close to the Fair Proportion index, as suggested by Hartmann (2013)? Furthermore when used in a prioritization problem, do these two values predict drastically different rankings or are they, as suggested by Wicke and Fischer (2017), reasonably correlated?

This paper attempts to answer these questions. To do so, we perform a preliminary step in which we explicitly provide simple formulas for these two Shapley values, as in Haake et al. (2007) but unlike Fuchs and Jin (2015), who simply show the equivalence to the Fair Proportion index. While this step might appear redundant, our approach, based on Shapley's original proof, provides a unified method which simplifies comparisons. Actually, there are three steps to his argument. He first constructs a linearly independent family of cooperative games which spans the set of all games. He then shows that each member of this family induces, under axioms (i)–(iii), a unique distribution of the individual contributions. Finally, he extends his result to all games by observing, under axiom (iv), that the imputation rule is a linear operator. Our approach uses the same rationale. We identify the basis of the subspace of phylogenetic games, compute the contribution to biodiversity of each taxon for games that are part of this basis and extend, by linearity, this measure to the set of all phylogenetic games. This method has several advantages.

It first provides a simple way to compute these two Shapley values by pointing out the differences between the two approaches. Haake et al. (2007) use unrooted trees and therefore do not, by construction, include the root in the computation of the phylogenetic diversity of a set of taxa, while Fuchs and Jin (2015), who use rooted trees, include it. This basic distinction leads to two different subsets of potential characteristic functions, hence to two different bases in the game space and to two distinct Shapley values. This computation exercise also recovers, using a different method, the equality between the Shapley value and the Fair Proportion index obtained by Fuchs and Jin (2015).

Our method also provides a natural parametrization of the two sets of potential characteristic functions. In fact, the two bases that we identify are mainly related to the split structure of the tree, while the potential characteristic functions are obtained by a linear combination of the components of these bases, the weights being those for each edge of the phylogenetic tree under consideration. This means, under the linearity assumption, that the two Shapley values are linear in weight, which incidentally provides a natural way to compute the difference and to compare the two values. We notably identify the linear subspace of weights for which the two Shapley values are

the same. We even show that this linear subspace intersects the interior of the positive orthant, meaning that the result extends to the case in which the weights are assumed to be non-negative. When the two Shapley values are different, we revisit the result of Hartmann (2013), who suggests that the contribution of each edge to individual biodiversity converges, as the number of taxa increases, to the contribution of each edge to the Fair Proportion index. This convergence result was revisited more recently by Fuchs and Paningbatan (2019). They consider a random tree model generated by a β -splitting model and show that the correlation between these two values goes to 1 as the number of taxa increases.

Finally, a direct computation of these two Shapley values also enables us to examine the species ranking they induce, especially their degree of correlation when the two metrics are known to be different. This problem is more empirical and requires simulations. Due to Hartmann's result (2013), we only consider trees with a low number of leaves and randomly select the split structures for each set of species. Weights are also chosen randomly but in the subset in which the two Shapley values are known to be different. We then compute the rank correlation coefficients between the species ranking induced by the two Shapley values. We especially consider the Tau statistic (Kendall 1938) and the Rho statistic (Spearman 1904). A kernel estimate of the distributions of these correlation coefficients suggests that the rankings are strongly correlated. These simulation results support those of Wicke and Fischer (2017), who reach the same conclusion with a different method.

Our argument will be organized as follows. Section 2 reviews the main notations concerning phylogenetic trees and recalls Shapley's result. Section 3 presents a simple example illustrating the main steps of the argument. Sections 4 and 5 are devoted to an explicit derivation of the Shapley value for phylogenetic trees when, respectively, the phylogenetic diversity is or is not independent of the root. Section 6 compares the two Shapley values and shows under which conditions they are the same. Section 7 explores the impact of the species ranking provided by the two metrics in the case in which they are known to be different. Section 8 concludes the present manuscript.

2 Notations and preliminary results

In this section, we briefly review the notion of phylogenetic trees and recall the main results related to the notion of the Shapley value.

2.1 Phylogenetic trees

A phylogenetic tree, T , of n taxa is a *weighted binary tree* whose set $I = \{1, \dots, n\}$ of leaves is identified with the n taxa. Its graph $(V, E, w(\cdot))$ is composed of (i) a set V of vertices, $v \in V$, including the n taxa, (ii) a set E of edges which describes the adjacent vertices $e = \{v, v'\} \in E$ with $v \neq v'$, and (iii) a map $w : E \rightarrow \mathbb{R}$ which assigns a weight to each edge $e \in E$. Being a tree, this graph is connected and free of cycles so that there exists a unique path $\{v \rightarrow v'\} \subset E$ between two vertices. The number of vertices adjacent to vertex v is called the degree of v , $\deg(v)$. The degree of

each taxon, i.e. of an *external vertex*, is 1. If the degree of all the other vertices called *internal* is 3, the tree is said to be an unrooted binary tree. If there exists at most one vertex of degree 2, called the root, ρ , while the others are either of degree 1 or 3, the tree is said to be a rooted binary tree.

Each unrooted binary tree is usually non-oriented and contains $(2n - 3)$ edges. This tree can be split into two subtrees by removing one edge e . This induces a partition of the set of leaves/taxa, $s_e = \{I_e \mid \bar{I}_e\}$. We call $S_T = \{s_e\}_{e \in E}$ the split structure of a tree T . Now let S be any set of $(2n - 3)$ different bipartitions of I . This set is called a pairwise compatible split structure if for each $s = \{I \mid \bar{I}\}$ and $s' = \{I' \mid \bar{I}'\}$ at least one of these sets, $I \cap I', I \cap \bar{I}', \bar{I} \cap I'$ or $\bar{I} \cap \bar{I}'$ is empty. In this case by the Buneman theorem (1971) and its consequences for phylogenetics (Semple and Steel 2003), the split equivalence theorem tell us that an unrooted binary tree can be characterized by its pairwise compatible split structure $S_T = \{s_e\}_{e \in E}$. In contrast, a rooted binary tree is usually oriented, since there exists a root ρ , and contains $(2n - 2)$ edges. The orientation makes it possible to associate with each edge e the set I_e of leaves/taxa descending from this edge. This set is called the cluster of edge e .

2.2 Cooperative games and Shapley value

Let us now consider sub-groups of the set, $I = \{1, \dots, n\}$, of taxa. Each sub-group of taxa or *coalition*, C , belongs to 2^I , the set of 2^n subsets of I . By convention, \emptyset and I are respectively called the *empty* and the *grand* coalition. Cooperative games assign a score to each coalition, i.e. an additive measure, which evaluates the benefit that this group can achieve by itself. This *characteristic function* $v : 2^I \rightarrow \mathbb{R}$ associates each coalition C with a real, $v(C)$. By convention, the image of the empty set is zero, i.e. $v(\emptyset) = 0$. Given this characteristic function, these games aim to propose an *imputation* $\phi_v : I \rightarrow \mathbb{R}$ measuring the contribution, $\phi_v(i)$, of each individual to the score, $v(I)$, of the grand coalition. Since the set of potential imputations is large, Shapley (1953) adds additional restrictions on this mapping. He first requires that this imputation distributes the wealth obtained by the grand coalition, implying that nothing is lost. This efficiency axiom says that:

Axiom 1 (Efficiency) $\sum_{i \in I} \phi_v(i) = v(I)$.

In addition to this axiom, Shapley (1953) requires that two individuals who contribute in the same way to every coalition obtain the same imputation. This symmetry axiom is given by:

Axiom 2 (Symmetry) If for all $C \subseteq I \setminus \{i, j\}$, $v(C \cup \{i\}) = v(C \cup \{j\})$ then $\phi_v(i) = \phi_v(j)$.

He also assumes that an individual contributing to no coalition receives no payment. In fact, he says:

Axiom 3 (Null Player) If for all $C \subseteq I \setminus \{i\}$, $v(C \cup \{i\}) = v(C)$ then $\phi_v(i) = 0$.

Finally, Shapley (1953) adds a somewhat more technical assumption. He requires that the imputation obtained from the sum of two characteristic functions simply be the sum of the two initial imputations:

Axiom 4 (Additivity) For all v_1, v_2 , and all $i \in I$, $\phi_{(v_1+v_2)}(i) = \phi_{v_1}(i) + \phi_{v_2}(i)$.

Under these four axioms, he shows that:

Proposition (Shapley 1953) *There exists a unique imputation rule, $Sh_v(i)$, called the Shapley value which satisfies these four axioms.*

His proof is essentially based on the idea that the space of characteristic functions can be spanned by a family of linearly independent games (winning coalitions) with the property that each of these characteristic functions induces, under Axioms 1–3, a unique imputation. Since these imputations also satisfy a scalar multiplication property, the additivity axiom extends this preliminary result, in a unique way, from this independent family of games to the set of all characteristic functions. Finally, he supplements his main result by providing one of the most popular ways to compute this value² using an average value of the marginal contribution of each individual to the different coalitions, i.e.:

$$Sh_v(i) = \sum_{C \subseteq I, i \in C} \frac{(|C|-1)!(n-|C|)!}{n!} (v(C) - v(C \setminus \{i\})) \quad (1)$$

The probability distribution used in this computation is based on the idea that individuals are randomly ordered. Given his rank, player i forms a coalition with the previously ranked players and receives a gain corresponding to his marginal contribution to the coalition formed. The probability that when he enters he will find coalition $C \setminus \{i\}$ there already is $\frac{(|C|-1)!(n-|C|)!}{n!}$.

