Is Level-1 Blob Reconstruction Under the Network Multispecies Coalescent Easy?

Junyan Dai
Department of Computer Science, University of Maryland, College Park, MD, USA

Erin K. Molloy[1]
Department of Computer Science, University of Maryland, College Park, MD, USA
University of Maryland Institute for Advanced Computer Studies, College Park, MD, USA

Abstract

Hybridization is an important evolutionary process, commonly modeled by the network multispecies coalescent. Reconstructing evolutionary histories under this model is notoriously costly, even for level-1 networks where hybridization events are isolated from each other. The widely used methods that combine speed with statistical guarantees rely on quartet concordance factors computed for all subsets of four species, resulting in an $O(n^4k)$ bottleneck that severely limits scalability to large numbers of species (n) and genes (k). Among quartet-based methods, NANUQ+ is notable because it decomposes the problem into two steps: first reconstructing a *tree of blobs*, which compresses each non-treelike part of the network, called a blob, into a single vertex, and second reconstructing the internal structure of each level-1 blob, specifically its circular order and hybrid vertex. Here, we investigate whether level-1 blob reconstruction is difficult once the tree of blobs is known. We present a fast and statistically consistent algorithm, called NetCS, based on two simple primitives: majority voting and merge sort, circumventing the bottleneck of computing all quartet concordance factors. In simulations, NetCS achieved comparable accuracy to NANUQ+ and was dramatically faster, enabling analyses of 200 taxa and 1000 genes in only a few minutes. Both methods attained near-perfect accuracy when given the true tree of blobs; however, their performance degraded in end-to-end pipelines due to errors in tree of blobs reconstruction. Strikingly, even methods that reconstruct level-1 networks directly struggled to accurately predict hybrid ancestry. Our results suggest that reconstructing level-1 blobs is unexpectedly easy once the tree of blobs is known, and that a major challenge for phylogenetic network inference lies in accurate tree of blobs reconstruction.

2012 ACM Subject Classification Applied computing → Molecular evolution

Keywords and phrases Phylogenetic Networks, Circular Orders, Quartets, Network Multispecies Coalescent

Digital Object Identifier 10.4230/LIPIcs.WABI.2026.1

Supplementary Material *Software (NetCS)*: `https://github.com/molloy-lab/TREE-QMC` archived at `swh:1:dir:4823b273f9bf1f647cc3e128f7ac10e051ee98d4`
Software (Analysis Scripts and Data Files): `https://github.com/junyandai/netcs-study` archived at `swh:1:dir:e2768f0fae2e6947629dacd14039401a098d4611`

Funding This material is based upon work supported by the U. S. National Science Foundation (NSF) under Grant No. 2441458 (to EKM) as well as the State of Maryland. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the sponsors.

1 Introduction

Hybridization and admixture are important evolutionary scenarios that must be represented as phylogenetic networks rather than trees. A phylogenetic network governs the ancestry of unlinked, recombination-free genomic regions, called genes, commonly modeled by the

[1] Corresponding author

© Junyan Dai and Erin K. Molloy;
licensed under Creative Commons License CC-BY 4.0

Network Multispecies Coalescent (NMSC) [40, 41]. Reticulations in the network enable genes from the same species' genome to trace back to different ancestral populations, resulting in conflicting genealogies, called gene trees. Additionally, conflicts among gene trees can arise from incomplete lineage sorting (ILS), similar to the traditional MSC model [29]. Although gene trees are not directly observed, they govern the evolution of DNA sequences, from which they can be estimated with fast methods, such as RAxML [36] and IQ-TREE [26].

Reconstructing phylogenetic networks under the NMSC is notoriously costly even for level-1 networks where hybridization events are isolated from each other. The popular methods that combine speed with statistical guarantees rely on *quartet concordance factors (qCFs)* computed for all subsets of four species, called taxa. A quartet is an unrooted phylogenetic tree on four species, and its concordance factor is simply the number of gene trees that have the same topology after restriction to the same species. Performing this precomputation for all four-taxon subsets results in an $O(n^4k)$ cost that severely limits scalability to large numbers of species (n) and genes (k), although methods based on qCFs are still significantly cheaper than those that based on full likelihood calculations (e.g., [41]).

The first network reconstruction method to leverage qCFs was SNaQ [34, 35], which conducts a search over the space of (semi-directed) level-1 networks via edit moves, seeking a network to maximize its pseudo-likelihood function (see [24] for a related composite likelihood function for DNA sequences). To improve scalability beyond $\sim$30 species, the most recent version of SNaQ attempts to approximate pseudo-likelihood by randomly sampling four-taxon subsets [22]; this approach currently has no statistical guarantees, and there are intricacies to random sampling, as highlighted by [10]. Alternatively, SNaQ can be leveraged within the divide-and-conquer framework InPhyNet [23], but then the goal is to run SNaQ on pairwise disjoint subsets of taxa that are as large as possible.

Another recently developed method that leverages qCFs is CAMUS [38], accepted to ISMB 2026. Unlike SNaQ, CAMUS takes a rooted base tree as input (typically estimated via ASTRAL-IV [43]) and then adds reticulations edges to maximize the number of quartets also displayed by the (directed) level-1 network (also see [28]). Although CAMUS is more scalable than SNaQ, it still requires qCF precomputation. CAMUS does not currently have statistical guarantees under the NMSC and can be statistically inconsistent when ASTRAL-IV is used to construct the base tree, as recently shown by Dinh and Baños [11] (similar issues for tree-based network inference arise in the context of admixture graph reconstruction [27]).

