Not all phylogenetic networks are leaf-reconstructible

Erdős, Péter L.; van Iersel, Leo; Jones, Mark

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Mathematical Biology

Not all phylogenetic networks are leaf-reconstructible

Péter L. Erdős[1] · Leo van Iersel[2] · Mark Jones[2]

Abstract
Unrooted phylogenetic networks are graphs used to represent reticulate evolutionary relationships. Accurately reconstructing such networks is of great relevance for evolutionary biology. It has recently been conjectured that all unrooted phylogenetic networks for at least five taxa can be uniquely reconstructed from their subnetworks obtained by deleting a single taxon. Here, we show that this conjecture is false, by presenting a counter-example for each possible number of taxa that is at least 4. Moreover, we show that the conjecture is still false when restricted to binary networks. This means that, even if we are able to reconstruct the unrooted evolutionary history of each proper subset of some taxon set, this still does not give us enough information to reconstruct their full unrooted evolutionary history.

Keywords Graph reconstruction · Phylogenetics · Undirected graphs · Leaf removal · Ulam's Conjecture · Phylogenetic Networks

Mathematics Subject Classification 05C60 Isomorphism problems · 92D15 Problems related to evolution

✉ Mark Jones
  markelliotlloyd@gmail.com

  Péter L. Erdős
  erdos.peter@renyi.mta.hu

  Leo van Iersel
  L.J.J.vanIersel@tudelft.nl

[1] Alfréd Rényi Institute of Mathematics, Reáltanoda u 13–15, Budapest 1053, Hungary

[2] Delft Institute of Applied Mathematics, Delft University of Technology, Van Mourik Broekmanweg 6, 2628 XE Delft, The Netherlands

1 Introduction

The reconstruction conjecture, introduced in 1941 by Kelly and Ulam (see Bondy and Hemminger 1977), conjectures that each graph with at least three vertices is uniquely reconstructable from its multiset of vertex-deleted subgraphs. Despite more than seven decades of research, the conjecture is still open.

Recently, a variant of this conjecture was introduced that is relevant for the field of *phylogenetics*, the study of evolutionary relationships. Such relationships among a set X of entities (e.g. biological species or languages) are traditionally described by a tree with no degree-2 vertices and its leaves bijectively labelled by the elements of X; this is called a *phylogenetic tree* on X. More recently, evolutionary histories are more and more often described by phylogenetic *networks* (Bapteste et al. 2013), which are basically (directed or undirected) graphs with their leaves bijectively labelled by the elements of X. These networks are able to describe more complex evolutionary relationships than trees.

To find out whether it may be possible to accurately reconstruct phylogenetic networks, an important question to answer is which substructures uniquely define a phylogenetic network. For example, although there is much research directed at reconstructing rooted phylogenetic networks from embedded trees [see e.g. Van Iersel et al. (2016) and Whidden et al. (2013)], these trees do not uniquely define a network [see e.g. Pardi and Scornavacca (2015)]. Hence, no method based on embedded trees can be guaranteed to reconstruct the right network, even when it gets error-free and complete trees as input. Moreover, it has recently been shown that rooted phylogenetic networks also cannot be reconstructed uniquely from their subnetworks obtained by deleting one or more leaves and transforming the result into a valid rooted phylogenetic network (Huber et al. 2014). A similar reconstruction question for pedigrees has also been answered negatively (Thatte 2008).

Here, we focus on *unrooted phylogenetic networks*, which are undirected graphs with leaves labelled by the elements of some taxon set X. Although real evolutionary histories are rooted, it is not always possible to identify the root location and the directions of all arcs. Therefore, just like unrooted phylogenetic trees are studied in addition to rooted phylogenetic trees, unrooted phylogenetic networks are studied increasingly. Van Iersel and Moulton (2018) studied reconstructing such networks from their X-deck, which consists of the graphs obtained by deleting a single taxon from the network (see Fig. 1 for an example). Several promising results were obtained, including a proof that all phylogenetic trees and all decomposable networks (i.e. networks that can be decomposed into two nontrivial subnetworks by deleting a single edge) are reconstructable from their X-deck, assuming $|X| \geq 5$. Moreover, the same was shown for networks that can be turned into a tree by deleting at most four edges, and for all networks with sufficiently many leaves. The only known networks not reconstructible from their $|X|$-decks were ones for which $|X| \leq 4$. It was conjectured that all unrooted phylogenetic networks on X, with $|X| \geq 5$, can be uniquely reconstructed from their X-deck.

Here, we show that this conjecture is false. To do so, we present, for each finite set X containing at least four elements, two unrooted phylogenetic networks on X that are not isomorphic but have the same X-deck. Moreover, we also give binary networks with
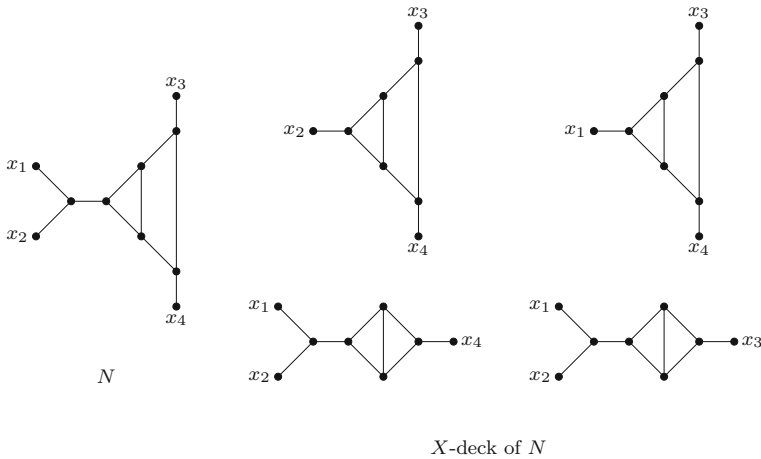
X-deck of N

Fig. 1 Example of a phylogenetic network N and its X-deck

these properties, hence showing that the conjecture restricted to binary networks is still false. These results can be seen as the unrooted counterpart to the results from Huber et al. (2014). However, we also note that there are important differences between the rooted and unrooted case, which make it impossible to directly transform the rooted counter-examples to the unrooted case, see Sect. 2.1.

Our result may have consequences for developing "supernetwork" methods, which attempt to reconstruct phylogenetic networks from subnetworks. Supertree methods work well for phylogenetic trees, which can be explained from the fact that a phylogenetic tree is uniquely determined by its induced set of four-leaved trees (or three-leaved trees in the case of rooted trees). Since phylogenetic networks are not uniquely determined by their subnetworks, developing supernetwork methods will be significantly more challenging than in the tree-case, even for unrooted networks.

The structure of the paper is as follows. We start off by giving formal definitions related to phylogenetic networks and binary sequences, which are central to the construction of our counter-examples, in Sect. 2. In Sect. 2.1, we explain why unrooting the counter-example for the rooted case from Huber et al. (2014) does not give a counter-example for the conjecture considered here. Then, in Sect. 3, we present our counter-examples for the unrooted, non-binary case. Finally, in Sect. 4 we show how these can be transformed into counter-examples for the unrooted, binary case.

2 Preliminaries

A *phylogenetic tree* on X is an undirected simple tree, with no degree-2 vertices, such that each leaf is bijectively labelled by an element from X. A *biconnected component* of a graph is a maximal 2-edge-connected subgraph and it is called a *blob* if it contains at least two edges. Let X be a finite set with $|X| \geq 2$, and let N be an undirected simple graph in which the *leaves* (degree-1 vertices) are bijectively labelled by the elements

of X. We say N is an *unrooted phylogenetic network* on X if contracting each blob into a single vertex gives phylogenetic tree (or equivalently, each cut-edge induces a unique partition of the leaves). In addition, we say that N is *binary* if every vertex has degree 1 or 3. In what follows, we will refer to unrooted phylogenetic networks as *networks* for short.