3 A preliminary example

This illustration is built on two very simple trees involving a set of three taxa, $I = \{1, 2, 3\}$. The first tree, T^u , is unrooted while the second, T^r , has a root. To spare some notations, we identify, for the moment, each edge with the split induced by the removal of this edge and assign a weight to each split. Knowledge of these two elements provides all the information relevant to our illustration. The tree, i.e. the split structure, is taken as given and the weights are used as parameters. Figure 1 summarizes these choices.

Let us now move to the game space. Since $I = \{1, 2, 3\}$, the set of all coalitions (i.e. the subsets of taxa) is given by:

$$2^I = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\} \quad (2)$$

where \emptyset and $\{1, 2, 3\}$ are called the *empty* and the *grand* coalition. A characteristic function $v : 2^I \rightarrow \mathbb{R}$ can therefore be viewed as a vector $(v(C))_{C \in 2^I} \in \mathbb{R}^8$ with, by convention, $v(\emptyset) = 0$. In other words, the set of potential characteristic functions studied by Shapley (1953) is, in this example, a linear subspace of dimension 7 in \mathbb{R}^8 .

² For other expressions of the Shapley value see for instance Kleinberg and Weiss (1985) or Rothblum (1988).

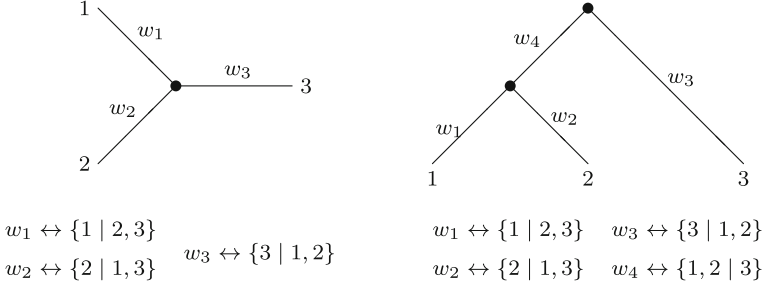


Fig. 1 Unrooted and rooted tree with associated weights and splits

But in phylogenetics, these characteristic functions are deduced from trees by using a phylogenetic diversity measure. For the unrooted trees, this measure, $PD^u(C)$, is given by the sum of the weights of the edges contained in the minimal subtree spanned by coalition C while, for the rooted trees, this quantity, $PD^r(C)$, is the sum of the weights, this time on the minimal subtree spanned by coalition C and the root ρ . To compute these two vectors $(PD^u(C))_{C \in 2^I}$ and $(PD^r(C))_{C \in 2^I}$, we maintain the ordering of the coalitions given in Eq. (2). In this case, these quantities write:

$$(PD^u(C))_{C \in 2^I} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ w_1 + w_2 \\ w_1 + w_3 \\ w_2 + w_3 \\ w_1 + w_2 + w_3 \end{pmatrix}, (PD^r(C))_{C \in 2^I} = \begin{pmatrix} 0 \\ w_1 + w_4 \\ w_2 + w_4 \\ w_3 \\ w_1 + w_2 + w_4 \\ w_1 + w_3 + w_4 \\ w_2 + w_3 + w_4 \\ w_1 + w_2 + w_3 + w_4 \end{pmatrix} \quad (3)$$

Taking the weights as parameters, we also observe that the game spaces \mathcal{P}^u and \mathcal{P}^r associated with, respectively, unrooted and rooted trees are generated by the following bases of characteristic functions:

$$\mathcal{P}^u = \left\langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \text{ and } \mathcal{P}^r = \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \quad (4)$$

$\{1|2,3\}$
 $\{2|1,3\}$
 $\{3|1,2\}$
 $\{1|2,3\}$
 $\{2|1,3\}$
 $\{3|1,2\}$
 $\{1,2|3\}$

We even observe that each vector composing these bases is associated with a particular split which characterizes the topology of each tree. To understand this, remember that each split is obtained by removing an edge, say e_0 . We can therefore take the initial tree structure, either in the unrooted or in the rooted case, and set all the weights to

0, except for edge e_0 whose weight is 1, and compute characteristic function $b(e_0)$ associated with this new tree by taking the suitable PD function.³ By repeating the argument for each element of the split structure of the two trees, we find the two bases of Eq. (4). This means that the dimension of the characteristic function space can be related to the number of splits or edges and is respectively 3 and 4 for \mathcal{P}^u and \mathcal{P}^r . This clearly shows that these game subspaces differ somewhat, depending on whether trees are unrooted or rooted.

When we move to the computation of the Shapley values, it becomes important to remember that the Shapley operator is a linear operator from the game space into the value space, here \mathbb{R}^3 . This means that we only need to compute the images of the vectors composing each basis to obtain what Haake et al. (2007) called the transformation matrices. We later provide simpler formulas to compute these values; however, at that point, the reader can simply use the Shapley value given by Eq. (1). In our example, these transformation matrices are:

$$M_{T^u} = \frac{1}{6} \begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix} \text{ and } M_{T^r} = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & 0 \end{bmatrix} \quad (5)$$

where, for instance, the first column of, say, M_{T^u} is composed of the Shapley values of taxon $i = 1, \dots, 3$ obtained with the characteristic function associated with the split $\{1 | 2, 3\}$. Moreover, by the linearity of the Shapley operator, the Shapley values of taxon $i = 1, \dots, 3$ for the two phylogenetic trees depicted in Fig. 1 are:

$$(Sh^u(i))_{i=1}^3 = \frac{1}{6} \begin{pmatrix} 4w_1 + w_2 + w_3 \\ w_1 + 4w_2 + w_3 \\ w_1 + w_2 + 4w_3 \end{pmatrix} \text{ and } (Sh^r(i))_{i=1}^3 = \frac{1}{2} \begin{pmatrix} 2w_1 + w_4 \\ 2w_2 + w_4 \\ 2w_3 \end{pmatrix} \quad (6)$$

Let us now compare these two Shapley values. This calls for equivalent trees. This is why we introduce, in line with Wicke and Fischer (2017), the notion of an unrooted rooted tree. This operation associates each rooted tree with an equivalent unrooted tree. This new tree, T^{ur} , is obtained by deleting the two incident edges to the root and by replacing these missing edges by a single edge whose weight is the sum of the weights of the two initial edges. In our example (Fig. 1), the unrooted rooted tree is obtained by taking the unrooted tree and replacing w_3 by $w'_3 = (w_3 + w_4)$. This means that the Shapley values of our taxa for the equivalent unrooted rooted tree are:

$$(Sh^{ur}(i))_{i=1}^3 = \frac{1}{6} \begin{pmatrix} 4w_1 + w_2 + (w_3 + w_4) \\ w_1 + 4w_2 + (w_3 + w_4) \\ w_1 + w_2 + 4(w_3 + w_4) \end{pmatrix} = \frac{1}{6} \underbrace{\begin{bmatrix} 4 & 1 & 1 & 1 \\ 1 & 4 & 1 & 1 \\ 1 & 1 & 4 & 4 \end{bmatrix}}_{M_{T^{ur}}} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} \quad (7)$$

³ To perform this computation, take, for instance, our unrooted tree and the split $\{1 | 2, 3\}$ associated with the removal of the edge of weight w_1 . If we set $w_1 = 1$ and $w_2 = w_3 = 0$ in this tree and apply the $PD^u(C)$ function given in Eq. (3), we obtain the first vector of the basis of \mathcal{P}^u [Eq. (4)].

The transformation matrix $M_{T^{ur}}$ associated with an unrooted rooted tree is, in our example, obtained by duplicating the last column of M_{T^u} [see Eq. (6)]. Moreover, since $M_{T^{ur}}$ and M_{T^r} are now two matrices of the same dimension, we can say that:

$$(Sh^{ur}(i))_{i=1}^3 - (Sh^r(i))_{i=1}^3 = \left(\frac{1}{6} \begin{bmatrix} 4 & 1 & 1 & 1 \\ 1 & 4 & 1 & 1 \\ 1 & 1 & 4 & 4 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & 0 \end{bmatrix} \right) \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} \quad (8)$$

$$= \frac{1}{3} \underbrace{\begin{bmatrix} -1 & \frac{1}{2} & \frac{1}{2} & -1 \\ \frac{1}{2} & -1 & \frac{1}{2} & -1 \\ \frac{1}{2} & \frac{1}{2} & -1 & 2 \end{bmatrix}}_A \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} \quad (9)$$

The two Shapley values are therefore the same for all vectors of weights which belong to the kernel of A . Moreover, since both Shapley values satisfy the efficiency axiom, we have that $\varepsilon'_3 \cdot A = 0$, where ε_3 denotes the vector of \mathbb{R}^3 with all entries equal to 1. This means that the kernel is given by the weights satisfying:

$$\begin{bmatrix} -1 & \frac{1}{2} \\ \frac{1}{2} & -1 \end{bmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + \begin{bmatrix} \frac{1}{2} & -1 \\ \frac{1}{2} & -1 \end{bmatrix} \begin{pmatrix} w_3 \\ w_4 \end{pmatrix} = 0 \quad (10)$$

$$\Leftrightarrow \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} \begin{pmatrix} w_3 \\ w_4 \end{pmatrix} \quad (11)$$

At this point, we can conclude that, in our example, the Shapley values computed on a rooted tree or on its associated unrooted rooted tree are the same on a subset of weights of dimension 2 satisfying Eq. (11). The reader will even note that the weights (1, 1, 3, 1) satisfy this condition, meaning that the result remains true if the weights are assumed to be non-negative. The set $\ker(A) \cap \mathbb{R}_+^4$ simply becomes a polytope of dimension 2.