In contrast, NANUQ+ [4] decomposes (semi-directed) level-1 network reconstruction into two steps: (1) reconstructing a tree of blobs, which compresses each non-treelike part of the network, called a blob, into a single vertex and then (2) reconstructing the internal structure of each level-1 blob, specifically its circular order and hybrid vertex (also see [16]). Each step as implemented by NANUQ+ is provably *statistically consistent* under the NMSC with some assumptions [3, 4]; however, both steps require qCF precomputation, limiting scalability.

Recent work has focused on scalable tree of blobs reconstruction from qCFs with statistical and/or correctness guarantees (e.g., [3, 10, 17, 13, 30]). This motivates us to investigate whether level-1 blob reconstruction is difficult once the tree of blobs is known. The remainder of the paper is organized as follows. We present the terminology and background in Section 2. In Section 3, we present our new algorithm for level-1 blob reconstruction from qCFs, called NetCS, along with its theoretical properties. NetCS is based on two simple primitives: majority voting and sorting; thus it differs dramatically from NANUQ+'s approach, which computes distances between all pairs of taxa based on the input qCFs and then evaluates the goodness of fit between expected and observed distances for all possible level-1 blob structures.

In Section 4, we describe the results of benchmarking NetCS against NANUQ+ and CAMUS. We conclude in Section 5 with a discussion of whether level-1 blob reconstruction is difficult based on our theoretical and empirical results.

2 Background

To present Net-CS, we review terminology and background for phylogenetic networks.

Phylogenetic Networks. A *directed* phylogenetic network N^+ is a directed simple acyclic finite graph with the property that there exists a directed path from the root, a unique vertex with in-degree 0, to all other vertices. The leaf vertices of N^+, which have out-degree 0, are bijectively labeled by a set S of species. A *hybrid* is a vertex with in-degree > 1, and a *reticulation edge* is an arc whose child is a hybrid. A *wedge* path in N^+ is a sequence of distinct vertices $u_0, u_1, \ldots, u_{k-1}, u_k$ such that there exists some vertex u_i that "roots" the path so $u_i, u_{i-1}, \ldots, u_2, u_1$ and $u_i, u_{i+1}, \ldots, u_{k-1}, u_k$ are both directed paths in N^+ [17]. A directed network N^+ can be converted into a *semi-directed* network N^- [34] by removing all edges and vertices above the lowest stable ancestor s of S (i.e., the lowest vertex on all directed paths from the root to each leaf in S; see [37], page 263), undirecting all non-reticulation edges, and suppressing s if it has degree 2. The wedge path terminology is used for semi-directed networks because there is a bijection between wedge paths in N^+ and paths in N^- that do not traverse two reticulation edges incident to the same hybrid vertex. If there are no hybrid vertices, N^- is simply an unrooted phylogenetic tree. We assume that semi-directed networks and unrooted trees are *binary*, specifically leaves have degree 1 and all other vertices have degree 3.

Cut Edges, Blobs, and Partitions. A bridge, also called a *cut edge*, is an edge whose deletion increases the number of connected components of the graph. The maximal bridgeless subgraphs of N^-, when ignoring the directionality of reticulation edges, are called *blobs*. The *level* of a blob is the number of reticulation edges in it minus the number of hybrid vertices in it; we assume that all blobs are level-1. The degree of a blob is the number of cut edges adjacent to it. An m-blob is a blob with degree m. The *tree of blobs (TOB)* [15] is obtained from N^- by contracting the edges in each blob into a single vertex and then suppressing any degree-2 vertices (Fig. 1A–B). Let $v_{\mathcal{B}}$ denote the vertex in the tree of blobs T_N that corresponds to blob $\mathcal{B}$ in the underlying network N^-. The deletion of vertex $v_{\mathcal{B}}$ from T_N produces a collection of rooted subgraphs whose leaf sets partition the taxon set S, denoted $Part(v_{\mathcal{B}}) = P_1|P_2|\ldots|P_{m_{\mathcal{B}}}$, where $m_{\mathcal{B}}$ is the degree of $v_{\mathcal{B}}$; this is equivalent to the partition induced by deleting all edges and vertices in blob $\mathcal{B}$ from N^-, denoted $Part(\mathcal{B})$. Likewise, the deletion of a cut edge e from N^- or its TOB yields two rooted subgraphs whose leaves partition S into two subsets, called a *bipartition* or *split*, denoted $A|B$. A bipartition is trivial if either $|A| = 1$ or $|B| = 1$. If both endpoints of e have degree 3, then deleting e, along with its endpoints, induces a *quadrapartition*: $Quad(e) = A_1|A_2|B_1|B_2$.

