*Syst. Biol.* 53(2):327–332, 2004

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150490423809

# Distances that Perfectly Mislead

DANIEL H. HUSON[1] AND MIKE STEEL[2]

[1]*Center for Bioinformatics, Tübingen University, Tübingen, Germany; E-mail: huson@informatik.uni-tuebingen.de*
[2]*Biomathematics Research Centre, University of Canterbury, Private Bag 4800, Christchurch, New Zealand; E-mail: m.steel@math.canterbury.ac.nz*

*Abstract*.—Given a collection of discrete characters (e.g., aligned DNA sites, gene adjacencies), a common measure of distance between taxa is the proportion of characters for which taxa have different character states. Tree reconstruction based on these (uncorrected) distances can be statistically inconsistent and can lead to trees different from those obtained using character-based methods such as maximum likelihood or maximum parsimony. However, in these cases the distance data often reveal their unreliability by some deviation from additivity, as indicated by conflicting support for more than one tree. We describe two results that show how uncorrected (and miscorrected) distance data can be simultaneously perfectly additive and misleading. First, multistate character data can be perfectly compatible and define one tree, and yet the uncorrected distances derived from these characters are perfectly treelike (and obey a molecular clock), only for a completely different tree. Second, under a Markov model of character evolution a similar phenomenon can occur; not only is there statistical inconsistency using uncorrected distances, but there is no evidence of this inconsistency because the distances look perfectly treelike (this does not occur in the classic two-parameter Felsenstein zone). We characterize precisely when uncorrected distances are additive on the true (and on a false) tree for four taxa. We also extend this result to a more general setting that applies to distances corrected according to an incorrect model. [Additive metric; distance-based phylogeny reconstruction; inconsistency.]

When distances between species have been estimated from character data and then either corrected or left uncorrected, they often fail to fit exactly on any tree. It is easy to measure how treelike (additive) any collection of distances is by, for example, looking at their splits graph (Bandelt and Dress, 1992; Huson, 1998) or other quartet-based approaches such as delta-plots (Holland et al., 2002). When the distances are perfectly additive (i.e., there is some tree $\mathcal{T}$ with positive edge weights so that each given distance between two tip species exactly matches the path distance in the tree), then the splits graph will display $\mathcal{T}$ exactly. Often however the splits graph will be a nontree network, and in the case of four species we may always represent any set of distances exactly by path lengths in a rectangle with one short side (of length $S$, equal to zero precisely when the distances fit a tree) and one long side ($L \geq S$) (Fig. 1).

The extent to which distances are treelike (as measured, for example, by $1 - [S/L]$) is often taken as an indication of their suitability and likely accuracy for tree reconstruction. Investigators are generally less happy inferring trees from distances when they display conflicting signals, as indicated, for example, by fat rectangles in the splits graph, than when the distances fit near perfectly on a tree.

Here, we provide two results that demonstrate that additivity is no guarantee of accuracy for uncorrected distances. Informally, these uncorrected distances simply count the proportion of characters for which the taxa have different states. More formally, given a sequence $\mathcal{C} = (f_1, \ldots, f_k)$ of characters on a set $X$ of species, let

$$d_{\mathcal{C}}(i, j) = \frac{|[s : f_s(i) \neq f_s(j)]|}{k},$$

the (normalized) Hamming distance, or sequence dissimilarity between species $i$ and $j$.

The distance measure $d_{\mathcal{C}}$ has some appealing properties. For example, if the characters evolve under any of the usual stochastic models (e.g., allowing a general time reversible stationary process, with rates across sites) and a molecular clock applies, then $d_{\mathcal{C}}$ is statistically consistent for tree topology reconstruction. In one recent simulation study, Holland et al. (2003) found the use of $d_{\mathcal{C}}$ more accurate for tree reconstruction in this setting than the use of corrected distances. One does not need to know the fine details of the model (e.g., for aligned DNA sequences, these would include parameters relating to rate variation across sites or nucleotide transition frequencies; for gene order, relevant parameters would relate to the relative rates of different rearrangement processes); one can simply apply methods such as neighbor joining directly to the uncorrected values $d_{\mathcal{C}}(i, j)$. The branch lengths in this case may not be consistently estimated (they will generally be underestimated, and consistent estimation requires knowing more of the fine details of the model).