In the rest of this manuscript, we first re-examine the construction of a Shapley value successively for unrooted and rooted trees. We then introduce unrooted rooted trees to compare these values. Finally, we perform simulations to explore how the species rankings change when the two Shapley values are known to be different.

4 Shapley value for unrooted trees

In this case, the phylogenetic diversity, $PD^u(C)$, of a coalition $C \in 2^I$ is given by the sum of the weights of the edges contained in the minimal subtree spanned by coalition C . Since there exists, for any tree, a unique path, $\{i \rightarrow j\} \subseteq E$ which joins each pair $\{i, j\}$ of taxa, the edges of this subtree are simply given by $E_C^u = \cup_{\{i,j\} \subseteq C} \{i \rightarrow j\}$, the union of all paths joining two taxa in coalition C , and the phylogenetic diversity of a coalition C becomes:

$$PD^u(C) = \sum_{e \in E_C^u} w(e) \quad (12)$$

Let us now identify the function $PD^u(C)$, for $C \in 2^I$, as the characteristic function of a cooperative game and denote by \mathcal{P}^u the set of all characteristic functions obtained by changing the weights while keeping the tree unchanged. This set, \mathcal{P}^u , is obviously a subset of the set of all characteristic functions of cooperative games explored by Shapley (1953). We even claim that \mathcal{P}^u is a linear subspace of dimension $(2n - 3)$ corresponding to the number of edges of an unrooted binary tree with n leaves. The intuition behind this result is quite simple. Since the tree structure remains unchanged, we can introduce a family of $(2n - 3)$ new phylogenetic trees. Each of these trees, say of type T_e , is associated with a given edge $e \in E$ and has the property that the weight of edge e is $w(e) = 1$, while for all $e' \neq e$, $w(e') = 0$. With this property, the phylogenetic diversity, $b_e^u(C)$, of a coalition C associated with tree T_e is either 0 or 1 and, from Eq. (12), it is immediate that this quantity is 1 if and only if there exist two taxa in C connected by a path containing edge e . More formally, we can say that:

$$b_e^u(C) = \begin{cases} 1 & \text{if there exists } \{i, j\} \subseteq C \text{ with } e \in \{i \rightarrow j\} \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

Moreover, we know that the removal of edge e from the tree causes the tree to split into two subtrees isolating two subsets of taxa. This partition $s_e = \{I_e \mid \bar{I}_e\}$ produces a new interpretation of Eq. (13) mainly based on the set of taxa. In fact, if there exists a path between two elements of C which contains e , this also means that coalition C has a non-empty intersection with both subsets composing the partition s_e , i.e. Eq. (13) becomes:

$$b_e^u(C) = \begin{cases} 1 & \text{if } I_e \cap C \neq \emptyset \text{ and } \bar{I}_e \cap C \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

It is also immediate that the family, $\{b_e^u\}_{e \in E} \subset \mathcal{P}^u$, of characteristic functions has the property that for each $C \in 2^I$:

$$PD^u(C) = \sum_{e \in E} w(e) b_e^u(C) \quad (15)$$

This last formula is typically a new reading of the split decomposition property of phylogenetic diversity (see for instance Minh et al. (2009), or Volkmann et al. (2014)). In fact, we simply say that each element of the split structure generates a characteristic function in the game space and that the linear combination of these characteristic functions spans the game space. We can even go a step further by showing that:

Lemma 1 $\{b_e^u\}_{e \in E} \subset \mathcal{P}^u$ is a basis of the linear set \mathcal{P}^u . Moreover, the dimension of \mathcal{P}^u is $(2n - 3)$.

Proof Let $(\alpha_e)_{e \in E}$ be a vector of scalars in \mathbb{R}^{2n-3} which satisfies:

$$\sum_{e \in E} \alpha_e b_e^u(C) = 0 \text{ for all } C \in 2^I \quad (16)$$

Now select an edge e with induced split $\{I_e \mid \bar{I}_e\}$. If this edge is internal, construct the partition $\{E_1 \mid e \mid E_2\}$ of the set of edges which isolates e and the non-empty sets of edges, E_1 and E_2 , associated with the subtrees, respectively, spanned by I_e and \bar{I}_e . In this case, we observe that: (i) for each $e' \in E_1$, the associated split $\{I_{e'} \mid \bar{I}_{e'}\}$ has the property that $I_{e'} \subset I_e$ so that, by Eq. (14), $b_{e'}^u(I_e) = 1$, (ii) for each $e' \in E_2$ the associated split $\{I_{e'} \mid \bar{I}_{e'}\}$ verifies $I_e \subset I_{e'}$ so that $b_{e'}^u(I_e) = 0$ and finally (iii) $b_e^u(I_e) = 0$ since I_e cannot intersect \bar{I}_e . It follows from Eq. (16) that:

$$\sum_{e' \in E} \alpha_{e'} b_{e'}^u(I_e) = \sum_{e' \in E_1} \alpha_{e'} = 0 \quad (17)$$

By a symmetric argument, Eq. (16) also implies that:

$$\sum_{e' \in E} \alpha_{e'} b_{e'}^u(\bar{I}_e) = \sum_{e' \in E_2} \alpha_{e'} = 0 \quad (18)$$

Moreover, if we consider the grand coalition, Eq. (14) says that $\forall e' \in E$, $b_{e'}^u(I) = 1$. It follows from Eq. (16):

$$\sum_{e' \in E} \alpha_{e'} b_{e'}^u(I) = \sum_{e' \in E} \alpha_{e'} = 0 \quad (19)$$

By combining Eqs. (17), (18) and (19), we obtain that $\alpha_e = 0$. This argument applies for each internal edge. Moreover, if our selected edge, e , is external we observe that $E_2 = \emptyset$ and $E_1 = E \setminus \{e\}$. This means, in the previous argument, that point (ii) and Eq. (18) are vacuous but Eqs. (17) and (19) now imply that $\alpha_e = 0$. Since the previous point is independent of selection of edge e , we can, by repeating the argument for all $e \in E$, conclude that $(\alpha_e)_{e \in E} = 0$. \square

Remark 1 From Eq. (14), it can be seen that the basis vector $(b_e^u(C))_{C \in 2^I}$, associated with edge e when restricted to coalitions, C , composed of two taxa $\{a, b\} \in I_2 \subset 2^I$, corresponds to the split metric $\sigma^e = \left(\sigma_{\{a,b\}}^e \right)_{\{a,b\} \in I_2}$ associated with edge e extensively studied by Bandelt and Dress (1992). It is, in particular, known that the family $\{\sigma^e\}_{e \in E}$ of split metrics associated with a tree forms a family of $(2n - 3)$ linearly independent vectors (see, for instance, Desper and Gascuel 2005). This result provides an alternative proof of Lemma 1.

The next step in our analysis is the construction of the Shapley value, $Sh_v^u(i)$, of a taxon $i \in I$, using the decomposition of $PD^u(\cdot)$ provided by Eq. (15). This requires a preliminary step consisting in computing the Shapley value, $Sh_{b_e^u}^u(i)$, of a taxon $i \in I$ for each characteristic function, $b_e^u(\cdot)$ composing the basis of the game space. By the standard Shapley formula [see Eq. (1)], this quantity is:

$$Sh_{b_e^u}^u(i) = \sum_{C \subseteq I, i \in C} \frac{(|C|-1)!(n-|C|)!}{n!} (b_e^u(C) - b_e^u(C \setminus \{i\})) \quad (20)$$

Let us now concentrate on the marginal contribution $(b_e^u(C) - b_e^u(C \setminus \{i\}))$ of taxon i to coalition C . This quantity belongs, a priori, to $\{-1, 0, 1\}$. But $b_e^u(C) = 0$ obviously implies by Eq. (14) that $b_e^u(C \setminus \{i\}) = 0$, meaning that this marginal contribution is either 0 or 1. Moreover, using Eq. (14) again, if $i \in I_e$, this quantity is equal to 1 if and only if $C \setminus \{i\}$ is non-empty and a subset of \bar{I}_e , while if $i \in \bar{I}_e$, this occurs if and only if the non-empty set $(C \setminus \{i\}) \subseteq I_e$. We now need a combinatorial analysis to identify the number of occurrences of these cases, concluding after simplification that:

Lemma 2 *If $\{I_e \mid \bar{I}_e\}$ denotes the split associated with edge e , the Shapley value of taxon $i \in I$ for the characteristic function $b_e^u(\cdot)$ is:*

$$Sh_{b_e^u}^u(i) = \left(\mathbf{I}_{i \in I_e} \frac{|\bar{I}_e|}{n |I_e|} + (1 - \mathbf{I}_{i \in I_e}) \frac{|I_e|}{n |\bar{I}_e|} \right) \\ \text{with } \mathbf{I}_{i \in I_e} = \begin{cases} 1 & \text{if } i \in I_e \\ 0 & \text{otherwise} \end{cases} \quad (21)$$

Proof Let us first observe that Eq. (20) can also be written as:

$$Sh_{b_e^u}^u(i) = \sum_{C \subseteq I \setminus \{i\}} \frac{|C|!(n-|C|-1)!}{n!} \underbrace{(b_e^u(C \cup \{i\}) - b_e^u(C))}_{=\Delta_e} \quad (22)$$

Let us now assume that $i \in I_e$. In this case Eq. (14) says that $\Delta_e = 1$ if and only if $C \subseteq \bar{I}_e$ and $C \neq \emptyset$. To verify this point, note that (i) if $C \neq \emptyset$ and $C \subseteq \bar{I}_e$, then $b_e^u(C) = 0$ and $b_e^u(C \cup \{i\}) = 1$ since $i \in I_e$ so that $\Delta_e = 1$ and (ii) if not, either C intersects both I_e and \bar{I}_e so that $b_e(C) = b_e(C \cup \{i\}) = 1$, or $C \subset I_e$ which implies that $b_e^u(C) = b_e^u(C \cup \{i\}) = 0$, i.e. in both cases $\Delta_e = 0$. This preliminary observation clearly says that the terms composing $Sh_{b_e^u}^u(i)$ are different from 0 if and only if $C \subseteq \bar{I}_e$ and $C \neq \emptyset$. Let us now observe that (i) there exist $\binom{|\bar{I}_e|}{|C|}$ potential choices of a subset of $c = |C|$ elements in \bar{I}_e and (ii) the size of this subset, C , can be $c = 1, \dots, |\bar{I}_e|$. It follows that the Shapley value becomes:

$$Sh_{b_e^u}^u(i) = \sum_{c=1}^{|\bar{I}_e|} \frac{c!(n-c-1)!}{n!} \binom{|\bar{I}_e|}{c} = \frac{|\bar{I}_e|!}{n!} \sum_{c=1}^{|\bar{I}_e|} \frac{(n-c-1)!}{(I_e-c)!} \\ = \frac{|\bar{I}_e|!(n-|\bar{I}_e|-1)!}{n!} \sum_{c=1}^{|\bar{I}_e|} \binom{n-c-1}{n-|\bar{I}_e|-1} \quad (23)$$

Moreover, an iterative use of the Pascal formula gives:

$$Sh_{b_e^u}^u(i) = \frac{|\bar{I}_e|!(n-|\bar{I}_e|-1)!}{n!} \binom{n-1}{n-|\bar{I}_e|} \quad (24)$$

and since $|I_e| + |\bar{I}_e| = n$, we have:

$$Sh_{b_e^u}^u(i) = \frac{|\bar{I}_e|}{n(n - |\bar{I}_e|)} = \frac{|\bar{I}_e|}{n|I_e|} \quad (25)$$

Finally, if $i \in \bar{I}_e$ we simply need to apply the same argument by replacing \bar{I}_e by I_e . This gives $Sh_{b_e^u}^u(i) = \frac{|I_e|}{n|I_e|}$ and, combining both results, Eq. (21) follows. \square

To go a step further, let us observe that Eq. (20) is homogeneous of degree 1 with respect to the characteristic function $b_e^u(\cdot)$, meaning that multiplying $b_e^u(\cdot)$ by any scalar $\alpha \in \mathbb{R}$, yields $\forall i \in I$, $Sh_{\alpha \cdot b_e^u}^u(i) = \alpha \cdot Sh_{b_e^u}^u(i)$. Since the construction of the Shapley value also requires additivity (see Axiom 4), it follows that:

$$Sh_{PD^u}^u(i) = Sh^u\left(\sum_{e \in E} w(e)b_e^u\right)(i) = \sum_{e \in E} w(e)Sh_{b_e^u}^u(i) \quad (26)$$

Since $\{b_e^u\}_{e \in E}$ is a basis of \mathcal{P}^u , we even say that this imputation rule is unique, meaning that there exists a unique function which attributes to each phylogenetic characteristic function $PD^u(\cdot)$ individual contributions to biodiversity.⁴ To conclude this discussion, we can say that:

Proposition 1 *Let T^u be an unrooted phylogenetic tree with n leaves whose split structure is given by $S_{T^u} = \{\{I_e | \bar{I}_e\}_{e \in E}\}$. Under the standard Shapley approach based on Axioms 1, 2, 3 and 4, the Shapley value of species i is given by:*

$$Sh_{T^u}^u(i) = \sum_{e \in E} w(e) \underbrace{\left(\mathbf{I}_{i \in I_e} \frac{|\bar{I}_e|}{n|I_e|} + (1 - \mathbf{I}_{i \in I_e}) \frac{|I_e|}{n|\bar{I}_e|} \right)}_{m_{i,e}^{T^u}} \quad (27)$$

with $\mathbf{I}_{i \in I_e} = \begin{cases} 1 & \text{if } i \in I_e \\ 0 & \text{otherwise} \end{cases}$

The previous proposition also shows that, for a given tree T_u , the relation between the Shapley value $(Sh_{T^u}^u(i))_{i=1}^n$ and the weights is linear and is given by the so-called Shapley transformation matrix, $M_{T^u} = [m_{i,e}^{T^u}]_{\substack{i \in I \\ e \in E}}$ of dimension $(n, 2n - 3)$. In particular, if $c(i, e)$ denotes the number of taxa that are on the same side of a split, s_e , as taxon i and $f(i, e)$ denotes its complement (i.e. $f(i, e) = n - c(i, e)$), we reached, with a different proof, the formula proposed by Haake et al. (2007) or Volkmann et al. (2014), the latter highlighting the importance of the split structure. In other words, we can say:

Corollary 1 (Haake et al. 2007, Th. 4) *The Shapley operator is linear in weights, i.e. $Sh_{T^u}^u(\cdot) = M_{T^u} \cdot w(\cdot)$ and the (i, e) th entry of the Shapley transformation matrix, M_{T^u} , of dimension $(n, 2n - 3)$ is given by:*

⁴ This result is not in contradiction with Haake et al. (2007), Th. 6., stating that several weighting structures can induce the same Shapley value. Our uniqueness result simply says that the imputation rule is unique but not that this map is injective.

$$m_{i,e}^{T^u} = \frac{f(i,e)}{n \times c(i,e)} \text{ for } i \in I \text{ and } e \in E \quad (28)$$

This linearity property, which is less clear from the early definition of the Shapley value [see Eq. (1)], comes from (i) the linear decomposition of the characteristic function of the PD^u game with respect to the weights, and (ii) the linearity of the Shapley operator as a function mapping the set of characteristic functions into the imputation space (*i.e.* the vector of Shapley values). These two properties combined show that the Shapley transformation map introduced by Haake et al. (2007) is a linear operator with respect to the weights (for a further discussion on the transformation matrix see Wicke and Fischer 2019) and then the following Remark 2:

Remark 2 In line with Haake et al. (2007), we observe that the Shapley transformation matrix, M_{T^u} , is related to the topology of the tree. In fact a column of this matrix is the Shapley imputation associated with one element of our basis $\{b_e^u\}_{e \in E}$, *i. e.* $\left(m_{i,e}^{T^u}\right)_{i=1}^n = Sh_{b_e^u}^u(\cdot)$. Moreover, each of the games composing this basis is related to a given element of the split structure of the tree.

5 Shapley value for rooted trees

To capture the main distinctions between Shapley values computed on unrooted and rooted trees, let us now perform the same kind of analysis for rooted trees. The main difference appears at the beginning: the definition of the phylogenetic diversity of a coalition is not exactly the same. This induces a different game space with, naturally, different bases and therefore different Shapley values.

Let us first turn back to the definition of phylogenetic diversity of a coalition in the case of rooted trees. This measure remains the sum of the weights along a subtree, but this sum is now computed on the minimal subtree spanned by coalition C and the root, ρ . This means that the relevant edges are those belonging to any path, $\{\rho \rightarrow i\}$, from the root to a taxon $i \in C$. If $E_c^r = \cup_{i \in C} \{\rho \rightarrow i\}$ denotes this set of edges, this new characteristic function writes:

$$PD^r(C) = \sum_{e \in E_c^r} w(e) \quad (29)$$

Both this function, $PD^r(\cdot)$ and the subset, \mathcal{P}^r , of characteristic functions obtained by changing the weights are completely different from $PD^u(\cdot)$ and \mathcal{P}^u introduced in the previous section. For instance, any coalition, $\{i\}$, formed by a single taxon $i \in I$ now has a non-zero phylogenetic diversity. Moreover \mathcal{P}^r is expected to be a linear subspace of dimension $(2n - 2)$, since any rooted binary tree with n leaves contains $(2n - 2)$ edges.