Blob structure. Each blob $\mathcal{B}$ in a binary, semi-directed, level-1 network is either a single vertex or a cycle when ignoring the directionality of reticulation edges; the former is called trivial and the latter non-trivial. Moreover, no two blobs can share an edge or vertex, and each non-trivial blob contains exactly one hybrid vertex incident to exactly two reticulation edges. The **circular order** of $\mathcal{B}$ is the cyclic ordering of vertices encountered when traversing the cycle ignoring edge direction, considered up to rotation and reversal. There is a bijection

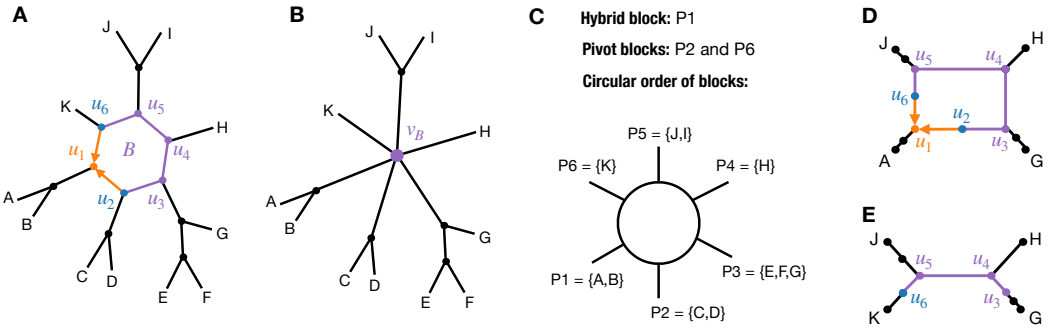

Figure 1 **A)** A binary, semi-directed, level-1 network N^- on species set S with one non-trivial blob $\mathcal{B}$. **B)** Tree of blobs for N^-. Blob vertex $v_{\mathcal{B}}$ induces an 6-partition of S denoted $Part(v_{\mathcal{B}}) = P_1|\ldots|P_6$. **C)** Information about $\mathcal{B}$. **D)** Subnetwork on 4 taxa from distinct blocks of $Part(v_{\mathcal{B}})$, including hybrid block P_1. The subnetwork displays two quartets: $A, J|H, G$ and $A, G|H, J$. **E)** Subnetwork on 4 taxa from distinct blocks of $Part(v_{\mathcal{B}})$, excluding hybrid block P_1.

between the vertices in the circular order of blob $\mathcal{B}$ and the blocks of $Part(\mathcal{B})$; thus, we refer to them interchangeably. N^- can be reconstructed from its TOB by (1) determining the circular order of each non-trivial blob, and (2) identifying the hybrid of each non-trivial blob, as this gives the directionality of reticulation edges (Fig. 1C) [4].

Quarnets and quartets. A *subnetwork* of N^- on $Y \subseteq S$, denoted $N^-|_Y$, is the subgraph induced by the union of vertices and edges on all wedge paths in N^- between all pairs of taxa in Y [5]. The *restriction* of N^- to $Y \subseteq S$ is formed by computing the subnetwork on Y and then repeatedly suppressing degree-2 vertices, parallel arcs, and 2-blobs [17]. For $|Y| = 4$, subnetworks are referred to as **quarnets**. **Quarnet-splits** are the set of bipartitions induced by the cut edges of $N^-|_Y$ or equivalently the edges in the TOB of $N^-|_Y$ [17]. A quarnet either contains a 4-blob or not [3, 17]. In the former case, the TOB of $N^-|_Y$ is an (unresolved) star tree (Fig. 1D), and in the latter case, the TOB of $N^-|_Y$ is a (binary) **quartet** tree (Fig. 1E); thus, we say that the evolutionary relationship is "treelike" [3]. A quartet on species set $\{a, b, c, d\}$ has one of three possible topologies, indicated by the one non-trivial bipartition: $a, b|c, d$ or $a, c|b, d$, or $a, d|b, c$. Any quarnet that contains a 4-blob **displays** exactly two quartets formed by deleting each of the reticulation edges incident to the hybrid vertex given our assumption that N^- is binary and level-1 [34, 7].

Model and Data. The Network Multispecies Coalescent (NMSC) model is parameterized by $\mathcal{N}^+ = (N^+, S, \Theta)$, where N^+ is a directed phylogenetic network on species set S and Θ is a set of numerical parameters: branch lengths in coalescent units, reticulation edge inheritance proportions, and the inheritance correlation ρ (note that we assume independent inheritance, i.e., $\rho = 0$). We only briefly summarize the NMSC here and refer readers to [34,10] for a complete description. A gene genealogy is generated under the NMSC by sampling one gene per species and tracing ancestry backward in time until all lineages coalesce (i.e., trace back to a common ancestor), forming a rooted, binary tree on S [40, 12]. We let $\mathbb{P}_{\mathcal{N}^+}(t)$ denote the probability of an unrooted, binary gene tree t on species set $Y \subseteq S$ given $\mathcal{N}^+$. For $|Y| = 4$, $\mathbb{P}_{\mathcal{N}^+}(q_i^Y)$, where $i = \{1, 2, 3\}$ indexes the three possible quartets on Y, are the *expected* **quartet concordance factors (qCFs)**. Expected qCFs are fully determined by the semi-directed version of $\mathcal{N}^+$, denoted $\mathcal{N}^-$ (Proposition 3 in [34]). *Observed* qCFs are

computed by counting the number of k *gene trees* whose restriction to Y is isomorphic to q_i^Y, denoted $w_k(q_i^Y)$, and then dividing each element in the 3-vector by its sum. Observed qCFs converge simultaneously to their expected values with probability 1 as the number of gene trees goes to infinity by the strong law of large numbers.