We show here how the characters $\mathcal{C}$ can perfectly favor one tree (through the eyes of character-based methods such as compatibility, parsimony, and likelihood), yet the induced dissimilarities $d_{\mathcal{C}}$ can appear to perfectly fit a totally different tree. We also describe precisely when $d_{\mathcal{C}}$ is perfectly additive on an incorrect tree when the characters evolve under a certain Markov model. The proofs of our main results appear in the Appendix.

## PERFECT CHARACTERS WITH PERFECTLY MISLEADING DISTANCES

For any sequence $\mathcal{C}$ of perfectly compatible binary characters, the induced distance $d_{\mathcal{C}}$ is additive on the tree(s) that these characters support (see Semple and Steel, 2003, proposition 7.1.9). In the case of nonbinary characters, this statement no longer holds; the distance $d_{\mathcal{C}}$ derived from perfectly compatible characters may no

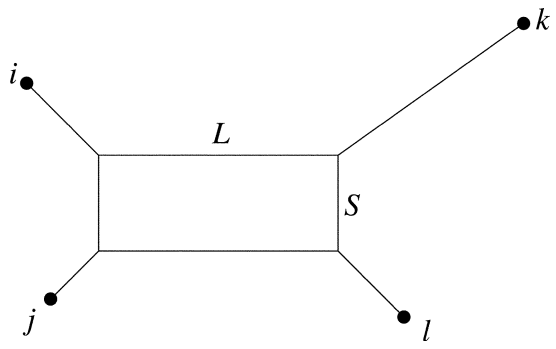


FIGURE 1. Any set of distances for four taxa can be represented by a splits graph of this form.

longer be additive on all or even any of the trees on which the characters in $\mathcal{C}$ are homoplasy free. It might be hoped that in these cases the induced distances would reveal their misleading nature by being not particularly additive (on any tree). However, the following theorem shows just how wrong this intuition can be. There exist data sets where the distances look perfectly additive (indeed clocklike) on one tree, yet they are homoplasy free only on a completely different tree. This condition holds for any number of taxa provided the there is a large enough number of character states. A binary phylogenetic tree is one that has no polytomies (i.e., it is fully resolved).

**Theorem 2.1**

Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be any two binary phylogenetic trees on the same set $X$ of species. Then there exists a sequence $\mathcal{C}$ of multistate characters for which

1. $\mathcal{C}$ is homoplasy free for $\mathcal{T}_1$ and for no other phylogenetic tree on $X$ (including $\mathcal{T}_2$).
2. The distances $d_{\mathcal{C}}$ derived from $\mathcal{C}$ are perfectly additive on $\mathcal{T}_2$ yet they are not additive on any other phylogenetic tree on $X$ (including $\mathcal{T}_1$). Furthermore, we may insist that the distances $d_{\mathcal{C}}$ fit a molecular clock (i.e., are ultrametric) on $\mathcal{T}_2$.

Figure 2 illustrates Theorem 2.1 in the simplest case where $|X| = 4$.

We may modify Theorem 2.1 to obtain a sequence $\mathcal{C}$ of multistate characters for which $\mathcal{T}_1$ has a strictly larger likelihood score under a symmetric Poisson model than does any other tree, yet $d_{\mathcal{C}}$ is additive only on $\mathcal{T}_2$. This follows by adding to $\mathcal{C}$ a sufficiently large number of constant sites and then invoking theorem 7 of Tuffley and Steel (1997). Our proof of Theorem 2.1 (Appendix) uses a construction for which the number of character states is 1 less than the number of taxa; however, we suspect this is merely an artifact of our proof, and a more intricate argument might be possible using fewer states.