Let G and H be two partially labelled undirected multigraphs with the same label set, such that $|V(G)| = |V(H)|$. Let $f : V(G) \to V(H)$ be a bijective function. We say that f is an *isomorphism* between G and H if it is both label-preserving (that is, vertex $a \in V(G)$ has label l if and only if $f(a)$ has label l) and edge-preserving (that is, for any $a, b \in V(G)$ the number of edges between a and b in G is equal to the number of edges between $f(a)$ and $f(b)$ in H). We say G and H are *equivalent*, denoted $G \sim H$, if there is an isomorphism between G and H.

Given an undirected multigraph G with no vertices of degree 2, and a vertex $a \in V(G)$, we denote by G_a the undirected multigraph derived from G by deleting a and all incident edges, and then suppressing any degree-2 vertices. We say G_a is derived from G by *removing* the vertex a. For a label x, we may write G_x to refer to G_a, where a is the unique vertex in G with label x.

Given a network N on X, an X-reconstruction of N is a network N' on X such that $N'_x \sim N_x$ for all $x \in X$. We call a phylogenetic network N *leaf-reconstructible* if $N' \sim N$ for every X-reconstruction N' of N. That is, all X-reconstructions of N are isomorphic to each other.

It was conjectured in Van Iersel and Moulton (2018) that all unrooted phylogenetic networks with 5 or more leaves are leaf-reconstructible. (We note that phylogenetic trees on 5 or more leaves are leaf-reconstructible, as it is clearly possible to reconstruct every quartet in the tree.)

In this paper, we show that the conjecture is false. More precisely, we will show that for each $r \geq 4$, there exist binary unrooted phylogenetic networks N and N' on X with $|X| = r$, such that $N \nsim N'$, but $N_x \sim N'_x$ for all $x \in X$. Thus, N and N' are not leaf-reconstructible.[1]

Finally, for an integer k, let $[k]$ denote the set $\{1, 2, \ldots, k\}$.

2.1 Unrooting the rooted counter-example

Huber et al. (2014) showed that for any $r \geq 3$, there exist rooted binary networks M and M' on X with $|X| = r$, such that $M \nsim M'$, but $M|_{X'} \sim M|_{X'}$ for any strict subset X' of X. Here $M|_{X'}$ denotes the subnet of M *induced* by X'; roughly speaking, $M|_{X'}$ is derived from M by deleting any vertices not on a directed path from the root to an element of X', then suppressing any degree-2 vertices and parallel arcs [see Huber et al. (2014) for full details].

We note that one cannot create a counterexample to the leaf-reconstruction conjecture by simply taking the directed networks M, M' given by Huber et al. and replacing them with their underlying undirected graphs G, G'. A key observation here is that

---

[1] It was previously known that networks on $r = 4$ leaves are not leaf-reconstructible in general. We nevertheless include the case $r = 4$ in our paper, as it allows us to give simpler figures than for the $r = 5$ case.
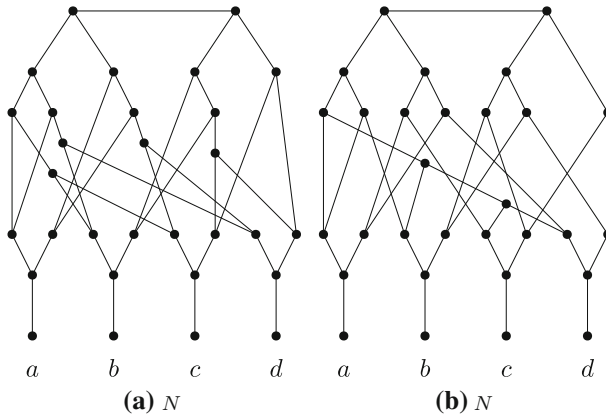
Fig. 2 The underlying undirected networks N and N' of two rooted networks which, in Huber et al. (2014), were shown to have the same induced subnetworks for any strict subset of the leaf set $X = \{a, b, c, d\}$. We observe that $N_d \not\simeq N'_d$, since the shortest path between a and b has length 7 in N_d and length 6 in N'_d. Hence, these networks can not be used as counter-example for the leaf-reconstruction conjecture

for any $x \in X$, the network $M|_{X \setminus \{x\}}$ may have many fewer vertices and arcs than M, whereas the graph G_x has at most two fewer edges and two fewer vertices than G. Indeed, Fig. 2 gives two networks N, N' on $X = \{a, b, c, d\}$ that correspond to the undirected versions (after suppressing degree-2 vertices) of the networks given by Huber et al. for $r = 4$. We observe that the distance between a and b is 7 in N_d, and 6 in N'_d, and thus these networks do not have the same X-deck. Thus the approach of Huber et al. cannot be naively used to give our result. However, the two papers do use similar ideas, in particular the use of binary sequences in the construction of a network (see Sect. 2.2).

2.2 Binary sequences

Given an alphabet Σ, let $w \in \Sigma^*$ be a sequence of elements with elements drawn from Σ. If $\Sigma = \{0, 1\}$ then we call w a *binary sequence*. The *length* of the sequence w, denoted $l(w)$, is the number of elements in w. We write w_i to denote the i'th element of w. We often write $e_1 e_2 \ldots e_l$ to denote the sequence w such that $l(w) = l$ and $w_i = e_i$ for each $i \in [l]$. (Thus, for example, 1011 denotes the length-4 binary sequence whose second element is 0 and whose first, third and fourth elements are 1.) Given a binary sequence w, the *weight* of w is the number of 1's in w. For an integer l, we write $\mathcal{B}_l$ to denote the set of binary sequences of length l. Given a sequence $w \in \mathcal{B}_r$ and $i \in [r]$, let $w^{\sim i}$ be the sequence derived from w by replacing the i'th element with $1 - w_i$ (for example, if $w = 1001$ and $i = 3$, then $w^{\sim i} = 1011$).

Central to the proof of our result is the idea that for a binary sequence w, one needs to know all elements of w in order to decide whether w has odd or even weight. (Note that here and in the rest of the paper, we consider a sequence of weight 0 to have even weight.) For some integer r, consider the set $\mathcal{B}_r^{\text{even}}$ of all length-r binary sequences of even weight, and the set $\mathcal{B}_r^{\text{odd}}$ of all length-r binary sequences of odd

weight. Given a length-r binary sequence w and integer $i \in [r]$, let w^{-i} denote the sequence on $\{0, 1, *\}$ derived from w by replacing the i'th element with $*$. Then for each $w \in \mathcal{B}_r^{\mathrm{even}}$, there exists a sequence $w' \in \mathcal{B}_r^{\mathrm{odd}}$ such that $(w')^{-i} = w^{-i}$ (indeed, $w^{\sim i}$ is such a sequence). For a set of sequences S and $i \in [r]$, let $\mathcal{B}_r^{-i} = \{w^{-i} ; w \in S\}$. Then it follows that for each $i \in [r]$, the sets $(\mathcal{B}_r^{\mathrm{odd}})^{-i}$ and $(\mathcal{B}_r^{\mathrm{even}})^{-i}$ are the same.