Appart from the Shapley value itself, the method of construction remains the same. We now consider a family of size $(2n - 2)$ of rooted trees each with a unit weight assigned to a specific edge and zero elsewhere. But the derivation of the characteristic functions $\{b_e^r(\cdot)\}_{e \in E}$ induced by these trees changes. For a coalition C , $b_e^r(C)$ is now

equal to 1 if and only if at least one path from the root to a taxon $i \in C$ contains edge $e \in E$, i.e.:

$$b_e^r(C) = \begin{cases} 1 & \text{if there exists } i \in C \text{ with } e \in \{\rho \rightarrow i\} \\ 0 & \text{otherwise} \end{cases} \quad (30)$$

Since the tree is rooted, each edge $e \in E$ can also be associated with the cluster $I_e \subset I$, which contains the subset of all taxa descending from this edge. This observation again leads to a simpler definition of the family $\{b_e^r(\cdot)\}_{e \in E}$, since claiming the existence of a path from the root to a taxon $i \in C$ containing edge e amounts to claiming that coalition C has a non-empty intersection with cluster I_e . It follows that $b_e^r(C)$ becomes:

$$b_e^r(C) = \begin{cases} 1 & \text{if } C \cap I_e \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \quad (31)$$

and it can be shown that:

Lemma 3 $\{b_e^r\}_{e \in E} \subset \mathcal{P}^r$ is a basis of the linear set \mathcal{P}^r . Moreover, the dimension of \mathcal{P}^r is $(2n - 2)$.

Proof Let $(\alpha_e)_{e \in E}$ be a vector of scalars in \mathbb{R}^{2n-2} such that:

$$\sum_{e \in E} \alpha_e b_e^r(C) = 0 \text{ for all } C \in 2^I \quad (32)$$

To show that $(\alpha_e)_{e \in E} = 0$, we proceed by induction, starting from the leaves and moving back to the root of the tree. So let us denote by e_i the external edge leading to leaf $i \in I$ and let I_e be the cluster induced by each edge $e \in E$. From Eq. (31), $b_e^r(I \setminus \{i\}) = 0$ if and only if $(I \setminus \{i\}) \cap I_e = \emptyset$. This only occurs if $I_e = \{i\}$, that is for the cluster associated with edge e_i . We can therefore say, by Eq. (32), that:

$$\sum_{e \in E} \alpha_e b_e^r(I \setminus \{i\}) = \sum_{e \in E \setminus \{e_i\}} \alpha_e = 0 \quad (33)$$

With a similar argument as in the unrooted case [see Eq. (19)], we can say that $\sum_{e \in E} \alpha_e = 0$ and we immediately conclude that $\alpha_{e_i} = 0$ for all $b_{e_i}^r(\cdot)$ associated with an external edge, e_i .

Now let us consider an internal edge e whose cluster is I_e , let us denote by E_s the subset of edges contained in the subtree which follows edge e , and let us assume that for all $e' \in E_s$, $\alpha_{e'} = 0$. To extend our result by induction, we now need to compute $b_{e'}^r(\bar{I}_e)$ for all $e' \in E$. From Eq. (31), $b_{e'}^r(\bar{I}_e) = 0$ if and only if $(\bar{I}_e) \cap I_{e'} = \emptyset$. This situation only occurs if $I_{e'} \subseteq I_e$, i.e. for clusters associated with edges $e' \in E_s$. From Eq. (32) estimated at $C = \bar{I}_e$, we can therefore say that:

$$\sum_{e' \in E} \alpha_{e'} b_{e'}^r(\bar{I}_e) = \sum_{e' \in E \setminus \{E_s \cup \{e\}\}} \alpha_{e'} = 0 \quad (34)$$

By using again the fact that $\sum_{e \in E} \alpha_e = 0$ and since we have assumed that for all $e_s \in E_s$, $\alpha_{e_s} = 0$, we get $\sum_{e' \in E \setminus E_s} \alpha_{e'} = 0$. Comparing with Eq. (34), we conclude that $\alpha_e = 0$.

Using this argument by induction, we can say that $(\alpha_e)_{e \in E} = 0$. \square

The next step consists in computing of the Shapley values for each characteristic function $b_e^r(\cdot)$ as in Eq. (20) where $b_e^r(\cdot)$ replaces $b_e^u(\cdot)$. It is immediate that the marginal contribution $(b_e^r(C) - b_e^r(C \setminus \{i\}))$ of taxon i to coalition C is again either 0 or 1, since $b_e^r(C) = 0$ always implies that $b_e^r(C \setminus \{i\}) = 0$. Moreover, if $i \notin I_e$, the path $\{\rho \rightarrow i\}$ never meets edge e , meaning that i 's marginal contribution to any coalition is always 0. In the opposite case, i.e. $i \in I_e$, this taxon has a unit marginal contribution if and only if the rest of coalition C is included in \bar{I}_e (the subset $C \setminus \{i\}$ could be empty). It remains, by a combinatorial argument, to identify the number of cases in which this last situation occurs and to reach the conclusion that:

Lemma 4 *If I_e denotes the cluster associated with edge e , the Shapley value of taxon $i \in I$ for the characteristic function $b_e^r(\cdot)$ is:*

$$Sh_{b_e^r}(i) = \mathbf{I}_{i \in I_e} \frac{1}{|I_e|} \text{ with } \mathbf{I}_{i \in I_e} = \begin{cases} 1 & \text{if } i \in I_e \\ 0 & \text{otherwise} \end{cases} \quad (35)$$

Proof As in the proof of Lemma 2, we now need to identify the cases in which Δ_e of Eq. (22) is equal to one (when, of course, $b_e^r(\cdot)$ is replaced by $b_e^r(\cdot)$).

Let us first assume that $i \in I_e$. From Eq. (31), (i) $b_e^r(C \cup \{i\}) = 1$ for any $C \subseteq I \setminus \{i\}$ and (ii) $b_e^r(C) = 0$ if and only if $C \subseteq \bar{I}_e$ (C being possibly empty), meaning that

$\Delta_e = 1$ if and only if $C \subseteq \bar{I}_e$. Moreover, there again exist $\binom{|\bar{I}_e|}{|C|}$ potential choices of a subset of $c = |C|$ elements in \bar{I}_e but now $c = 0, \dots, |\bar{I}_e|$ since C can be empty. The Shapley value is therefore:

$$\begin{aligned} Sh_{b_e^r}(i) &= \sum_{c=0}^{|\bar{I}_e|} \frac{c!(n-c-1)!}{n!} \binom{|\bar{I}_e|}{c} = \frac{|\bar{I}_e|!}{n!} \sum_{c=0}^{|\bar{I}_e|} \frac{(n-c-1)!}{(|\bar{I}_e|-c)!} \\ &= \frac{|\bar{I}_e|!(n-|\bar{I}_e|-1)!}{n!} \sum_{c=0}^{|\bar{I}_e|} \binom{n-c-1}{n-|\bar{I}_e|-1} \end{aligned} \quad (36)$$

By adapting the previous iterative Pascal formula [see Eq. (24)], we obtain:

$$\begin{aligned} Sh_{b_e^r}(i) &= \frac{|\bar{I}_e|!(n-|\bar{I}_e|-1)!}{n!} \left(\binom{n-1}{n-|\bar{I}_e|} + \binom{n-1}{n-|\bar{I}_e|-1} \right) \\ &= \frac{1}{n} \left(\frac{|\bar{I}_e|}{(n-|\bar{I}_e|)} + 1 \right) = \frac{1}{|I_e|} \text{ (since } |I_e| + |\bar{I}_e| = n) \end{aligned} \quad (37)$$

Now assume that $i \notin I_e$. In this case, either $b_e^r(C) = 1$ or 0 depending on whether $C \cap I_e \neq \emptyset$ or not; but adding i to C makes no difference anyway, since $i \notin I_e$. It

follows from the null player axiom (see Axiom 3) that $Sh_{b_e^r}(i) = 0$. By combining the two results, we get Eq. (35). \square

We observe, with the same argument as in the preceding section, that the Shapley value is still a linear operator on \mathcal{P}^r and we can conclude that:

Proposition 2 *Let T^r be a rooted phylogenetic tree with n leaves whose cluster structure is given by $\mathcal{C}_{T^r} = \{I_e\}_{e \in E}$. Under the standard Shapley approach based on Axioms 1, 2, 3 and 4, the Shapley value of species i is given by:*

$$Sh_{T^r}^r(i) = \sum_{e \in E} \frac{\mathbf{I}_{i \in I_e}}{|I_e|} w(e) \text{ with } \mathbf{I}_{i \in I_e} = \begin{cases} 1 & \text{if } i \in I_e \\ 0 & \text{otherwise} \end{cases} \quad (38)$$

This expression can even be simplified. Due to the presence of the indicator function $\mathbf{I}_{i \in I_e}$, the sum only runs over the set of edges, e , whose cluster, I_e , contains taxon i . Moreover, it is straightforward to verify that this set can be identified with the set of all edges contained in the path $\{\rho \rightarrow i\}$. This means that the Shapley value is the sum along $\{\rho \rightarrow i\}$ of the weights associated with these edges divided by $D_e = |I_e|$, the number of taxa descending from each of these edges. But this is the definition of the Fair Proportion index introduced by Redding and Mooers (2006).