## EVOLVING PERFECTLY MISLEADING UNCORRECTED DISTANCES

The data sets just described were deliberately chosen to illustrate an extreme phenomenon. Here, we investigate when precisely characters that evolve under a Markov process can also provide perfectly deceptive induced distances. We describe precise conditions for a four-taxon tree.

Suppose that $n$ sites or characters are generated independently on the tree $\mathcal{T}_1$ (Fig. 3a). Under the usual models of DNA evolution, uncorrected distances are generally not exactly additive; however, they can be additive both on the "true" tree that generated them and, in another slice of parameter space, on a false tree. In the later case, their appearance of being exactly additive may be deceptive, because it shows no evidence of causing a problem, nevertheless the wrong tree will be inferred from them.

Uncorrected distances will usually fail to be exactly treelike, but when they are treelike there are two possibilities:

1. The distances are exactly treelike on the true topology that generated the data.
2. The distances are exactly treelike on an incorrect topology.

Situation 1 occurs when there is a molecular clock (Steel and Penny, 2000: theorem 4); however, there is a strictly larger region of parameter space where this can occur.

Regarding situation 2, distance methods can be inconsistent (Felsenstein, 1978). However, situation 2 requires much more than mere inconsistency; not only must the distances favor a false tree over the true tree, but they must show no support for any tree other than the false tree. The subspace of parameter space where this occurs is the region of perfect inconsistency. This region does not intersect with the original two-parameter zone described by Felsenstein (1978); in that zone there is always

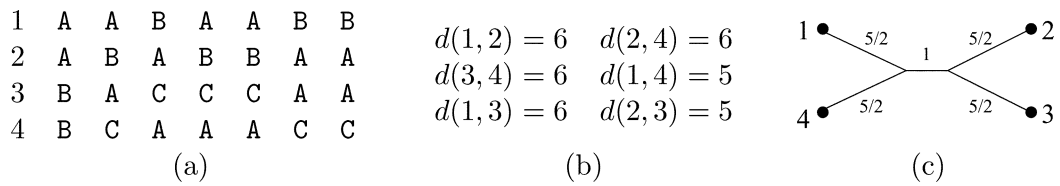


FIGURE 2. (a) Seven characters are homoplasy free on the tree 12|34. (b) Corresponding distances $d = 7d_{\mathcal{C}}$. (c) These distances are perfectly additive (and clocklike) on the tree 14|23.

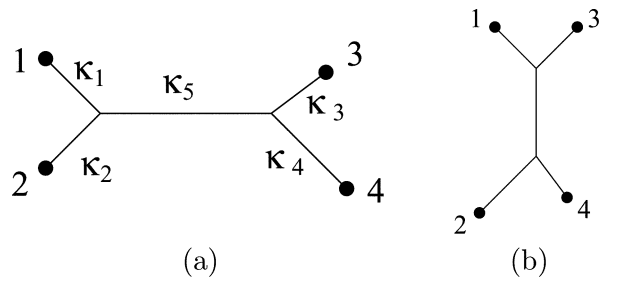


FIGURE 3. (a) Tree $\mathcal{T}_1$ used to generate sequence data. (b) An alternative tree $\mathcal{T}_2$ on the same set of four taxa.

conflicting support but more support for a false tree. Near the region of perfect inconsistency, the close fit of the uncorrected distances could mislead the unwary investigator into concluding that the reconstructed tree was accurate (or that the distances did not need correcting). We illustrate this region with an example and end by describing a more general result.

First, we will analyze exactly when situations 1 and 2 can arise when the underlying model is the equal input model on $r$ states. This is a time-reversible model whose transition rate matrix has as its $i$th row a common nondiagonal entry $\pi_i$ (which represents the expected frequency of state $i$ in the data). When $r = 4$, this is the Tajima–Nei (1984) model, which has a three-parameter rate matrix, allowing for an arbitrary base composition. As a special case, when all the base compositions are equal this model becomes the symmetric Poisson model (Jukes–Cantor [1969] when $r = 4$, and the Mk model of Lewis [2001] when $r = k \geq 4$). For the tree $\mathcal{T}_1$ displayed in Figure 3a the values $\kappa_i$ represent the expected number of substitutions on edge $i$.