Assumptions and Identifiability. We make standard assumptions about the model network; see [34, 7, 3, 4, 10]. Specifically, we assume that $\mathcal{N}^-$ is binary, semi-directed, level-1, and metric (i.e., non-reticulation edges have positive lengths, reticulation edges have lengths greater than or equal to zero, reticulation edges incident to the same target have inheritance proportions that sum to 1 and take on values between 0 and 1, non-inclusive). Under these assumptions, the structure of 3-blobs is not identifiable from qCFs [7]. However, the circular order of 4-blobs is identifiable, and for large blobs with degree > 4, the hybrid vertex also is identifiable (Theorem 4 in [7]). The crux behind this identifiability result is that for any 4-taxon subset Y such that $N^-|_Y$ contains a 4-blob, the qCFs on Y are distinct, i.e., $\mathbb{P}_{\mathcal{N}^-}(q_1^Y) \neq \mathbb{P}_{\mathcal{N}^-}(q_2^Y) \neq \mathbb{P}_{\mathcal{N}^-}(q_3^Y)$ (Def. 17 and Theorem 3 in [7]). Moreover, the lowest CF corresponds to the only quartet that is not displayed by $N^-|_Y$ (Proposition 11 in [7]). These properties hold under generic numerical parameter settings except those in a set of measure zero (an example in the set of measure zero can be obtained by setting all edges in Fig. 1D to have the same length in coalescent units and both inheritance proportions to 0.5).

Oracle for 4-blobs. Lastly, our approach requires an oracle that takes a 4-taxon subset on Y as input, specifically its associated 3-vector of qCFs, and returns whether the quarnet $N^-|_Y$ contains a 4-blob or not. In practice, we employ the same hypothesis testing framework used for tree of blobs reconstruction in NANUQ+ [4, 3, 2]. There are two versions of the hypothesis test: T3 and cut. The null hypothesis of both is that the $N^-|_Y$ does not contain a 4-blob, thus low p-values indicate the null should be rejected (i.e., $N^-|_Y$ does contain a 4-blob). The test assumes that expected qCFs are distinct for every 4-taxon subset $Y \subseteq S$ such that $N^-|_Y$ contains a 4-blob [3], which is satisfied as noted in the previous section. However, the T3 version of the hypothesis test additionally assumes that networks are *class-1 quartet-nonanomalous*, that is, for every 4-taxon subset $Y \subseteq S$ such that $N^-|_Y$ does not contain a 4-blob, $\mathbb{P}_{\mathcal{N}^-}(q_1^Y) > \mathbb{P}_{\mathcal{N}^-}(q_2^Y)$ and $\mathbb{P}_{\mathcal{N}^-}(q_1^Y) > \mathbb{P}_{\mathcal{N}^-}(q_3^Y)$, where q_1^Y corresponds to the TOB of $N^-|_Y$ [3]. This assumption is satisfied for common inheritance ($\rho = 1$) but not necessarily independent inheritance ($\rho = 0$) [5], although strong empirical performance of the T3 option is still possible even under model violations [3, 10].

3 Methods

We are now ready to present our method NetCS, which has the following input and output.

- **Input:**
  - Unrooted tree of blobs T_N on a set S of n species
  - Set $\{G_1, \ldots, G_k\}$ of unrooted, binary gene trees, each on species set S_i, where $S = \cup_{i=1}^k S_i$
- **Output:** A binary, semi-directed, level-1 network on S with TOB isomorphic to T_N (note that only non-trivial blobs of degree at least 5 are reconstructed)

NetCS processes each vertex with degree 5 or more in T_N *independently* with two primitives: majority voting to detect the hybrid and sorting to resolve the circular order.

3.1 Hybrid Majority Voting

Consider some vertex $v_{\mathcal{B}}$ in the input TOB T_N with degree $m_{\mathcal{B}} \geq 5$. The first step in NetCS is to find the hybrid block of $Part(v_{\mathcal{B}})$, which is the unique block induced by the cut edge incident to the reticulation vertex. Our approach is based on the following result.

▶ **Lemma 1.** *Let $\mathcal{N}^- = (N^-, S, \Theta)$ be a binary, semi-directed, level-1, metric species network. Let $\mathcal{O}_k(Y)$ denote the p-value of NANUQ+ hypothesis test applied to some 4-taxon subset $Y \subseteq S$, specifically the qCFs computed from k independent (unrooted) gene trees on species set S sampled from the distribution induced by the NMSC given $\mathcal{N}^-$. Let $v_{\mathcal{B}}$ be a vertex in the TOB for N^- with degree $m_{\mathcal{B}} \geq 5$. Now consider two 4-taxon subsets constructed such that every taxon is drawn from distinct blocks of $Part(v_{\mathcal{B}})$, with the first subset Y_{+h} drawing one taxon from the hybrid block of $Part(v_{\mathcal{B}})$ and the second subset Y_{-h} drawing zero taxa from the hybrid block. Then, with probability $\to 1$ as $k \to \infty$, $\mathcal{O}_k(Y_{+h}) < \mathcal{O}_k(Y_{-h})$ for all pairs of 4-taxon subsets Y_{+h}, Y_{-h} constructed as described above.*