We will use this concept to guide our construction of two networks N^{even} and N^{odd} on a set $X = \{x_1, \ldots, x_r\}$. Roughly speaking, N^{even} can be thought of as a representation of $\mathcal{B}_r^{\mathrm{even}}$, and N^{odd} can be thought of as a representation of $\mathcal{B}_r^{\mathrm{odd}}$. Then for each $i \in [r]$, $(N^{\mathrm{even}})_{x_i}$ corresponds to $(\mathcal{B}_r^{\mathrm{even}})^{-i}$, and $(N^{\mathrm{odd}})_{x_i}$ corresponds to $(\mathcal{B}_r^{\mathrm{odd}})^{-i}$. Just as $(\mathcal{B}_r^{\mathrm{even}})^{-i} = (\mathcal{B}_r^{\mathrm{odd}})^{-i}$, we will be able to show that $(N^{\mathrm{even}})_{x_i}$ and $(N^{\mathrm{odd}})_{x_i}$ are equivalent, while originally N^{even} and N^{odd} are different.

3 A non-binary example

In order to demonstrate the main concepts of our construction, we first give a construction using non-binary graphs. In the next section, we will construct an example with binary phylogenetic networks, using these non-binary graphs as a guide.

For some integer $r \geq 4$, let X denote the set of labels $\{x_1, \ldots, x_r\}$. We will construct two graphs M^{even} and M^{odd}, in which the leaves are bijectively labelled by the elements of X. As in the previous section, let $\mathcal{B}_r^{\mathrm{even}}$ denote the set of all length-r binary sequences of even weight, and let $\mathcal{B}_r^{\mathrm{odd}}$ denote the set of all length-r binary sequences of odd weight.

The graph M^{even} is constructed as follows. For each $i \in [r]$, let M^{even} contain vertices $v_{i,0}$ and $v_{i,1}$, and a leaf labelled with x_i, such that x_i is adjacent to $v_{i,0}$.[2] For each $w \in \mathcal{B}_r^{\mathrm{even}}$, let M^{even} contain a vertex u_w. For each $w \in \mathcal{B}_r^{\mathrm{even}}$ and $i \in [r]$, let u_w be adjacent to $v_{i,0}$ if $w_i = 0$, and let u_w be adjacent to $v_{i,1}$ if $w_i = 1$. This completes the construction of M^{even} (see Fig. 3a).

The construction of M^{odd} is identical to that of M^{even}, except that we have a vertex u_w for each $w \in \mathcal{B}_r^{\mathrm{odd}}$ rather than each $w \in \mathcal{B}_r^{\mathrm{even}}$. For completeness, the full construction is as follows: For each $i \in [r]$, let M^{odd} contain vertices $v_{i,0}$ and $v_{i,1}$, and a leaf labelled with x_i, such that x_i is adjacent to $v_{i,0}$. For each $w \in \mathcal{B}_r^{\mathrm{odd}}$, let M^{odd} contain a vertex u_w. For each $w \in \mathcal{B}_r^{\mathrm{odd}}$ and $i \in [r]$, let u_w be adjacent to $v_{i,0}$ if $w_i = 0$, and let u_w be adjacent to $v_{i,1}$ if $w_i = 1$. This completes the construction of M^{odd} (see Fig. 3b).

Lemma 1 *M^{even} and M^{odd} are not equivalent.*

Proof Suppose for a contradiction that M^{even} and M^{odd} are equivalent, and let $f : V(M^{\mathrm{even}}) \to V(M^{\mathrm{odd}})$ be an isomorphism between M^{even} and M^{odd}. Let $\mathbf{0}$ denote the all-0 sequence from $\mathcal{B}_r^{\mathrm{even}}$. Observe that for each $i \in [r]$, the distance between $u_{\mathbf{0}}$ and x_i is 2 (as both $u_{\mathbf{0}}$ and x_i are adjacent to $v_{i,0}$). It follows that $f(u_{\mathbf{0}})$ must have distance 2 to $f(x_i) = x_i$ in M^{odd}, for each $i \in [r]$. We will show that no such $f(u_{\mathbf{0}})$ exists in M^{odd}, a contradiction to the existence of f.

---

[2] We note that in this section and next, we will often give names to particular vertices in the graphs we construct. This is done to differentiate between vertices, in order to aid in the description of the construction and help define isomorphisms. However, this is not the same as labelling the vertices; the only labelling that will occur is the labelling of leaves with elements of X.
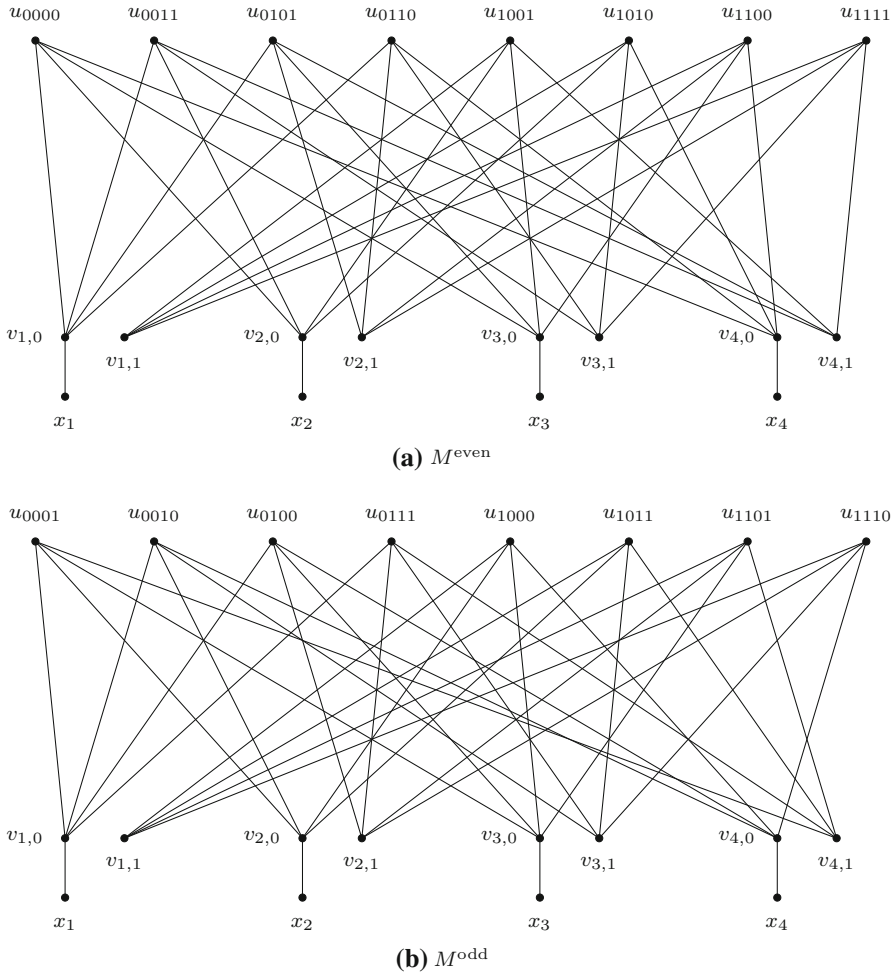
(a) M^{even}

(b) M^{odd}

Fig. 3 Non-binary example of M^{even} and M^{odd} for the case when when $r = 4$. Vertices u_w are adjacent to vertices $v_{i,h}$ if and only if $w_i = h$

Observe that by construction of M^{odd} (in particular, the fact that it is a bipartite graph with one side consisting of vertices $v_{j,0}$ or $v_{j,1}$), the distance between any leaf x_i and any vertex $v_{j,0}$ or $v_{j,1}$ is odd. It follows that $f(u_{\mathbf{0}})$ must be the vertex u_w, for some $w \in \mathcal{B}_r^{\text{odd}}$ (any other vertex is either a leaf, which has distance 0 from itself, or has odd distance from any leaf). However, for any $w \in \mathcal{B}_r^{\text{odd}}$ there exists $i \in [r]$ such that $w_i = 1$, and so u_w is not adjacent to $v_{i,0}$. As $v_{i,0}$ is the only vertex adjacent to x_i, it follows that the distance between u_w and x_i is greater than 2, and so $f(u_{\mathbf{0}}) \neq u_w$.