Let us generate a sequence $\mathcal{C}$ of $k$ characters independently under this model. As $k$ becomes large, $d_{\mathcal{C}}(i, j)$ converges to its expected value, i.e., the probability that $i$ and $j$ are in different states under the model, which we denote $p(i, j)$. Let $p = [p(i, j)]$ be the matrix of values of $p$.

We now describe conditions on the $\kappa_i$ values for which the values of $p$ will be perfectly treelike in situations 1 (=treelike on $\mathcal{T}_1$) and 2 (=treelike on a different tree, say the tree $\mathcal{T}_2$ with topology 13|24, shown in Fig. 3b). To describe these results it is useful to let $p_\infty = 1 - \sum_{i=1}^{r} \pi_i^2$, and for $i = 1, \ldots, 5$, let $x_i = \exp(-\kappa_i / p_\infty)$. Note that $p_\infty$ is the limiting probability that two species are in different states as their evolutionary distance tends to infinity.

**Theorem 3.1**

Suppose that $p$ is the matrix of uncorrected distances generated by tree $\mathcal{T}_1$ with edge parameters $x_i = \exp(-\kappa_i / p_\infty)$ under the equal-input model.

1. The matrix $p$ is additive on $\mathcal{T}_1$ (and only on $\mathcal{T}_1$) if and only if both of the following two conditions apply:

$$x_1 = x_2 \text{ or } x_3 = x_4 \text{ (or both)}$$

and

$$x_5 < \frac{x_3 x_4 + x_1 x_2}{x_1 x_3 + x_2 x_4}.$$

2. The matrix $p$ is additive on $\mathcal{T}_2$ (and only on $\mathcal{T}_2$) if and only if the following two conditions apply:

$$(x_1 - x_2)(x_3 - x_4) > 0$$

and

$$x_5 = \frac{x_1 x_2 + x_3 x_4}{x_1 x_4 + x_2 x_3}.$$

We can express the conditions described in Theorem 3.1 purely in terms of the $p(i, j)$ values because the $x_k$ values are determined by the $p(i, j)$ values. For example,

$$x_1 = \sqrt{\frac{p(1,2)p(1,3)}{p(2,3)}},$$

and

$$x_5 = \sqrt{\frac{p(1,3)p(2,4)}{p(1,2)p(3,4)}}.$$

In part 1, the constraint on parameter space described is strictly weaker than the constraint imposed by a molecular clock. For example, suppose we set $x = (x_1, x_2, x_3, x_4, x_5) = (0.5, 0.5, 0.2, 0.1, 0.1)$. These values for $x$ satisfy the condition in part 1 of Theorem 3.1 and so are additive on $\mathcal{T}_1$. However, if we consider the values of the function $p$ on the pairs of elements chosen from $\{1, 3, 4\}$, we obtain three distinct values (i.e., $p(1, 3)$, $p(1, 4)$ and $p(3, 4)$ are all different, as can be verified by applying Equation 5 of the Appendix), and so $p$ is not an ultrametric. Consequently, because $p$ is a monotone increasing function of evolutionary distance, the $\kappa$ values cannot satisfy a molecular clock.

The phenomena described by part 2 of Theorem 3.1 cannot arise under the simple two-parameter Felsenstein zone setting for which $x_1 = x_3$ and $x_2 = x_4 = x_5$ (i.e., two long equal-length edges and three short equal-length edges).

The region for which we have exact additivity (either on the true tree $\mathcal{T}_1$ or on an alternative tree $\mathcal{T}_2$) requires an equality to hold and so is unlikely to hold exactly in practice. Nevertheless, the result suggests there will be situations where one will be very close (perhaps indistinguishably close) to additivity.