Proof. For every pair of 4-taxon subsets Y_{+h}, Y_{-h} constructed as described above, $N|_{Y_{+h}}$ contains a 4-blob but $N|_{Y_{-h}}$ does not, by Lemma 5 in [10]. The remainder of the proof follows from the statistical consistency of the hypothesis test $\mathcal{O}_k$; specifically, Theorem 4 in [3] states that there exists a sequence of significance thresholds $\alpha_k \to 0$ such that the hypothesis test applied to k gene trees will, with probability tending to 1 as $k \to \infty$, return the correct result for all 4-taxon subsets of S (also see [30]). This implies the returned p-value will be less than α_k for any 4-taxon subset Y such that $N|_Y$ contains a 4-blob and will be greater than or equal to α_k otherwise, giving us our result. ◀

The innovation behind NetCS is applying the hypothesis test $\mathcal{O}_k$ to carefully selected 4-taxon subsets so that the precomputation of qCFs for all 4-taxon subsets of S can be avoided. Specifically, we create *re-quadrapartitions* of $Part(v_{\mathcal{B}})$ and then apply $\mathcal{O}_k$ to four taxa drawn from distinct blocks.

▶ **Definition 1** (Re-quadraparition). *Let $\mathcal{P} = P_1|P_2|\dots|P_m$ be a m-partition with $m \geq 5$. A* re-quadrapartition *is a coarsening of $\mathcal{P}$ into exactly 4 blocks, denoted $P'_1|P'_2|P'_3|P'_4$, such that each original block P_i is contained entirely within exactly one new block P'_j.*

Given a re-quadrapatition of $Part(v_{\mathcal{B}})$, denoted $P'_1|P'_2|P'_3|P'_4$, the hypothesis test $\mathcal{O}_k$ is applied within the **3-fix, 1-alter (3f1a) algorithm**, introduced at RECOMB 2026 [10]:

1. Select one taxon arbitrarily from each block of the requadrapartition: $x_1 \in P'_1, x_2 \in P'_2, x_3 \in P'_3, x_4 \in P'_4$.
2. Apply $\mathcal{O}_k$ to the qCFs for $\{x_a, x_2, x_3, x_4\}$ for all $x_a \in P'_1$.
3. Repeat step (2) three times but instead of fixing the three taxa from P'_1, P'_2, P'_3 and rotating through all taxa in P'_1, swap out P'_1 with P'_2, P'_3, P'_4.
4. Return the minimum p-value found from all applications of $\mathcal{O}_k$, along with the associated minimizer 4-taxon subset.

Unlike its original application, we leverage 3f1a for voting, allowing each taxon in the minimizer subset vote for its respective block of $Part(v_{\mathcal{B}})$. This process is repeated for c rounds, each corresponding to an arbitrary re-quadrapartition of $Part(v_{\mathcal{B}})$. The block of $Part(v_{\mathcal{B}})$ with the most votes is selected as the hybrid, with some additional steps for tie breaking (Algorithm 1); the tie breaking step is closely related to hybrid identifiability [7, 13].

Algorithm 1 Hybrid Majority Voting Algorithm.

Input : Partition of species set S induced by vertex $v_{\mathcal{B}}$, denoted $Part(v_{\mathcal{B}}) = P_1|P_2|\ldots|P_{m_{\mathcal{B}}}$, Set of k unrooted, binary gene trees each on (a subset of) S, Hypothesis test $\mathcal{O}_k$ that takes qCFs as input and evaluates if the associated quarnet has a 4-blob or not (e.g., the cut test from [3]), Quadrapartition minimizer search algorithm $\mathcal{A}$ (e.g., 3f1a from [10])

Output : Hybrid block of $Part(v_{\mathcal{B}})$

if *fast mode* **then**
  $c \leftarrow 1$;
  Select one taxon arbitrarily from P_i to use for all qCF calculations for all $i \in \{1, 2, \ldots, m_{\mathcal{B}}\}$
else
  $c \leftarrow m_{\mathcal{B}} - 3$
Build a function $F(x)$ that returns i if taxon $x \in P_i$ for all $x \in \cup_{i=1}^{m_{\mathcal{B}}} P_i$;
Initialize an $m_{\mathcal{B}}$-vector, called *votes*, to zeros;
for c *voting rounds* **do**
  Generate re-quadrapartition $P'_1|P'_2|P'_3|P'_4$ by randomly shuffling the blocks of $Part(v_{\mathcal{B}}) = P_1|P_2|\ldots|P_{m_{\mathcal{B}}}$ and then assigning P_i to P'_j if i mod 4 equals j for all $i \in \{1, 2, \ldots, m_{\mathcal{B}}\}$;
  $(p, \{x_1, x_2, x_3, x_4\}) \leftarrow \mathcal{A}(\{P'_1, P'_2, P'_3, P'_4\}, \mathcal{O}_k)$;
  Add one to $votes[F(x_1)]$, $votes[F(x_2)]$, $votes[F(x_3)]$, $votes[F(x_4)]$;
Let P_a, P_b, P_c, P_d, P_e be the five blocks with the most votes, sorted from highest to lowest, breaking ties arbitrarily;
if *votes[a] > votes[b]* **then**
  return P_a;
else
  Apply $\mathcal{O}_k$ to all ways of choosing four blocks from P_a, P_b, P_c, P_d, P_e and store the 4-block subset with highest p-value, denoted P_s, P_t, P_y, P_z;
  return $\{P_a, P_b, P_c, P_d, P_e\} \setminus \{P_s, P_t, P_y, P_z\}$;