As there is no choice for $f(u_{\mathbf{0}})$ that satisfies the conditions of an isomorphism, we have that there is no possible isomorphism between M^{even} and M^{odd}, and so M^{even} and M^{odd} are not equivalent. □

Lemma 2 *For each $i \in [r]$, $(M^{\text{even}})_{x_i} \sim (M^{\text{odd}})_{x_i}$.*

Proof Observe that $v_{i,0}$ and $v_{i,1}$ each have $2^{r-2} \geq 4$ neighbors in M^{even} not including x_i (as $|\mathcal{B}_r^{\text{even}}| = 2^{r-1}$ and exactly half of the sequences in $\mathcal{B}_r^{\text{even}}$ have 1 as their i'th element). Also any vertex u_w has $r \geq 4$ neighbors in M^{even}. It follows that if x_i is deleted from M^{even}, the remaining graph has no vertices of degree 2, and thus $(M^{\text{even}})_{x_i}$ is exactly M^{even} with x_i deleted. By a similar argument, $(M^{\text{odd}})_{x_i}$ is exactly M^{odd} with x_i deleted.

Now define a bijective function $f : V((M^{\text{even}})_{x_i}) \to V((M^{\text{odd}})_{x_i})$ as follows. For each $w \in \mathcal{B}_r^{\text{even}}$, let $f(u_w) = u_{w^{\approx i}}$. Observe that this defines a bijection between $\{u_w : w \in \mathcal{B}_r^{\text{even}}\}$ and $\{u_w : w \in \mathcal{B}_r^{\text{odd}}\}$. Let $f(v_{i,0}) = v_{i,1}$ and $f(v_{i,1}) = v_{i,0}$. For $j \in [r] \setminus \{i\}$, let $f(v_{j,0}) = v_{j,0}$, $f(v_{j,1}) = v_{j,1}$ and $f(x_j) = x_j$ (recall that the leaf x_i does not appear in $(M^{\text{even}})_{x_i}$ or $(M^{\text{odd}})_{x_i}$, so we do not need to define $f(x_i)$).

By construction, f is a bijective function from $V((M^{\text{even}})_{x_i})$ to $V((M^{\text{odd}})_{x_i})$. It remains to show that f is label-preserving and edge-preserving. As f is the identity on all labelled vertices, f is label-preserving. As $(M^{\text{even}})_{x_i}$ and $(M^{\text{odd}})_{x_i}$ are simple graphs, to show that f is edge-preserving it is enough to show that two vertices a, b are adjacent in $(M^{\text{even}})_{x_i}$ if and only if $f(a)$ and $f(b)$ are adjacent in $(M^{\text{odd}})_{x_i}$.

So consider any $a, b \in V((M^{\text{even}})_{x_i})$. Suppose first that $a = u_w$ for some $w \in \mathcal{B}_r^{\text{even}}$ and that $b = v_{j,h}$ for some $j \in [r] \setminus \{i\}$ and $h \in \{0, 1\}$. Then a and b are adjacent if and only if $w_j = h$. By definition of f, we have $f(a) = u_{w^{\approx i}}$, and we note that $(w^{\approx i})_j = w_j$. Finally, we have that $f(a)$ and $f(b) = v_{j,h}$ are adjacent if and only if $(w^{\approx i})_j = h$. Putting it together, we have that $ab \in E((M^{\text{even}})_{x_i}) \Leftrightarrow w_j = h \Leftrightarrow w_j^{\approx i} = h \Leftrightarrow f(a)f(b) \in E((M^{\text{even}})_{x_i})$. Thus a and b are adjacent if and only $f(a)$ and $f(b)$ are adjacent.

Next suppose that $a = u_w$ for some $w \in \mathcal{B}_r^{\text{even}}$ and that $b = v_{i,h}$ for some $h \in \{0, 1\}$. Then a and b are adjacent if and only if $w_i = h$. Furthermore $f(a) = u_{w^{\approx i}}$ where $w_i^{\approx i} = 1 - w_i$, and $f(a)$ and $f(b) = v_{i,1-h}$ are adjacent if and only if $w_i^{\approx i} = 1 - h$. Thus $ab \in E((M^{\text{even}})_{x_i}) \Leftrightarrow w_j = h \Leftrightarrow w_j^{\approx i} = 1 - h \Leftrightarrow f(a)f(b) \in E((M^{\text{even}})_{x_i})$.

If a and b are $u_w, u_{w'}$ for some $w, w' \in \mathcal{B}_r^{\text{even}}$, then a and b are not adjacent, and neither are $f(a)$ and $f(b)$ (which are both vertices $u_{w''}, u_{w'''}$ for some $w', w'' \in \mathcal{B}_r^{\text{odd}}$). By a similar argument, if a and b are both vertices $v_{j,h}$ for some $j \in [r]$ and $h \in \{0, 1\}$, then a, b are not adjacent and $f(a), f(b)$ are not adjacent. If $b = x_j$ for some $j \in [r] \setminus j$, then a and b are adjacent if and only if $a = v_{j,0}$, which holds if and only if $f(a) = v_{j,0}$, which in turn holds if and only if $f(a)$ is adjacent to $x_j = f(b)$. This covers all possible cases, and so we have that a and b are adjacent if and only if $f(a)$ and $f(b)$ are adjacent. This completes the proof that f is an isomorphism, and so $(M^{\text{even}})_{x_i} \sim (M^{\text{odd}})_{x_i}$. $\qquad\square$

4 A binary example

In this section, we show how to construct two binary networks on X that are X-reconstructions of each other but are not equivalent, for $|X| \geq 4$. (An example of two such networks for the case when $|X| = 4$ is given in Fig. 7.) This is enough to show that networks on $r \geq 4$ leaves are not leaf-reconstructible.

Fig. 4 The caterpillar $Cat(w)$ for the case $r = 5$
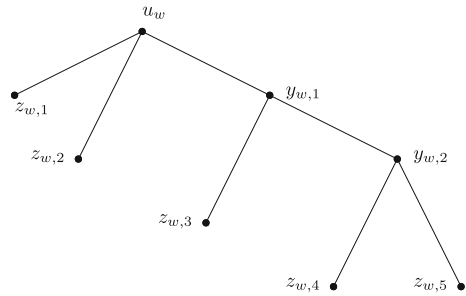
Given the non-binary networks M^{even} and M^{odd} constructed in the previous section, we proceed to construct two graphs G^{even} and G^{odd} in the following way. For each binary sequence $w \in \mathcal{B}_r$, u_w will be expanded into a caterpillar $Cat(w)$ (details of the construction are given below). Each vertex $v_{i,h}$ will be expanded into a lexicographic tree $Lex(i, h)^{\text{even}}$ or $Lex(i, h)^{\text{odd}}$ (defined below). These subgraphs contain leaves denoted $z_{w,i}$, for $w \in \mathcal{B}_r$ and $i \in [r]$. Two subgraphs $Cat(w)$ and $Lex(i, h)^{\text{even}}$ (or $Cat(w)$ and $Lex(i, h)^{\text{odd}}$) will share a vertex $z_{w,i}$ if and only if $w_i = h$ (analogous to how in M^{even} and M^{odd}, the vertices u_w and $v_{i,h}$ are adjacent if and only if $w_i = h$).

Similarly to M^{even} and M^{odd}, we will show that G^{even} and G^{odd} are not equivalent, but that they become equivalent if a single leaf x_i is deleted.