Corollary 2 (Fuchs and Jin 2015, Th. 1) *The Shapley value of an individual in a rooted binary tree T^r is equal to the Fair Proportion index, i.e.*

$$Sh_{T^r}^r(i) = \sum_{e \in \{\rho \rightarrow i\}} \frac{w(e)}{D_e} = FP_{T^r}^r(i) \quad (39)$$

To conclude this section on the Shapley values of rooted trees, we will briefly look at the notion of transformation matrices again.

Remark 3 For rooted trees, the transformation matrix $M_{T^r}^r$ is given by the $(n, 2n - 2)$ matrix $M_{T^r}^r = \left[\frac{\mathbf{I}_{i \in I_e}}{|I_e|} \right]_{i \in I, e \in E}$ [see Eq. (38)], column e being the Shapley imputation induced by the PD^r game defined by the basis vector $(b_e^r(C))_{C \in 2^I}$ [see Eq. (35)]. The Shapley imputation is given by $(Sh_{T^r}^r(i))_{i \in I} = M_{T^r}^r \cdot (w(e))_{e \in E}$. Finally, this matrix $M_{T^r}^r$ is deduced from the topological structure of the tree, precisely, from the cluster structure contrary to the unrooted case where the split structure matters.

6 Comparing the two Shapley values

This comparison calls for identical trees, however the two Shapley values are typically defined on different trees (i.e. unrooted vs rooted). To overcome this problem, we introduce, in line with Wicke and Fischer (2017), the notion of unrooted rooted trees. The idea is the following. We start with a rooted phylogenetic tree $T^r = (V, E, w(\cdot))$ and transform the two edges incident to the root, say e_1 and e_2 , into a single edge, say

e'_1 whose weight is the sum of the weights of the initial incident edges, i.e. $w(e'_1) = w(e_1) + w(e_2)$. The new tree

$$T^{ur} = (V \setminus \{\rho\}, (E \setminus \{e_1, e_2\}) \cup \{e'_1\}, ((w(e))_{e \in E \setminus \{e_1, e_2\}}, w(e'_1))) \quad (40)$$

is called the unrooted rooted tree induced by the rooted tree T^r . The construction of the rooted Shapley values on T^r directly follows from Eq. (38), while the unrooted Shapley value on T^{ur} can be written as:

Proposition 3 *The unrooted Shapley value, $Sh_{T^{ur}}^u(i)$, of an unrooted rooted tree, T^{ur} , is*

$$Sh_{T^{ur}}^u(i) = \sum_{e \in E} w(e) \left(\mathbf{I}_{i \in I_e} \frac{|\bar{I}_e|}{n |I_e|} + (1 - \mathbf{I}_{i \in I_e}) \frac{|I_e|}{n |\bar{I}_e|} \right) \quad (41)$$

where E , $w(\cdot)$ and $s = \{\{I_e \mid \bar{I}_e\}\}_{e \in E}$ are the edges, weights and split structure of the rooted tree, T^r

Proof From Eq. (27), the unrooted Shapley value on T^{ur} is :

$$\begin{aligned} Sh_{T^{ur}}^u(i) &= w(e'_1) \underbrace{\left(\mathbf{I}_{i \in I_{e'_1}} \frac{|\bar{I}_{e'_1}|}{n |I_{e'_1}|} + (1 - \mathbf{I}_{i \in I_{e'_1}}) \frac{|I_{e'_1}|}{n |\bar{I}_{e'_1}|} \right)}_m \\ &\quad + \sum_{e \in E \setminus \{e_1, e_2\}} w(e) \left(\mathbf{I}_{i \in I_e} \frac{|\bar{I}_e|}{n |I_e|} + (1 - \mathbf{I}_{i \in I_e}) \frac{|I_e|}{n |\bar{I}_e|} \right) \end{aligned} \quad (42)$$

Now let us denote by $\{I_{e_1} \mid \bar{I}_{e_1}\}$ and $\{I_{e_2} \mid \bar{I}_{e_2}\}$ the split structures associated with edges e_1 and e_2 of the rooted tree T^r , with the first component corresponding to the cluster of each edge. Since these edges are incident to the root, $I_{e_1} = \bar{I}_{e_2}$. Moreover, if $\{I_{e'_1} \mid \bar{I}_{e'_1}\}$ denotes the split associated with e'_1 in T^{ur} we can say, w.l.o.g., that $I_{e_1} = I_{e'_1}$. It follows that:

$$m = \mathbf{I}_{i \in I_{e_1}} \frac{|\bar{I}_{e_1}|}{n |I_{e_1}|} + (1 - \mathbf{I}_{i \in I_{e_1}}) \frac{|I_{e_1}|}{n |I_{e_1}|} = (1 - \mathbf{I}_{i \in I_{e_2}}) \frac{|I_{e_2}|}{n |I_{e_2}|} + \mathbf{I}_{i \in I_{e_2}} \frac{|\bar{I}_{e_2}|}{n |I_{e_2}|}$$

Since $w(e'_1) = w(e_1) + w(e_2)$ and, T^r , T^{ur} have identical weights and splits except for edges e_1 and e_2 , Eq. (41) is satisfied. \square

To compare the two values, let us now compute their difference, given by:

$$\begin{aligned} \forall i \in I, \Delta_{sh}(i) &= Sh_{T^{ur}}^u(i) - Sh_{T^r}^r(i) \\ &= \frac{1}{n} \sum_{e \in E} w(e) \left(-\mathbf{I}_{i \in I_e} + (1 - \mathbf{I}_{i \in I_e}) \frac{|I_e|}{|\bar{I}_e|} \right) \end{aligned} \quad (43)$$

Since both Shapley operators are linear with respect to the weight vector, $w = (w(e))_{e \in E}$, the difference, $\Delta_{sh} = (\Delta_{sh}(i))_{i \in I}$, shares the same property and can be written as:

$$\Delta_{sh} = \frac{1}{n} A \cdot w, \text{ the generic term of } A \text{ being } a_{i,e} = \begin{cases} -1 & \text{if } i \in I_e \\ \frac{|I_e|}{|I|} & \text{otherwise} \end{cases} \quad (44)$$

where A is a $(n, 2n - 2)$ matrix.

The properties of this matrix are crucial to understanding the difference between these two Shapley values. First, note that both values satisfy efficiency (see Axiom 1). This implies that:

Proposition 4 *The image, $\langle A \rangle$, of the linear operator A is orthogonal to $\varepsilon_n \in \mathbb{R}^n$, the vector in which all entries are equal to 1. It follows that the rank of matrix A is at most $(n - 1)$.*

Proof From Axiom 1, we know that the sum over the taxa of the Shapley values is always equal to the phylogenetic diversity of the grand coalition I , i.e. $\sum_{i \in I} Sh_{Tr}^r(i) = v_{Tr}(I)$, and $\sum_{i \in I} Sh_{Tur}^u(i) = v_{Tur}(I)$. Moreover, a rooted tree and its associated unrooted rooted tree share the same phylogenetic diversity, i.e. $v_{Tr}(I) = v_{Tur}(I)$. It follows that $\sum_{i \in I} \Delta_{sh}(i) = 0$ regardless of the weight vector. From Eq. (44), we conclude that $\varepsilon_n' \cdot A = 0$ or that $\langle A \rangle$ is orthogonal to $\varepsilon_n \in \mathbb{R}^n$ which means that $rank(A) \leq n - 1$. \square

This orthogonality property has an additional consequence. We can claim that it is impossible to find a system of weights, $w = (w(e))_{e \in E}$, which satisfies either $Aw < 0$ or $Aw > 0$, without contradicting $\varepsilon_n' \cdot A = 0$. In other words, it is impossible to have for all i , $Sh_{Tur}^u(i) \geq Sh_{Tr}^r(i)$ or $Sh_{Tur}^u(i) \leq Sh_{Tr}^r(i)$ with at least one strict inequality, meaning that neither the unrooted nor the rooted Shapley value dominates the other.

Since $\sum_{i \in I} \Delta_{sh}(i) = 0$, the previous Lemma also suggests that one component of Δ_{sh} is redundant. We dismiss the last component of this vector and call $\tilde{\Delta}_{sh}$ the new difference vector. Let us now denote by E_{ext} and E_{int} the sets of relevant external and internal edges and by e_n the edge incident to taxon n . After a suitable reordering of the edges (i.e. the columns of A and the weight space) $\tilde{\Delta}_{sh}$ can be written as:

$$\left(\tilde{\Delta}_{sh}(i) \right)_{i=I \setminus \{n\}} = \frac{1}{n} \left[\underbrace{[a_{i,e}]_{i=I \setminus \{n\}, e \in E_{ext} \setminus \{e_n\}}]_{=B_{ext}} \mid \underbrace{[a_{i,e_n}]_{i=I \setminus \{n\}}]_{=B_{e_n}} \mid \underbrace{[a_{i,e}]_{i=I \setminus \{n\}, e \in E_{int}}]_{=B_{int}} \right] \cdot w \quad (45)$$

It remains to prove that the $(n - 1, n - 1)$ matrix B_{ext} is invertible in order to claim that:

Proposition 5 A is a linear mapping of rank $(n - 1)$. Its kernel, also of dimension $(n - 1)$, is given by the weights, $w \in \mathbb{R}^{2n-2}$, which satisfy:

$$\begin{aligned} & (w(e))_{e \in E_{ext} \setminus \{e_n\}} \\ &= \frac{n-1}{n} (I_{n-1} + \varepsilon_{n-1} \cdot (\varepsilon_{n-1})') \cdot [B_{e_n} \mid B_{int}] \cdot \begin{pmatrix} w(e_n) \\ (w(e))_{e \in E_{int}} \end{pmatrix} \end{aligned} \quad (46)$$

with I_{n-1} the identity matrix of \mathbb{R}^{n-1} and ε_{n-1} the vector of \mathbb{R}^{n-1} in which all entries are equal to 1.