▶ **Theorem 1.** *Assume the same settings as Lemma 1, and let $v_{\mathcal{B}}$ be some vertex in the TOB for N^- with degree $m_{\mathcal{B}} \geq 5$. Then, Algorithm 1 with $c \geq 1$ voting rounds applied to $Part(v_{\mathcal{B}})$ and k gene trees will, with probability $\to 1$ as $k \to \infty$, return the (correct) hybrid block of $Part(v_{\mathcal{B}})$. The time complexity of applying Algorithm 1 to all $v_{\mathcal{B}}$ is $O(n \log(n)k)$ in fast mode or $O(n^2 k)$ otherwise, where n is the number of taxa.*

Proof. Vertex $v_{\mathcal{B}}$ is associated with a non-trivial blob $\mathcal{B}$ in N^- given our assumption that N^- is level-1 and binary, so the 3f1a algorithm applied to an arbitrary re-quadrapartition of $Part(v_{\mathcal{B}})$ is guaranteed to apply $\mathcal{O}_k$ to at least one 4-taxon subset Y such that $N^-|_Y$ contains a 4-blob, implying that Y contains a taxon from the hybrid block, by Theorem 2 in [10]. Then, by Lemma 1, the minimizer subset returned by 3f1a given an arbitrary re-quadrapartition of $Part(v_{\mathcal{B}})$ and k gene trees will, with probability going to 1 as $k \to \infty$, include the hybrid block of $Part(v_{\mathcal{B}})$. Repeating for c arbitrary re-quadrapartitions of $Part(v_{\mathcal{B}})$ will produce one of two outcomes: (1) the hybrid block is the only block that earns c votes or (2) the hybrid block is tied for c votes with one or more other blocks. Taking the five blocks with the most votes, breaking ties arbitrarily, ensures that exactly one way of selecting 4 out of these

5 blocks does not include the hybrid block. The p-value from applying the oracle $\mathcal{O}_k$ to this special 4-block subset will be higher than the other 4-block subsets by Lemma 1, revealing the hybrid block by exclusion. The argument above extends to *fast mode*, in which a single taxon is selected arbitrarily to represent each block of $Part(v_{\mathcal{B}})$ (Algorithm 1, line 3), as any 4-taxon subset that covers the same distinct blocks of $Part(v_{\mathcal{B}})$ yields the same expected qCFs under the NMSC given $\mathcal{N}^-$ [7] (also see [10]). It is worth noting that a 4-block tie occurs in fast mode because of setting $c = 1$. However, our implementation of NetCS uses $c = m_{\mathcal{B}} - 3$, as this does not increase the total time complexity of NetCS (Corollary 1).

Now we consider time complexity. Building the taxon-to-block mapping function F is $O(m_{\mathcal{B}})$ in fast mode or $O(n)$ otherwise. Applying 3f1a to c re-quadrapartitions requires $O(cm_{\mathcal{B}})$ or $O(cn)$ queries to the oracle $\mathcal{O}_k$, depending on whether fast mode is employed or not. The oracle is a hypothesis test that can be performed in constant time but requires qCFs as input. After an $O(n \log(n)k)$ preprocessing step, the well-known binary lifting algorithm can be used to retrieve the lowest common ancestor of any pair of leaves in an (arbitrarily) rooted binary gene tree in constant time [14], enabling qCF computation for any 4-taxon subset on the fly in $O(k)$ time. Therefore, when including the calculation of qCFs, each oracle query is $O(cm_{\mathcal{B}}k)$ in fast mode and $O(cnk)$ otherwise. The other calculations do not exceed these time costs. Repeating for all $v_{\mathcal{B}}$ gives $O(n \log(n)k + \sum_{v_{\mathcal{B}}} cm_{\mathcal{B}}k)$ in fast mode and $O(n \log(n)k + \sum_{v_{\mathcal{B}}} cnk)$ otherwise. Taking $c = 1$ in fast mode and $c = O(m_{\mathcal{B}})$ otherwise gives us our result, noting that $\sum_{v_{\mathcal{B}}} O(m_{\mathcal{B}}) = O(n)$. Note that hybrid detection requires qCFs to be computed for $O(n)$ 4-taxon subsets when running in fast mode; as expected qCFs reveal the quarnet topology, we say it makes $O(n)$ **quarnet queries**. ◀

3.2 Pivot Scanning

After determining the hybrid block of $Part(v_{\mathcal{B}})$, it remains to determine the circle order. NetCS leverages a *pivot* to linearize the circular order for sorting.

▶ **Definition 2** (Pivots). *Let N^- be a binary, semi-directed, level-1 network. Then, any non-trivial blob $\mathcal{B}$ in N^- has two reticulation edges whose source vertices are called* pivots.

The deletion of the cut edge incident to each pivot vertex induces a pivot block.