We note that G^{even} and G^{odd} are not technically networks, because while they have maximum degree 3, they contain some vertices of degree 2 (in particular, every vertex $z_{w,i}$ has degree 2). In the last part of this section, we will produce two networks N^{even} and N^{odd} from G^{even} and G^{odd}.

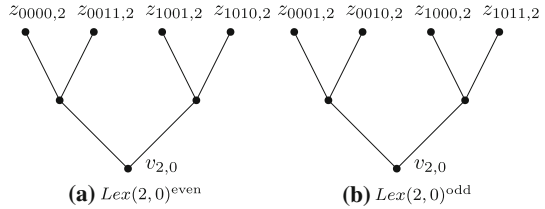We now define the two types of tree that will be used in our construction.

Definition 1 For any sequence $w \in \mathcal{B}_r$, the *caterpillar* $Cat(w)$ is the tree with internal vertices u_w and $y_{w,i}$ for each $i \in [r - 3]$, leaves $z_{w,i}$ for each $i \in [r]$, and edges $u_w z_{w,1}$, $u_w z_{w,2}$, $u_w y_{w,1}$, $y_{w,r-3} z_{w,r-1}$, $y_{w,r-3} z_{w,r}$, and $y_{w,i} z_{w,i+2}$, $y_{w,i} y_{w,i+1}$ for each $1 \leq i \leq r - 4$.

See Fig. 4 for an example. Observe that all internal vertices of $Cat(w)$ have degree 3.

Observation 1 *Given sequences $w, w' \in \mathcal{B}_r$, the trees $Cat(w)$ and $Cat(w')$ are equivalent. In particular, there exists an isomorphism f between $Cat(w)$ and $Cat(w')$ such that $f(u_w) = u_{w'}$ and $f(z_{w,i}) = z_{w',i}$ for all $i \in [r]$.*

Definition 2 Given a set S of binary sequences such that $|S| = 2^t$ for some positive integer t, and $i \in [r]$, the lexicographic tree $Lex(i, S)$ is a fully balanced binary tree with leaves $z_{w,i}$ for $w \in S$. All non-leaf vertices have degree 3 except for a single vertex, called the *root*, of degree 2, and all leaves are of distance exactly t from the root. Moreover, the leaves are arranged in such a way that there exists a depth-first search of the vertices of $Lex(i, S)$ that traverses the leaves $z_{w,i}$ in lexicographic order with respect to w. (Note that this uniquely determines $Lex(i, S)$.)

$z_{0000,2}$ $z_{0011,2}$ $z_{1001,2}$ $z_{1010,2}$ $z_{0001,2}$ $z_{0010,2}$ $z_{1000,2}$ $z_{1011,2}$

$v_{2,0}$

$v_{2,0}$

(a) $Lex(2, 0)^{even}$ **(b)** $Lex(2, 0)^{odd}$

Definition 3 Let $(\mathcal{B}_r^{even})_{i:h}$ be the set of all length-r binary sequences w of even weight such that $w_i = h$. Let $(\mathcal{B}_r^{odd})_{i:h}$ be the set of all length-r binary sequences w of odd weight such that $w_i = h$.

Definition 4 For any $i \in [r]$ and $h \in \{0, 1\}$, define $Lex(i, h)^{even} = Lex(i, (\mathcal{B}_r^{even})_{i:h})$, and define $Lex(i, h)^{odd} = Lex(i, (\mathcal{B}_r^{odd})_{i:h})$. (Thus the leaves of $Lex(i, h)^{even}$ are $z_{w,i}$ for $w \in (\mathcal{B}_r^{even})_{i:h}$, and the leaves of $Lex(i, h)^{odd}$ are $z_{w,i}$ for $w \in (\mathcal{B}_r^{odd})_{i:h}$). We refer to the root of $Lex(i, h)^{even}$ by $v_{i,h}^{even}$, and we refer to the root of $Lex(i : h)^{odd}$ by $v_{i,h}^{odd}$.

See Fig. 5 for some examples.

Lemma 3 *For any $j \in [r] \setminus \{i\}$ and $h \in \{0, 1\}$, there exists an isomorphism f between $Lex(j, h)^{even}$ and $Lex(j, h)^{odd}$ such that $f(v_{j,h}^{even}) = v_{j,h}^{odd}$, and $f(z_{w,j}) = z_{w^{\sim i},j}$ for all $w \in (\mathcal{B}_r^{even})_{j:h}$.*

Also, for any $h \in \{0, 1\}$ there exists an isomorphism f between $Lex(i, h)^{even}$ and $Lex(i, 1 - h)^{odd}$ such that $f(v_{i,h}^{even}) = v_{i,1-h}^{odd}$, and $f(z_{w,i}) = z_{w^{\sim i},i}$ for all $w \in (\mathcal{B}_r^{even})_{i:h}$.

Proof Observe that the root of a lexicographic tree is unique, as it is the only vertex of degree 2. Then for any integer l and leaf $z_{w,j}$ in a lexicographic tree, we may define the *depth-l ancestor* of $z_{w,j}$ as follows. The depth-l ancestor of $z_{w,j}$ is the unique vertex on a path between $z_{w,j}$ and the root, that has distance l from $z_{w,j}$. Note that we count the root itself as a depth-$(r - 2)$ ancestor of every leaf, and each leaf is the depth 0 ancestor of itself. Moreover, because a lexicographic tree is fully balanced, if a vertex a is the depth-l ancestor of one leaf and the depth-l' ancestor of another leaf then $l = l'$.

In order to prove the first claim, we first show that for any two sequences $w, w' \in (\mathcal{B}_r^{even})_{j:h}$ and integer l, the leaves $z_{w,j}, z_{w',j}$ share a depth-l ancestor in $Lex(j, h)^{even}$ if and only if $z_{w^{\sim i},j}, z_{w'^{\sim i},j}$ share a depth-l ancestor in $Lex(j, h)^{odd}$. Indeed, it is easy to see that $z_{w,j}, z_{w',j}$ share a depth-l ancestor if and only if w, w' agree on the first $r - 2 - l$ elements not including j. But if w, w' agree on these elements then so do $w^{\sim i}, w'^{\sim i}$, and so $z_{w^{\sim i},j}, z_{w'^{\sim i},j}$ also share a depth-l ancestor.

Thus, we may define a bijective function $f : V(Lex(j, h)^{even}) \to V(Lex(j, h)^{odd})$ as follows. For any vertex $a \in V(Lex(j, h)^{even})$ with distance $r - 2 - l$ from the root, choose any sequence $w \in (\mathcal{B}_r^{even})_{j:h}$ such that a is a depth-l ancestor of $z_{w,j}$, and let $f(a)$ be the depth-l ancestor of $z_{w^{\sim i},j}$ in $Lex(j, h)^{odd}$. Observe that f is well-defined, since we have just shown that if two leaves $z_{w,j}, z_{w',j}$ share a as a depth-l ancestor, then $z_{w^{\sim i},j}, z_{w'^{\sim i},j}$ also have the same depth-l ancestor.

By construction, it is clear that $f(v_{j,h}^{\text{even}}) = v_{j,h}^{\text{odd}}$, and $f(z_{w,j}) = z_{w^{\approx i},j}$ for all $w \in (\mathcal{B}_r^{\text{even}})_{j:h}$. To see that f is an isomorphism it remains to show that f is edge-preserving. To see this, observe that two vertices $a, b \in V(Lex(j, h)^{\text{even}})$ are adjacent if and only if one is the depth-l ancestor and the other the depth($l + 1$) ancestor of some leaf, and that this holds if and only if $f(a)$, $f(b)$ are also adjacent.