*Example.*—Suppose $x = (x_1, x_2, x_3, x_4, x_5) = (0.95, 0.5, 0.75, 0.7, 0.9615385)$. This value of $x$ satisfies the conditions described in part 2 of Theorem 3.1. Given a tree based on these values, we used the program SeqGen (Rambaut and Grassly, 1997) to generate 100 simulated data sets for each of the sequence lengths 100, 500, 1,000,

TABLE 1. For each of the sequence lengths, we generated 100 simulated data sets on the tree described in the text. For uncorrected and corrected distances, we list the percentage of true trees and false trees recovered, respectively. We also report the mean *tree-likeness* score: $1 - (S/L)$.

| | Uncorrected distances | | | | Corrected distances | | | |
|---|---|---|---|---|---|---|---|---|
| | True tree | | False tree | | True tree | | False tree | |
| Length | % | Mean score | % | Mean score | % | Mean score | % | Mean score |
| 100 | 22 | 0.69 | 78 | 0.56 | 51 | 0.62 | 49 | 0.53 |
| 500 | 24 | 0.44 | 76 | 0.50 | 83 | 0.59 | 17 | 0.40 |
| 1,000 | 16 | 0.46 | 84 | 0.56 | 92 | 0.62 | 8 | 0.25 |
| 5,000 | 0 | | 100 | 0.72 | 100 | 0.77 | 0 | |
| 10,000 | 1 | 0.07 | 99 | 0.78 | 100 | 0.83 | 0 | |
| 100,000 | 0 | | 100 | 0.92 | 100 | 0.95 | 0 | |

5,000, 10,000, and 100,000. Table 1 summarizes the results, which show an increasing tree-likeness score with sequence length, both for the uncorrected distances (which converge to a false tree) and the corrected distances (which converge to the true tree).

*An extension.*—The region of perfect inconsistency is not particular to the equal input model; it may also arise when distances are corrected according to an incorrect model. For most models of site substitution where uncorrected or miscorrected distances can be inconsistent, there is a corresponding and nonempty region of perfect inconsistency. By a standard model, we mean any model of nucleotide substitution for which $p(i, j)$ can be written as some function $f$ of the total evolutionary distance (sum of the edge lengths) separating sequences $i$ and $j$. Virtually all models used in molecular systematics are standard, including the general time reversible (GTR) model (with or without rate variation) and covarion-type models. To emphasize the dependence of $p$ on $\kappa = (\kappa_1, \ldots, \kappa_5)$, we will write $p_\kappa$. Suppose that $g$ is some continuous function that is applied to $d_{\mathcal{C}}$. We think of $g$ either as the identity function (i.e., uncorrected distances) or some function that attempts to correct distances but is subject to an incorrect model (perhaps due to undercorrection).

For any distance function $d$ on $X = \{1, 2, 3, 4\}$ and ordering $i, j, k, l$ of $X$, let $s_{ij|kl}(d) = g[d(i, j)] + g[d(k, l)]$, and let $S(d) = [s_{12|34}(d), s_{13|24}(d), s_{14|23}(d)]$. For a sequence $\mathcal{C}$ of characters on the set $X = \{1, 2, 3, 4\}$, most distance methods, such as neighbor joining, applied to the distances $g[d_{\mathcal{C}}(i, j)]$ will select tree $ij|kl$ precisely if $s_{ij|kl}(d_{\mathcal{C}})$ is the smallest value of $S(d_{\mathcal{C}})$ (Gascuel, 1997).

To investigate inconsistency, we replace $d_{\mathcal{C}}$ in the above expressions by its expected value, $p_\kappa$. Let $\min S(p_\kappa)$ and $\max S(p_\kappa)$ denote the minimal and maximal elements of $S(p_\kappa)$. Then $\kappa$ is a strong inconsistency parameter value under $g$-correction if the following two conditions hold:

1. $s_{12|34}(p_\kappa) \notin \{\min S(p_\kappa), \max S(p_\kappa)\}$ $(SI_1)$.
2. If $\kappa_i' = \kappa_i$ for $i = (3, 4, 5)$ and $\kappa_1'$ and $\kappa_2'$ both lie between $\kappa_1$ and $\kappa_2$, then $s_{12|34}(p_{\kappa'}) \neq \min S(p_{\kappa'})$ $(SI_2)$.