▶ **Lemma 2.** *Let $Part(\mathcal{B}) = P_1|\ldots|P_{m_{\mathcal{B}}}$ be a partition induced by a non-trivial blob $\mathcal{B}$ in a binary, semi-directed, level-1 network N^-, and let P_h be the hybrid block of $Part(\mathcal{B})$. Then, P_p is a pivot block of $Part(\mathcal{B})$ if and only if for all $s, t \in \{1, \ldots, m_{\mathcal{B}}\}$ such that $s \neq t \neq p \neq h$ and all $x_h \in P_h, x_p \in P_p, x_s \in P_s, x_t \in P_t$, $N^-|_{\{x_h,x_p,x_s,x_t\}}$ displays quartet $x_h, x_p|x_s, x_t$.*

The proof of Lemma 2 is in the Appendix A. The idea is that the taxa from the pivot and hybrid blocks must be siblings in one of the two quartets displayed by the quarnet. As the lowest of the three qCFs corresponds to the non-displayed quartet by Proposition 11 in [7], the qCFs identify one block, out of four, that is neither the hybrid nor one of the two pivots. Then, scanning through all blocks in $Part(v_{\mathcal{B}})$ gives the two pivot blocks by process of elimination (Algorithm 2 in Appendix A). This approach requires a method $\mathcal{M}_k$ for computing the qCFs based on four distinct blocks of $Part(v_{\mathcal{B}})$. In this setting, taxa from the same block are exchangeable, as previously noted, so in fast mode, one taxon is arbitrarily selected from each block in $\{P_h, P_p, P_s, P_t\}$ for qCF computation; however, computing the average qCFs from all 4-taxon subsets covering $\{P_h, P_p, P_s, P_t\}$ may increase robustness to noise, which occurs when qCFs are computed from a finite number of gene trees with estimation error and/or missing taxa. Both approaches enable consistent estimation.

▶ **Theorem 2.** *Assume the same settings as Lemma 1. Let π be the error tolerance:*

$$\pi = \min_{Y \in f_4^4(S)} \min\{|\mathbb{P}_{\mathcal{N}^-}(q_1^Y) - \mathbb{P}_{\mathcal{N}^-}(q_3^Y)|, |\mathbb{P}_{\mathcal{N}^-}(q_2^Y) - \mathbb{P}_{\mathcal{N}^-}(q_3^Y)|\} \quad (1)$$

where $f_4^4(S)$ is the set of all 4-taxon subsets $Y \subseteq S$ such that $N^-|_Y$ contains a 4-blob and q_3^Y is the quartet out of the three possible on Y that is not displayed by $N^-|_Y$. Let $v_{\mathcal{B}}$ be some vertex in the TOB for N^- with degree $m_{\mathcal{B}} \geq 5$. Then, for every $\epsilon > 0$, Algorithm 2 applied to $Part(v_{\mathcal{B}})$, with hybrid block P_h, and $k \geq \frac{2}{\pi^2} \log \frac{6|f_4^4(S)|}{\epsilon}$ gene trees returns the (correct) two pivot blocks of $Part(v_{\mathcal{B}})$ with probability at least $1 - \epsilon$. The time complexity of applying Algorithm 2 to all $v_{\mathcal{B}}$ is $O(nk)$ in fast mode and $O(n^2k)$ otherwise.

Proof. By Lemma 2, Algorithm 2 returns the two pivot blocks for all $v_{\mathcal{B}}$ when the lowest qCF on Y corresponds to the non-displayed quartet for all $Y \in f_4^4(S)$ such that $N^-|_Y$ contains a 4-blob. In fast mode or otherwise, this occurs when the error in the observed qCFs is bounded by $\pi/2$, specifically: $w_k(q_i^Y)/k > \mathbb{P}_{\mathcal{N}^-}(q_i^Y) - \pi/2$ and $w_k(q_i^Y)/k < \mathbb{P}_{\mathcal{N}^-}(q_i^Y) + \pi/2$ for all $i = \{1, 2, 3\}$ and all $Y \in f_4^4(S)$. Following the approach in [33], let $X_g^{i,Y}$ be a random variable which takes on value 1 when the restriction of gene tree G_g to Y is isomorphic to q_i^Y and takes on value 0 otherwise. Then, $(X_g^{i,Y})_{g=1}^k$ is a sequence of independent Bernoulli random variables with success probability $\mathbb{P}_{\mathcal{N}^-}(q_i^Y)$, and the observed qCF is $w_k(q_i^Y) = \sum_{g=1}^k X_g^{i,Y}$ with expected value $\mathbb{E}_{\mathcal{N}^-}[w_k(q_i^Y)] = k \cdot \mathbb{P}_{\mathcal{N}^-}(q_i^Y)$. Now, we can bound the probability that $w_k(q_i^Y)/k$ deviates from its expected value by $\pi/2$ or more using Hoeffding's inequality: $\mathbb{P}[|(w_k(q_i^Y) - \mathbb{E}_{\mathcal{N}^-}[w_k(q_i^Y)])/k| \geq \pi/2] \leq 2e^{-2k(\pi/2)^2} = 2e^{-k\pi^2/2}$ for all $i \in \{1, 2, 3\}$ and all $Y \in f_4^4(S)$. Let $E_{i,Y}$ denote the event that $|(w_k(q_i^Y) - \mathbb{E}_{\mathcal{N}^-}[w_k(q_i^Y)])/k| \geq \pi/2$. Then, by the union bound, $\mathbb{P}(\cup_{Y \in f_4^4(S)} \cup_{i=1}^3 E_{i,Y}) \leq \sum_{Y \in f_4(S)} \sum_{i=1}^3 \mathbb{P}(E_{i,Y}) \leq 6|f_4^4(S)|e^{-k\pi^2/2}$. Lastly, we solve $1 - 6|f_4^4(S)|e^{-k\pi^2/2} \leq 1 - \epsilon$ for k to get our result. To summarize, Algorithm 2 is consistent under the NMSC and, as $|f_4^4(S)| = O(n^4)$, has sample complexity $O(\frac{1}{\pi^2} \log \frac{n}{\epsilon})$.