The proof of the second claim is similar. □

We can now describe the structure of G^{even} and G^{odd}.

For each $w \in \mathcal{B}_r^{\text{even}}$, let G^{even} contain the caterpillar $Cat(w)$. For each $i \in [r]$ and $h \in \{0, 1\}$, let G^{even} contain the lexicographic tree $Lex(i, h)^{\text{even}}$. Finally, for each $i \in [r]$ let G^{even} contain the labelled leaf x_i adjacent to $v_{i,0}^{\text{even}}$.

The construction of G^{odd} is similar: For each $w \in \mathcal{B}_r^{\text{odd}}$, let G^{odd} contain the caterpillar $Cat(w)$. For each $i \in [r]$ and $h \in \{0, 1\}$, let G^{odd} contain the lexicographic tree $Lex(i, h)^{\text{odd}}$. Finally, for each $i \in [r]$ let G^{odd} contain the labelled leaf x_i adjacent to $v_{i,0}^{\text{even}}$.

Observe that in both G^{even} and G^{odd}, the vertices $z_{w,i}$ have degree 2 (as they appear as a leaf in the caterpillar $Cat(w)$ and in the lexicographic tree $Lex(i, w_i)^{\text{even}}$ or $Lex(i, w_i)^{\text{odd}}$). The vertices $v_{i,1}$ also have degree 2, and all other non-leaf vertices have degree 3.

We will later show that G^{even} and G^{odd} are not equivalent. First though, we will show that the multigraphs derived from G^{even} and G^{odd} by *deleting* (not removing) the same leaf are in fact equivalent. (Recall that the difference between deleting and removing a vertex v is that removing v involves the extra step of suppressing any degree-2 vertices left after deleting v.)

Lemma 4 *For $i \in [r]$, let $G^{\text{even}} - x_i$ be the graph derived from G^{even} by deleting x_i and its incident edge, and similarly let $G^{\text{odd}} - x_i$ be the graph derived from G^{odd} by deleting x_i and its incident edge. Then $G^{\text{even}} - x_i$ and $G^{\text{odd}} - x_i$ are equivalent.*

Proof We will describe a set of isomorphisms between subgraphs of $G^{\text{even}} - x_i$ and $G^{\text{odd}} - x_i$, then combine them to produce an isomorphism between $G^{\text{even}} - x_i$ and $G^{\text{odd}} - x_i$. Each isomorphism will be one that maps vertex $z_{w,j}$ to $z_{w^{\approx i},j}$.

For each $w \in \mathcal{B}_r^{\text{even}}$, Observation 1 implies that there exists an isomorphism f between $Cat(w)$ and $Cat(w^{\approx i})$ such that $f(z_{w,j}) = z_{w^{\approx i},j}$ for each $j \in [r]$. For each $j \in [r] \setminus \{i\}$ and $h \in \{0, 1\}$, Lemma 3 implies that there exists an isomorphism f between $Lex(j, h)^{\text{even}}$ and $Lex(j, h)^{\text{odd}}$, such that $f(v_{j,h}) = v_{j,h}$ and $f(z_{w,j}, j) = z_{w^{\approx i};j}$ for each leaf $z_{w,j}$. Finally, for each $h \in \{0, 1\}$, Lemma 3 implies that there exists an isomorphism f between $Lex(i, h)^{\text{even}}$ and $Lex(i, 1 - h)^{\text{odd}}$, such that $f(v_{i,h}^{\text{even}}) = v_{i,1-h}^{\text{odd}}$ and $f(z_{w,i}) = z_{w^{\approx i},i}$ for each leaf $z_{w,i}$.

Observe that all of these isomorphisms agree on $z_{w,j}$ for any $w \in \mathcal{B}_r^{\text{even}}$, $j \in [r]$ (that is, they each map this vertex to $z_{w^{\approx i},j}$), and such vertices are the only vertices that are shared between caterpillars and lexicographic trees. Thus we can combine these isomorphisms into a single edge-preserving function f that maps every non-leaf vertex of $G^{\text{even}} - x_i$ to a non-leaf vertex of $G^{\text{odd}} - x_i$. Moreover, as each caterpillar and lexicographic tree in $G^{\text{even}} - x_i$ is mapped to a different caterpillar or lexicographic tree in $G^{\text{odd}} - x_i$, this function is a bijection. Finally, set $f(x_j) = x_j$ for every $j \in [r] \setminus i$.

Then f is now a bijective function from $V(G^{\text{even}} - x_i)$ to $V(G^{\text{odd}} - x_i)$ that is both edge-preserving and label-preserving.   □

We note that we cannot extend the above graph isomorphism between $G^{\text{even}} - x_i$ and $G^{\text{odd}} - x_i$ to an isomorphism between G^{even} and G^{odd} by setting $f(x_i) = x_i$, because $f(v_{i,0}^{\text{even}}) = v_{i,1}^{\text{odd}}$, and so there would be no edge between $x_i = f(x_i)$ and $f(v_{i,0}^{\text{even}}) = v_{i,1}^{\text{odd}}$ in G^{odd}.

In fact, the next lemma shows that there is no isomorphism between G^{even} and G^{odd}.

Lemma 5 *Let $\mathbf{0}$ denote the all-0 sequence from $\mathcal{B}_r$. For two vertices a, b in G^{even}, let $dist^{\text{even}}(a, b)$ denote the distance between a and b in G^{even}. Similarly for two vertices a, b in G^{odd}, let $dist^{\text{odd}}(a, b)$ denote the distance between a and b in G^{odd}. Then for any vertex a in G^{odd}:*

1. *If $dist^{\text{odd}}(a, x_1) = dist^{\text{even}}(u_\mathbf{0}, x_1)$ then $a = u_w$ for some $w \in \mathcal{B}_r^{\text{odd}}$.*
2. *If $a = u_w$ for some $w \in \mathcal{B}_r^{\text{odd}}$ then there exists $i \in [r]$ such that $dist^{\text{odd}}(a, x_i) > dist^{\text{even}}(u_\mathbf{0}, x_i)$.*

This holds even if we suppress all degree-2 vertices in G^{even} and G^{odd}.

Proof We consider the two parts of the claim separately.

1. We first calculate the value of $dist^{\text{even}}(u_\mathbf{0}, x_1)$. Recall that in G^{even}, x_1 is adjacent to the root $v_{1,0}$ of $Lex(1, 0)^{\text{even}}$, and (by definition) every leaf of $Lex(1, 0)^{\text{even}}$ has distance $r - 2$ from $v_{1,0}$. As $u_\mathbf{0}$ is adjacent to a leaf $z_{\mathbf{0},1}$ of $Lex(1, 0)^{\text{even}}$, it follows that $dist^{\text{even}}(u_\mathbf{0}, x_1) = 1 + r - 2 + 1 = r$ (there is no shorter path from $u_\mathbf{0}$ to x_1, as any path must pass through $z_{w,1}$ for some w).

   As all leaves in $Lex(1, 0)^{\text{odd}}$ have distance $r - 2$ from $v_{1,0}$ in G^{odd}, and therefore distance $r - 1 = dist^{\text{even}}(u_\mathbf{0}, x_1) - 1$ from x_1, it follows that the only vertices in G^{odd} of distance $dist^{\text{even}}(u_\mathbf{0}, x_1)$ from x_1 are those which are not in $Lex(1, 0)^{\text{odd}}$ but adjacent to a leaf $z_{w,1}$ of $Lex(1, 0)^{\text{odd}}$. By construction, all such vertices are u_w for some $w \in \mathcal{B}_r^{\text{odd}}$ such that $w_1 = 0$.