One motivation for this definition is that strong inconsistency values arise in classic Felsenstein zone settings with uncorrected distances. For example, under the equal input model, if we set $x_1 = x_3 = \epsilon$ (long edges) and $x_2 = x_4 = x_5 = 1 - \epsilon$ (short edges), then for a sufficiently small positive value of $\epsilon$ it can be verified that conditions $SI_1$ and $SI_2$ both hold.

For a range of models, and at least some functions $g$ (e.g., those that are close to the identity function, which correspond to undercorrection), we may therefore expect there to be a corresponding strong inconsistency parameter value $\kappa$. In such cases, the following result guarentees there is a corresponding nonempty region of perfect inconsistency.

**Theorem 3.2**

For sequences generated by the standard model on $\mathcal{T}_1$, suppose that there is a strong inconsistency parameter value $\kappa$ under $g$-correction. Then, there exists *a* nonempty region of perfect inconsistency for this model, i.e., parameters for which the expected sequence dissimilarities, after they have been corrected according to $g$, are perfectly additive on a tree that is different from $\mathcal{T}_1$.

## CONCLUDING COMMENTS

For topology estimation under clocklike models, the use of $d_{\mathcal{C}}$ may be preferable to corrected distances that have higher variance or when the fine details of the model are unclear. The catch is the invocation of a molecular clock. Where evolution is not clocklike, the raw distance measure $d_{\mathcal{C}}$ can be seriously misleading. These distances can favor incorrect topologies. We have demonstrated that this can occur without any hint that there is a problem in terms of conflicting signal. Techniques such as SplitsTree (Huson, 1998) often provide a useful visual representation of conflicting phylogenetic signal present in distance data. Generally, the absence of conflict should provide some confidence in the output, because the conditions for perfect inconsistency are quite special. Nevertheless, that there even exists such a region should be taken as a caution. The message is simple: a (near) perfect fit of raw distance data to some tree does not necessarily confer confidence in the accuracy of the tree.

## ACKNOWLEDGMENTS

We thank the New Zealand Institute for Mathematics and Its Applications (Phylogenetic Genomics Programme). We also thank Junhyong Kim, Kevin Atteson, and an anonymous reviewer for several helpful comments on an earlier version of this manuscript.

## REFERENCES


Bandelt, H.-J., and A. W. M. Dress. 1992. A canonical decomposition theory for metrics on a finite set. Adv. Math. 92:47–105.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27:401–410.

Gascuel, O. 1997. Concerning the NJ algorithm and its unweighted version, UNJ. Pages 149–170 *in* DIMACS series in discrete mathematics and theoretical computer science, Volume 37. Mathematical hierarchies and biology (B. Mirkin, F. R. McMorris, F. S. Roberts, and A. Rzhetsky, eds.). American Mathematical Society, Providence, Rhode Island.

Holland, B., K. Huber, A. W. M. Dress, and V. Moulton. 2002. Delta-plots: A tool for the analysis of phylogenetic distance data. Mol. Biol. Evol. 19:2051–2059.

Holland, B., D. Penny, and M. D. Hendy. 2003. Outgroup misplacement and phylogenetic inaccuracy under a molecular clock—A simulation study. Syst. Biol. 52:229–238.

Huson, D. H. 1998. SplitsTree: A program for analyzing and visualizing evolutionary data. Bioinformatics 14(10):68–73.

Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Pages 21–132 *in* Mammalian protein evolution (H. N. Munro, ed.). Academic Press, New York.

Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50:913–925.

Rambaut, A., and N. C. Grassly. 1997. An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Comput. Appl. Biosci. 13:235–238.

Semple, C., and M. Steel. 2003. Phylogenetics. Oxford Univ. Press, New York.

Steel, M., and D. Penny. 2000. Parsimony, likelihood and the role of models in molecular phylogenetics. Mol. Biol. Evol. 17:839–850.

Tajima, F., and M. Nei. 1984. Estimation of evolutionary distance between nucleotide sequences. Mol. Biol. Evol. 1:269–285.