Proof The kernel, $\ker(A)$, of A satisfies $Aw = 0$. But we know that, for the vector $\varepsilon_n \in \mathbb{R}^n$, $(\varepsilon_n)' \cdot A = 0$, meaning that the last equation in $Aw = 0$ is redundant. It follows from Eq. (45) that:

$$w \in \ker(A) \Leftrightarrow [B_{ext} \mid B_{e_n} \mid B_{int}] \cdot \begin{pmatrix} (w(e))_{e \in E_{ext} \setminus \{e_n\}} \\ w(e_n) \\ (w(e))_{e \in E_{int}} \end{pmatrix} = 0 \quad (47)$$

Let us now concentrate on the $(n - 1, n - 1)$ matrix B_{ext} . To construct this matrix, we need to return to the early definition of the generic term a_{ie} of the matrix A [see Eq. (44)] and observe, since we are looking at external edges, that $I_e = \{i\}$. It follows that:

$$B_{ext} = \begin{bmatrix} -1 & \frac{1}{n-1} & \cdots & \frac{1}{n-1} \\ \frac{1}{n-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{1}{n-1} \\ \frac{1}{n-1} & \cdots & \frac{1}{n-1} & -1 \end{bmatrix} = -\frac{n}{n-1} (I_{n-1} - \frac{1}{n} \varepsilon_{n-1} \cdot (\varepsilon_{n-1})') \quad (48)$$

with I_{n-1} the identity matrix of \mathbb{R}^{n-1} and ε_{n-1} the unit vector of \mathbb{R}^{n-1} . Moreover, since $(\varepsilon_{n-1})' \cdot \varepsilon_{n-1} = n - 1$, a simple computation shows that:

$$(I_{n-1} - \frac{1}{n} \varepsilon_{n-1} \cdot (\varepsilon_{n-1})') \cdot (I_{n-1} + \varepsilon_{n-1} \cdot (\varepsilon_{n-1})') = I_{n-1} \quad (49)$$

It follows that $(B_{ext})^{-1} = -\frac{n-1}{n} (I_{n-1} + \varepsilon_{n-1} \cdot (\varepsilon_{n-1})')$. Eq. (46) of Proposition 5 directly follows from the block decomposition provided by Eq. (47). This also shows that $\dim(\ker(A)) = n - 1$. \square

The previous Lemma shows that the two Shapley values are the same on a subset of weights of dimension $(n - 1)$ as long as there are no restrictions on the set of weights. Usually, however, these weights are assumed to be non-negative. So if $\ker(A) \cap \mathbb{R}_+^{2n-2} = \{0\}$, the previous result is of no real interest. The next lemma shows that this is not the case. We even go a step further by proving that:

Proposition 6 The kernel of A intersects the interior of the positive orthant, i.e. $\ker(A) \cap \mathbb{R}_{++}^{2n-2} \neq \emptyset$. The relative interior of $\ker(A) \cap \mathbb{R}_{++}^{2n-2}$ is therefore also of dimension $n - 1$.

Proof This result follows from Farkas' Lemma, in particular from Stiemke's alternative, which states that either (i) there exists $w \in \mathbb{R}^{2n-2}$ satisfying $Aw = 0$ and $w \gg 0$, or (ii) there exists $x \in \mathbb{R}^n$ satisfying $x'A > 0$. We show that (ii) is impossible. We begin with the matrix A given by Eq. (44) and reorganize the columns such that the n first columns correspond to external edges. The matrix becomes $[A_{ext} \mid A_{int}]$. Since $x'A > 0$ assume first that $x'A_{ext} > 0$. From Eq. (44), this amounts to saying that for all $i = 1, \dots, n$

$$-x_i + \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^n x_j \geq 0 \text{ with at least one strict inequality} \quad (50)$$

or, equivalently, that for all $i = 1, \dots, n$,

$$\frac{1}{n} \sum_{j=1}^n x_j \geq x_i \text{ with at least one strict inequality} \quad (51)$$

Summing over i gives $\sum_{j=1}^n x_j > \sum_{j=1}^n x_j$, an impossibility.

Since $x'A > 0$, consider now the complementary case given by $x'A_{ext} = 0$ and $x'A_{int} > 0$. The first condition and Eq. (51) say that for all $i = 1, \dots, n$, $\frac{1}{n} \sum_{j=1}^n x_j = x_i$ so that each x_i is equal to the same value $\lambda = \frac{1}{n} \sum_{j=1}^n x_j$. In other words, we have $x = \lambda \varepsilon_n$ with ε_n the vector of \mathbb{R}^n in which all entries are equal to 1. Moreover, since $x'A > 0$, there exists an internal edge e and a column, a_{int}^e , in A_{int} which satisfies $\lambda \varepsilon_n' \cdot a_{int}^e > 0$. But by Proposition 4, $\varepsilon_n' \cdot a_{int}^e = 0$, which is again an impossibility. \square

The next remark summarizes this first discussion so far.

Remark 4 For a given tree structure, we can say that:

- (i) neither the unrooted nor the rooted Shapley value induces a higher evaluation of the contribution to biodiversity for all taxa,
- (ii) the two Shapley values are identical on a linear subset of dimension $(n - 1)$ of weights satisfying Eq. (46),
- (iii) the previous result extends to non-negative weights but the subset on which these values are the same becomes a polytope of dimension $(n - 1)$,
- (iv) for the other weights, at least two taxa with different Shapley values exist.

In line with Hartmann (2013), let us now compare the contribution of the weight of an edge to each Shapley value. This contribution is straightforward for the rooted Shapley value. Since the latter is equal to the Fair Proportion index, each edge only contributes to its descendants, I_e , in a proportion, $\frac{1}{|I_e|}$, inverse to their number. For an unrooted Shapley value [see Eq. (27)], the sharing rule for the weight associated with an edge is totally different. Each edge contributes to the Shapley value of each taxon under a two-step sharing rule. First, the weight is shared out between the descendants and the non-descendants: the descendants as a whole group receive a quantity proportional to the number of non-descendants, $\frac{n-|I_e|}{n}$, and vice-versa. These overall amounts are

then uniformly shared out among the members of each group. To summarize, we can say:

Proposition 7 *The contribution of the weight, $w(e)$, of an edge, $e \in E$, to the two Shapley values of taxon $i \in I$ is summarized in the following:*

	$i \in I_e$	$i \notin I_e$	
$Sh_{Tur}^u(i)$	$\frac{1}{ I_e } \binom{n- I_e }{n}$	$\frac{1}{n- I_e } \binom{ I_e }{n}$	(52)
$Sh_{Tr}^r(i)$	$\frac{1}{ I_e }$	0	

As a consequence, we immediately observe that:

Corollary 3 (Hartmann 2013, Th. 1) *As the number of taxa $n \rightarrow \infty$, the contribution of an edge e to the rooted and unrooted Shapley values becomes the same.*

7 Comparing the ordering induced by the two Shapley values

Up to now, Proposition 4 tells us that the two Shapley values are equal, and in fact equal to the Fair Proportion index, on a linear subset of weights of dimension $n - 1$, this subset being largely related to the split structure of the tree under consideration. But this makes us wonder what happens when the weight structures do not belong to this set. In particular, if we consider prioritization programs, do these two Shapley values induce drastically different species rankings or are they fairly similar? To answer this question, we proceed, like Wicke and Fischer (2017), by simulations⁵ and reach similar conclusions: both rankings are strongly correlated. Our method is nevertheless different.