   When degree-2 vertices are suppressed, a similar argument holds, except that $dist^{\text{even}}(u_\mathbf{0}, x_1)$ is reduced by 1 (as we suppress $z_{\mathbf{0},1}$). It remains the case that the vertices in G^{odd} of distance $dist^{\text{even}}(u_\mathbf{0}, x_1)$ from x_1 are those which are incident to a vertex from $Lex(1, 0)^{\text{odd}}$ but not in $Lex(1, 0)^{\text{odd}}$ themselves, and again all such vertices are u_w for some $w \in \mathcal{B}_r^{\text{odd}}$.

2. For any $w \in \mathcal{B}_r^{\text{odd}}$, there exists $i \in [r]$ such that $w_i = 1$. Any path from u_w to x_i must pass through a vertex $z_{w',i}$ where $w'_i = 0$, and all such vertices have equal distance from x_i. Thus, it is enough to show that the distance in G^{odd} between u_w and any such $z_{w',i}$ is greater than the distance between $u_\mathbf{0}$ and $z_{\mathbf{0},i}$ in G^{even}.

   To see this, consider a path P between u_w and $z_{w',i}$. As $w'_i = 0$, we note that $w' \neq w$ and so P must traverse at least one lexicographic tree. We construct a mapping $g : V(P) \to V(Cat(\mathbf{0}))$, as follows. For any $a \in V(P)$, if a is in $Cat(w'')$ for any $w'' \in \mathcal{B}_r^{\text{odd}}$ (including w or w'), set $g(a) = f(a)$, where f is the isomorphism between $Cat(w'')$ and $Cat(\mathbf{0})$ such that $f(u_w) = u_\mathbf{0}$ and $f(z_{w'',j}) = z_{\mathbf{0},j}$ for all $j \in [r]$ (such an isomorphism exists by Observation 1).

Otherwise, it must be the case that $a \in Lex(j, h)^{\text{odd}}$ for some $j \in [r]$, $h \in \{0, 1\}$. In this case, set $g(a) = z_{\mathbf{0},j}$. Let Q be the set of all $g(a)$ for any vertex a in P. Observe that for any vertices a, b in P, if a and b are adjacent then either $g(a) = g(b)$ or $g(a)$ and $g(b)$ are adjacent. It follows that Q forms a connected set of vertices in $Cat(\mathbf{0})$, and thus Q contains a path between $g(u_w) = u_{\mathbf{0}}$ and $g(z_{w',i}) = z_{\mathbf{0},i}$. Moreover, as P must traverse at least one lexicographic tree, there are consecutive vertices in P that are mapped to the same vertex by g. It follows that the path in Q is shorter than the path P, as required. It follows that the distance between u_w and x_i is greater than $dist^{\text{even}}(u_{\mathbf{0}}, x_i)$. We note that a similar argument applies even when vertices of degree 2 are suppressed.

$\square$

Corollary 1 G^{even} and G^{odd} are not equivalent.

The next lemma will be used to show that when we suppress degree-2 vertices in G^{even} and G^{odd}, the resulting graphs N^{even} and N^{odd} are networks.

Lemma 6 In both G^{even} and G^{odd}, there exists a single blob containing all non-leaf vertices.

Proof Observe that any non-leaf vertex is part of a path between u_w and $v_{i,h}$ for some $w \in \mathcal{B}_r$, $i \in [r]$, $h \in \{0, 1\}$. Furthermore every vertex $v_{i,h}$ appears on a path between u_w and $u_{w'}$ for some w, w'. Therefore it is enough to show that for any $w \neq w'$, u_w and u'_w appear in the same blob.

Let $00*, 01*, 11*, 10*$ be four sequences in $\mathcal{B}_r^{\text{even}}$ such that $hk*_1 = h, hk*_2 = k$ (such sequences exist as $r > 3$).

Then there exists a cycle

$$u_{00*}z_{00*,1} \cdots z_{01*,1}u_{01*}z_{01*,2} \cdots z_{11*,2}u_{11*}z_{11*,1} \cdots z_{10*,1}u_{10*}z_{10*,2} \cdots z_{00*,2}u_{00*}.$$

Here the path between $z_{00*,1}$ and $z_{01*,1}$ passes through $Lex(1, 0)^{\text{even}}$, the path between $z_{01*,2}$ and $z_{11*,2}$ passes through $Lex(2, 1)^{\text{even}}$, the path between $z_{11*,1}$ and $z_{10*,1}$ passes through $Lex(1, 1)^{\text{even}}$, and the path between $z_{10*,2}$ and $z_{00*,2}$ passes through $Lex(2, 0)^{\text{even}}$. See Fig. 6 for an example when $00* = 0000$, $01* = 0101$, $11* = 1100$ and $10* = 1001$.

As $00*, 01*, 11*, 10*$ appear on a cycle, they appear in the same blob of G^{even}. Moreover as any vertex u_w could fill the role of one of $00*, 01*, 11*, 10*$, we have that all u_w appear in the same blob. A similar argument holds for G^{odd}. $\square$

Now we are ready to construct the networks N^{even} and N^{odd}: Let N^{even} be derived from G^{even} by suppressing all vertices of degree 2. Similarly, let N^{odd} be derived from G^{odd} by suppressing all vertices of degree 2 (see Fig. 7 for the networks when $r = 4$).

Lemma 7 N^{even} and N^{odd} are networks on X.

Proof We show that N^{even} is a network on X (the proof for N^{odd} is similar). By construction, all vertices in N^{even} have degree 1 or 3 and the leaves are bijectively
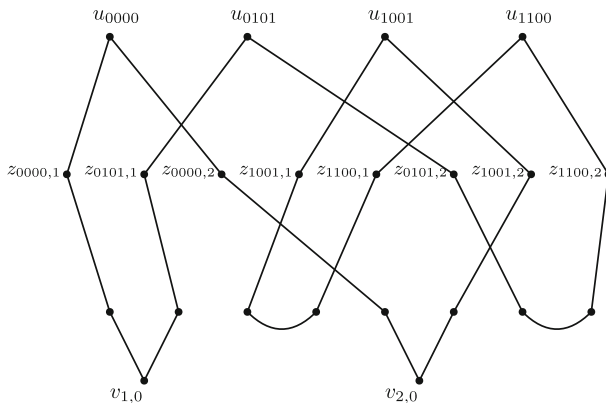
Fig. 6 A cycle containing the vertices $u_{0000}, u_{0101}, u_{1001}, u_{1100}$ in G^{even} for the case $r = 4$

labelled with the elements of X. It remains to show that contracting each blob into a single vertex gives a tree with no degree-2 vertices, which we will do by showing that N^{even} has only one blob. By Lemma 6, all non-leaf vertices in G^{even} are part of the same blob in G^{even}. Observe that if two degree-3 vertices are in the same blob, then they are still in the same blob after contracting degree-2 vertices. Thus, all non-leaf vertices in N^{even} are part of the same blob, and thus N^{even} has a single blob, as required. □

Lemma 8 N^{even} *and* N^{odd} *are not equivalent.*

Proof As N^{even} and N^{odd} are derived from G^{even} and G^{odd} by suppressing degree-2 vertices, Lemma 5 implies that there is no vertex in N^{odd} that has the same distance from each leaf x_i as u_0 has from x_i in N^{even}.