Tuffley, C., and M. Steel. 1997. Links between maximum likelihood and maximum parsimony under a simple model of site substitution. Bull. Math. Biol. 59:581–607.


*First submitted 27 August 2003; reviews returned 17 November 2003; final acceptance 28 November 2003*

*Associate Editor: Junhyong Kim*

## APPENDIX
## PROOF OF RESULTS

### *Proof of Theorem 2.1*

Throughout this proof, let $n = |X|$, $B = \binom{n}{2} - 1$. Let $D_1(i, j)$ denote the number of interior edges in $\mathcal{T}_1$ that separate $i$ and $j$. On $\mathcal{T}_2$, assign edge weight 1 to all interior edges, and assign to each exterior (pendant) edge a nonnegative integer edge weight in such a way that the induced metric on $X$ (denote by $D_2$) is an ultrametric. Thus, $D_2$ is additive on $\mathcal{T}_2$ and no other tree, and we can represent $D_2$ on $\mathcal{T}_2$ with a edgeweighting that satisfies a clock (i.e., the distance from some vertex or the midpoint of some edge to all the leaves is the same).

Let $S = \sum_{i,j}[D_2(i, j) - D_1(i, j)]$, and select a positive integer $r$ sufficiently large that for all $i, j \in X$ we have

$$r - B[D_2(i, j) - D_1(i, j)] + S \geq 0. \tag{1}$$

Let $\mathcal{C}_1$ be the set of $n - 3$ binary characters (one for each interior edge of $\mathcal{T}_1$) that correspond to the splits of $\mathcal{T}_1$. Thus, $\mathcal{T}_1$ is the only phylogenetic tree on $X$ on which $\mathcal{C}_1$ is homoplasy free. For a pair of species $i, j \in X$, a character is of type $ij$ if $i$ and $j$ are the only two species in $X$ that are assigned the same character state.

Let $\mathcal{C}_2$ consist of the following sequence of characters. For each distinct pair $i, j \in X$, place $n_{ij}$ characters of type $ij$ in $\mathcal{C}_2$ where

$$n_{ij} = r - B[D_2(i, j) - D_1(i, j)] + S \tag{2}$$

(by Eq. 1 this is a nonnegative integer). Let $\mathcal{C}$ be the concatenation of $\mathcal{C}_2$ and $B$ copies of $\mathcal{C}_1$. A character of type $ij$ is homoplasy free on any phylogenetic tree on $X$. Thus, $\mathcal{C}$ is homoplasy free on $\mathcal{T}_1$ and only on $\mathcal{T}_1$. It is easily checked that for each distinct pair $i, j \in X$, if we let $D_{\mathcal{C}}$ be the (nonnormalized) Hamming distance defined by $D_{\mathcal{C}}(i, j) = |[s : f_s(i) \neq f_s(j)]|$, then

$$D_{\mathcal{C}}(i, j) = B \cdot D_1(i, j) + N - n_{ij}, \tag{3}$$

where $N := \sum_{i,j} n_{ij}$. Let $D'_2(i, j) = B \cdot D_2(i, j) + r B$ for each $i, j \in X$. Then, $D'_2$ is an ultrametric that is additive on $\mathcal{T}_2$ and no other tree (because $D_2$ has this property). Furthermore, substituting Equation 2 into Equation 3 and reducing the resulting equation gives $D_{\mathcal{C}}(i, j) = D'_2(i, j)$ for all $i, j \in X$. Thus, $D_{\mathcal{C}}(i, j)$ is precisely the distance in $T_2$ with the edge weighting as specified. Consequently, $D_{\mathcal{C}}$ and thereby $d_{\mathcal{C}}$ are perfectly additive on $\mathcal{T}_2$, as claimed. By the assumption that $\mathcal{T}_2$ is binary and because all the edge weight assignments described are positive, it follows that $d_{\mathcal{C}}$ is additive only on $\mathcal{T}_2$. This completes the proof.