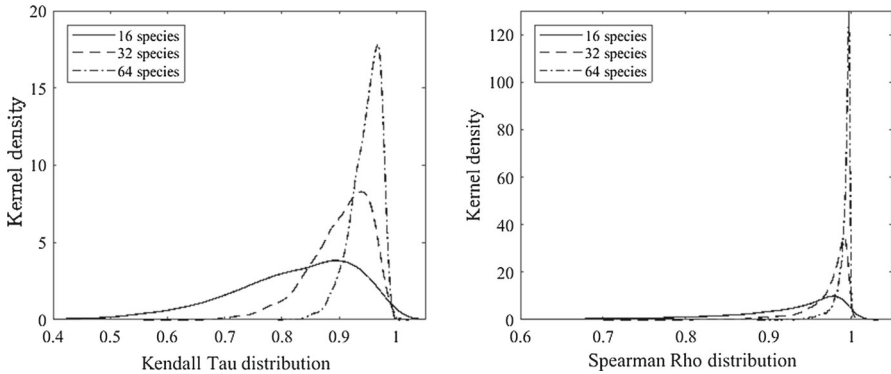
Contrary to Wicke and Fischer (2017), we do not generate a set of trees using an underlying random tree model, e.g. a pure birth process like the Yule model or a birth-death process. Our simulation method essentially builds on our previous results. Our preliminary example makes it clear that the computation of both Shapley values only requires three pieces of information: the number of taxa, the split structure and the weights. We fix the number of taxa respectively at 16, 32 and 64. We then construct a split structure by randomly splitting the set of species into two subsets, and recursively split each of these subsets into two new subsets until we reach singletons. Finally, we assign a random weight to each element of the split structure. The weight choice however satisfies two requirements. We ensure, using Proposition 5, that the two Shapley values induce different imputation vectors and we normalize the sum of the weights to 1 to eliminate redundant rankings.⁶ We then compute for each randomly-selected split structure both transformation matrices, using Eq. (52) of Proposition 7, and use the randomly-chosen weights to obtain the unrooted rooted and rooted Shapley imputation vectors.

⁵ The Matlab R2018a codes are available in an electronic supplement material associated with this manuscript (see Supplementary Material).

⁶ More precisely, this second restriction follows from the fact that the two Shapley values are linear in weights, meaning that if we multiply the weight structure by any $\lambda > 0$ we obtain the same ranking.

Table 1 Main characteristics of the Kendall and Spearman coefficients

	Kendall 16 species	Spearman 16 species	Kendall 32 species	Spearman 32 species	Kendall 64 species	Spearman 64 species
Mean	.8182	.9121	.9013	.9720	.9464	.9913
Median	.8333	.9441	.9113	.9824	.9514	.9948
SD	.1125	.0919	.0539	.0303	.0268	.0099

**Fig. 2** Kernel Density estimates: Kendal and Spearman rank correlation

The rankings induced by these two imputations are compared with some classical statistical tools, the Kendall (1938) and the Spearman (1904) rank correlation coefficients. Our first statistic essentially captures the inversions of elements in the two rankings, while the second highlights differences in the position of an element in each ranking. But in any case, if both statistics are close to 1, the two rankings are considered as strongly (positively) correlated.

Looking at trees with 16, 32 and 64 leaves, and randomly selecting 5000 trees in each case, we obtain 2×3 sets of 5000, respectively, Kendall and Spearman rank correlation coefficients associated with trees with respectively, 16, 32 and 64 species. The main characteristics of these distributions are summarized in Table 1.

These results clearly suggest that the two rankings induced by the two Shapley values are largely correlated and, as expected (see Hartmann 2013), the degree of correlation increases with the number of species. This intuition is confirmed by a kernel estimation of the density of these different distributions (see Fig. 2).

8 Concluding remarks

The purpose of this paper was to explore the main differences between the Shapley values for a taxon introduced by Haake et al. (2007) and Fuchs and Jin (2015). Although these two metrics are based on the same Shapley axiomatic, they do not use the same definition of phylogenetic diversity: the former does not include the root in the subtree while the latter does. This induces two different subsets of potential

characteristic functions and different Shapley values. To illustrate this, we explicitly compute the Shapley values in both cases by identifying, in line with Shapley, the bases of the two different sets of characteristic functions. This gives us the opportunity to formally compare the two values and to identify for each phylogenetic tree a set of weights for which these quantities are identical. This clearly raises a second question where prioritization problems are concerned. Although the two Shapley values are different, do they induce a similar ranking for the different species? To answer this question, we simulate alternative situations and show that the Kendall and Spearman rank correlation coefficients are both close to 1. Bearing in mind that the Shapley value introduced by Fuchs and Jin (2015) is equal to the Fair Proportion index constructed by Redding and Mooers (2006), this suggests that the latter, simpler index can be used in prioritization problems.

References

- Bandelt H-J, Dress AWM (1992) Split decomposition theory for metrics on a finite set. *Adv Math* 92:47–105. [https://doi.org/10.1016/0001-8708\(92\)90061-O](https://doi.org/10.1016/0001-8708(92)90061-O)
- Buneman P (1971) The recovery of trees from measures of dissimilarity. In: Hobson FR, Kendal DG, Tautu P (eds) *Mathematics in the archeological and historical sciences*. Edinburgh University Press, Edinburgh, pp 387–395
- Cadotte MW, Davies TJ, Regetz J, Kembel SW, Cleland E, Oakley TH (2010) Phylogenetic diversity metrics for ecological communities: integrating species richness, abundance and evolutionary history. *Ecol Lett* 13:96–105. <https://doi.org/10.1111/j.1461-0248.2009.01405.x>
- Crozier RH, Dunnett LJ, Agapow P (2005) Phylogenetic biodiversity assessment based on systematic nomenclature. *Evol Bioinform Online* 1:11–36
- Desper R, Gascuel O (2005) The minimum evolution distance-based approach to phylogenetic inference. In: Gascuel O (ed) *Mathematics of evolution and phylogeny*. Oxford University Press, Oxford, pp 1–32
- Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Biol Conserv* 61:1–10. [https://doi.org/10.1016/0006-3207\(92\)91201-3](https://doi.org/10.1016/0006-3207(92)91201-3)
- Faith DP (2006) The role of the phylogenetic diversity measure, PD, in bio-informatics: getting the definition right. *Evol Bioinform Online* 2:277–283
- Fuchs M, Jin EY (2015) Equality of Shapley value and fair proportion index in phylogenetic trees. *J Math Biol* 71:1133–1147. <https://doi.org/10.1007/s00285-014-0853-0>
- Fuchs M, Paningbatan AR (2019) Correlation between Shapley values of rooted phylogenetic trees under the beta-splitting model. *J Math Biol*. <https://doi.org/10.1007/s00285-019-01435-3>
- Haake CJ, Kashiwada A, Su FE (2007) The Shapley value of phylogenetic trees. *J Math Biol* 56:479–497. <https://doi.org/10.1007/s00285-007-0126-2>
- Hartmann K (2013) The equivalence of two phylogenetic biodiversity measures: the Shapley value and fair proportion index. *J Math Biol* 67:1163–1170. <https://doi.org/10.1007/s00285-012-0585-y>
- Jensen EL, Mooers AØ, Cacccone A, Russello MA (2016) I-HEDGE: determining the optimum complementary sets of taxa for conservation using evolutionary isolation. *PeerJ* 4:e2350. <https://doi.org/10.7717/peerj.2350>
- Kendall M (1938) A new measure of rank correlation. *Biometrika* 30:81–89. <https://doi.org/10.2307/2332226>
- Kleinberg NL, Weiss JH (1985) A new formula for the Shapley value. *Econ Lett* 18:311–315. [https://doi.org/10.1016/0165-1765\(85\)90249-6](https://doi.org/10.1016/0165-1765(85)90249-6)
- Martyn I, Kuhn TS, Mooers AØ, Moulton V, Spillner A (2012) Computing evolutionary distinctiveness indices in large scale analysis. *Algorithm Mol Biol* 7:6. <https://doi.org/10.1186/1748-7188-7-6>
- Minh BQ, Klaere S, von Haeseler A (2009) Taxon selection under split diversity. *Syst Biol* 57:586–594. <https://doi.org/10.1093/sysbio/syp058>
- Redding DW, Mooers AØ (2006) Incorporating evolutionary measures into conservation prioritization. *Conserv Biol* 20:1670–1678. <https://doi.org/10.1111/j.1523-1739.2006.00555.x>

- Redding DW, Mazel F, Mooers AØ (2014) Measuring evolutionary isolation for conservation. *PLoS ONE* 9(12):e113490. <https://doi.org/10.1371/journal.pone.0113490>
- Rothblum UG (1988) Combinatorial representations of the Shapley value based on average relative payoffs. In: Roth AE (ed) *The Shapley value: essays in honor of Lloyd S. Shapley*. Cambridge University Press, Cambridge, pp 121–126
- Semple C, Steel M (2003) *Phylogenetics*. Oxford University Press, Oxford
- Shapley LS (1953) A value for n-person games. In: Kuhn HW, Tucker AW (eds) *Contributions to the theory of games, volume II, annals of mathematics studies 28*. Princeton University Press, Princeton, pp 307–17
- Spearman C (1904) The proof and measurement of association between two things. *Am J Psychol* 15:72–101. <https://doi.org/10.2307/1412159>
- Volkman L, Martyn I, Moulton V, Spillner A, Mooers AØ (2014) Prioritizing populations for conservation using phylogenetic networks. *PLoS ONE* 9(2):e88945. <https://doi.org/10.1371/journal.pone.0088945>
- Weitzman ML (1998) The Noah's Ark problem. *Econometrica* 66:1279–1298. <https://doi.org/10.2307/2999617>
- Wicke K, Fischer M (2017) Comparing the rankings obtained from two biodiversity indices: the Fair Proportion Index and the Shapley Value. *J Theor Biol* 430:207–214. <https://doi.org/10.1016/j.jtbi.2017.07.010>
- Wicke K, Fischer M (2019) On the Shapley value of unrooted phylogenetic trees. *Bull Math Biol* 81:618–638. <https://doi.org/10.1007/s11538-018-0392-8>