This implies that there is no isomorphism between N^{even} and N^{odd}, as if f is edge-preserving and label-preserving then the distance between u_0 and x_i is equal to the distance between $f(u_0)$ and $f(x_i) = x_i$. □

Lemma 9 *For each* $i \in [r]$, $(N^{\text{even}})_{x_i}$ *and* $(N^{\text{odd}})_{x_i}$ *are equivalent.*

Proof Recall the definitions of $G^{\text{even}} - x_i$ and $G^{\text{odd}} - x_i$, and observe that $(N^{\text{even}})_{x_i}$ (respectively, $(N^{\text{odd}})_{x_i}$) can be derived from $G^{\text{even}} - x_i$ ($G^{\text{odd}} - x_i$) by suppressing degree-2 vertices. By Lemma 4, there exists an isomorphism f' between $G^{\text{even}} - x_i$ and $G^{\text{odd}} - x_i$. So define a bijective function $f : V((N^{\text{even}})_{x_i}) \to V((N^{\text{odd}})_{x_i})$ by setting $f(a) = f'(a)$ for all $a \in V((N^{\text{even}})_{x_i})$. Note that if a does not have degree 2 in $G^{\text{even}} - x_i$, $f'(a)$ also does not have degree 2 in $G^{\text{odd}} - x_i$. Thus if $a \in V((N^{\text{even}})_{x_i})$ then $f(a) = f'(a) \in V((N^{\text{odd}})_{x_i})$, and so f is well-defined.

By construction, f is label-preserving. To see that f is edge-preserving, consider some $a, b \in V((N^{\text{even}})_{x_i})$. Observe that the number of edges between a and b in $(N^{\text{even}})_{x_i}$ is equal to the number of paths between a and b in $G^{\text{even}} - x_i$ whose internal vertices have degree 2. As f' is an isomorphism, this is equal to the number of paths between $f'(a)$ and $f'(b)$ in $G^{\text{odd}} - x_i$ whose internal vertices have degree 2, which in turn is equal to the number of edges between $f(a)$ and $f(b)$ in $(N^{\text{odd}})_{x_i}$. Thus, f is edge-preserving, and so f is an isomorphism. □
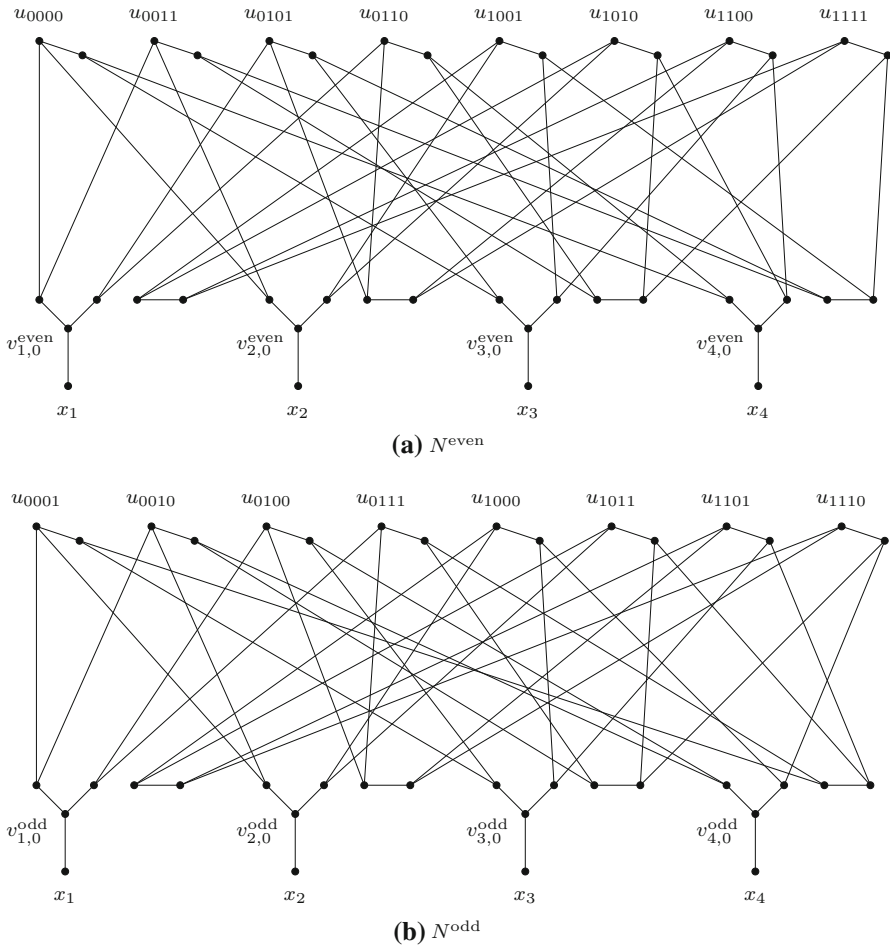
Lemmas 7, 8 and 9 imply the following theorem:

(a) N^{even}

(b) N^{odd}

Fig. 7 Binary example of N^{even} and N^{odd} for the case when when $r = 4$. The vertex u_{0000} in N^{even} has distance $d_1 = 3$ from x_1, $d_2 = 3$ from x_2, $d_3 = 4$ from x_3, and $d_4 = 4$ from x_4. Moreover there is no vertex in N^{odd} with distance d_i from x_i for each $i \in [4]$. Thus N^{even} and N^{odd} are not equivalent. However, for each $i \in [4]$ the multigraphs $(N^{\text{even}})_{x_i}$ and $(N^{\text{odd}})_{x_i}$ are equivalent, using an isomorphism that maps each vertex u_w to $u_{w \nsim i}$

Theorem 2 *For any $r \geq 4$, there exist networks N^{even}, N^{odd} on X with $|X| = r$, such that N^{odd} is a leaf-reconstruction of N^{even}, but N^{even} and N^{odd} are not equivalent. Thus, N^{even} is not leaf-reconstructible.*

5 Concluding remarks

Although we have shown that not all phylogenetic networks with five or more leaves are leaf-reconstructible, this does not mean that reconstructing networks from subnet-

works is completely hopeless. There are already some positive results for interesting restricted network classes (Van Iersel and Moulton 2018). Moreover, since the presented counter-examples are very complex, it is certainly possible that other reasonable network classes are also leaf-reconstructible.

For example, while it is known that all networks with at least five leaves and $|E| - |V| \leq 3$ are leaf-reconstructible, the counter-examples presented in this paper have $|E| - |V| = 2^{r-1}(r-1) - 2r$, with r the number of leaves. Hence, whether networks with $3 < |E| - |V| < 2^{r-1}(r-1) - 2r$ are leaf-reconstructible is still open. In particular, is it possible to construct counter-examples where $|E| - |V|$ is bounded by a linear function of the number of leaves?

References

Bapteste E, van Iersel L, Janke A, Kelchner S, Kelk S, McInerney JO, Morrison DA, Nakhleh L, Steel M, Stougie L et al (2013) Networks: expanding evolutionary thinking. Trends Genet 29(8):439–441

Bondy JA, Hemminger RL (1977) Graph reconstructiona survey. J Graph Theory 1(3):227–268

Huber KT, van Iersel L, Moulton V, Wu T (2014) How much information is needed to infer reticulate evolutionary histories? Syst Biol 64(1):102–111

Pardi F, Scornavacca C (2015) Reconstructible phylogenetic networks: do not distinguish the indistinguishable. PLOS Comput Biol 11(4):e1004135

Thatte BD (2008) Combinatorics of pedigrees I: counterexamples to a reconstruction question. SIAM J Discrete Math 22(3):961–970

van Iersel L, Moulton V (2018) Leaf-reconstructibility of phylogenetic networks. SIAM J Discrete Math 32(3):2047–2066

van Iersel L, Kelk S, Scornavacca C (2016) Kernelizations for the hybridization number problem on multiple nonbinary trees. J Comput Syst Sci 82(6):1075–1089

Whidden C, Beiko RG, Zeh N (2013) Fixed-parameter algorithms for maximum agreement forests. SIAM J Comput 42(4):1431–1466