### *Proof of Theorem 3.1*

This proof relies on elementary algebra, with repeated application of the formal identity:

$$(x_i - x_j)(x_k - x_l) = x_i x_k + x_j x_l - (x_i x_l + x_j x_k) \tag{4}$$

together with the relationship

$$p(i, j) = p_\infty \left(1 - \prod_{k \in P(\mathcal{T}_1; i, j)} x_k\right), \tag{5}$$

where $P(\mathcal{T}_1; i, j)$ is the set of edges in $\mathcal{T}_1$ on the path connecting $i$ and $j$ (Semple and Steel, 2003).

*Proof of part 1.*—By the four-point condition, $p$ is additive (only) on $\mathcal{T}_1$ precisely if

$$p(1, 2) + p(3, 4) < p(1, 3) + p(2, 4) = p(1, 4) + p(2, 3),$$

and this translates (by virtue of Eq. 5) to the conditions

$$(x_1 x_3 + x_2 x_4)x_5 = (x_1 x_4 + x_2 x_3)x_5 \tag{6}$$

and

$$x_1 x_2 + x_3 x_4 > (x_1 x_3 + x_2 x_4)x_5. \tag{7}$$

Now, applying Equation 4 (for $i, j, k, l = 1, 2, 3, 4$), Equation 6 is equivalent to the condition $(x_1 - x_2)(x_3 - x_4) = 0$, that is, $x_1 = x_2$ or $x_3 = x_4$ (or both). Similarly, Equation 7 is equivalent to the condition: $x_5 < (x_1 x_2 + x_3 x_4)/(x_1 x_3 + x_2 x_4)$. These are the conditions described in part 1 of Theorem 3.1.

*Proof of part 2.*—By the four-point condition, $p$ is additive (only) on $\mathcal{T}_2$ precisely if

$$p(1, 3) + p(2, 4) < p(1, 2) + p(3, 4) = p(1, 4) + p(2, 3),$$

and this translates to the condition

$$(x_1 x_3 + x_2 x_4)x_5 > x_1 x_2 + x_3 x_4 = (x_1 x_4 + x_2 x_3)x_5. \tag{8}$$

Comparing the first and last term in Equation 8, we obtain $x_1 x_3 + x_2 x_4 - x_1 x_4 - x_2 x_3 > 0$, which by Equation 4 is equivalent to the condition $(x_1 - x_2)(x_3 - x_4) > 0$. The equality of the last two terms in Equation 8 is equivalent to the condition $x_5 = (x_1 x_2 + x_3 x_4)/(x_1 x_4 + x_2 x_3)$. This completes the proof.

### *Proof of Theorem 3.2*

Let $\kappa$ be a strong inconsistency value under $g$-correction. Without loss of generality, by $SI_1$ we have

$$s_{13|24}(p_\kappa) < s_{12|34}(p_\kappa) < s_{14|23}(p_\kappa).$$

Let $\kappa^*$ be the parameter value obtained by interchanging $\kappa_1$ and $\kappa_2$ in $\kappa$:

$$s_{12|34}(p_{\kappa^*}) = s_{12|34}(p_\kappa)$$

$$s_{13|24}(p_{\kappa^*}) = s_{14|23}(p_\kappa)$$

and

$$s_{14|23}(p_{\kappa^*}) = s_{13|24}(p_\kappa).$$

For $t \in [0, 1]$, let $\kappa^t = t\kappa + (1-t)\kappa^*$ and let $h(t) = s_{14|23}(p_{\kappa^t}) - s_{12|34}(p_{\kappa^t})$. Thus, $h(0) < 0$, $h(1) > 0$, and so. Because $h$ is continuous, there exists a value $t_0 \in (0, 1)$ for which $h(t_0) = 0$. Let $\kappa^0 = \kappa^{t_0}$. Note that $\kappa_i^0 = \kappa_i$ for $i = (3, 4, 5)$, and so by $(SI_2)$ we have

$$s_{13|24}(p_{\kappa^0}) < s_{12|34}(p_{\kappa^0}) = s_{14|23}(p_{\kappa^0}),$$

and so $\kappa^0$ is a point of perfect inconsistency. This completes the proof.