Now consider time complexity. Algorithm 2 computes qCFs for $O(m_{\mathcal{B}})$ 4-block subsets. Each qCF computation takes $O(k)$ time in fast mode, taking advantage of the preprocessing performed for hybrid detection; otherwise, computing averages via the algorithm in Fig. S14 of [32] takes $O(nk)$ time. Repeating for all $v_{\mathcal{B}}$ gives $O(\sum_{v_{\mathcal{B}}} m_{\mathcal{B}}k) = O(nk)$ and $O(\sum_{v_{\mathcal{B}}} m_{\mathcal{B}}nk) = O(n^2k)$. Pivot scan makes $O(n)$ quartet queries in fast mode. ◀

3.3 Circular Order Sorting

Taking the hybrid block P_h and a pivot block P_p together enables us to linearize the circular order of blob $\mathcal{B}$ as the canonical vector $V_{\mathcal{B}}(P_h, P_p) = [P_h, P_p, ..., P_{p'}]$, where $P_{p'}$ is the block of $Part(\mathcal{B})$ induced by the other pivot. Then, mapping each block $P_s \in Part(\mathcal{B})$ to its index in the canonical vector $V_{\mathcal{B}}(P_h, P_p)$ yields a total ordering on the blocks, denoted $L_{\mathcal{B}}(P_h, P_p)$. As an example, traversing the only cycle in N^- of Fig. 1A from the hybrid vertex u_1 towards the pivot u_2, ignoring the direction of reticulation edges, gives $V_{\mathcal{B}}(P_1, P_2) = [P_1, P_2, P_3, P_4, P_5, P_6]$. Traversing in the opposite direction towards the other pivot u_6 gives $V_{\mathcal{B}}(P_1, P_6) = [P_1, P_6, P_5, P_4, P_3, P_2]$. Now consider the relation in $L_{\mathcal{B}}(P_h, P_p)$ for two distinct blocks $P_s, P_t \in Part(\mathcal{B})$ that are neither the hybrid P_h nor the focal pivot P_p.

▶ **Lemma 3.** *Let $Part(\mathcal{B}) = P_1|\dots|P_{m_{\mathcal{B}}}$ be a partition induced by a non-trivial blob $\mathcal{B}$ in a binary, semi-directed, level-1 network N^-. Let P_h be the hybrid block and P_p be a pivot block of $Part(\mathcal{B})$. Then, for all pairs of blocks $P_s, P_t \in Part(\mathcal{B})$ such that $s \neq t \neq h \neq p$, $P_s < P_t$ is in $L_{\mathcal{B}}(P_h, P_p)$ if and only if for all $x_h \in P_h, x_p \in P_p, x_s \in P_s, x_t \in P_t$, $N^-|_{\{x_h, x_p, x_s, x_t\}}$ displays quartet $x_t, x_h|x_p, x_s$; otherwise, $P_t < P_s$ is in $L_{\mathcal{B}}(P_h, P_p)$.*

The proof of Lemma 3 is in the Appendix A. The idea is that one of the two displayed quartets will put the hybrid and the focal pivot together and the other will separate the hybrid and the focal pivot, revealing the linearized order of the non-pivot, non-hybrid blocks P_s, P_t. As the lowest of the three qCFs corresponds to the non-displayed quartet by Proposition 11 in [7], the ordering of P_s, P_t can be determined by comparing the qCF values of the two quartets that separate the pivot and the hybrid (Algorithm 3 in Appendix A). This approach is closely related to the identifiability result for circular orders; however, the innovation behind NetCS is leveraging a pivot to linearize the order and then recovering it in an efficient manner via sorting (Algorithm 4 in Appendix A).

▶ **Theorem 3.** *Assume the same settings as Lemma 1. Let $v_{\mathcal{B}}$ be some vertex in the TOB for N^- with degree $m_{\mathcal{B}} \geq 5$. Then, for every $\epsilon > 0$, Algorithm 4 applied to $Part(v_{\mathcal{B}})$, with hybrid block P_h and pivot block P_p, and $k \geq \frac{2}{\pi^2} \log \frac{6|f_4^4(S)|}{\epsilon}$ gene trees returns the (correct) circular order of $Part(v_{\mathcal{B}})$ with probability at least $1-\epsilon$. The total time complexity of applying Algorithm 4 to all $v_{\mathcal{B}}$ is $O(n \log nk)$ in fast mode and $O(n^2 \log(n)k)$ otherwise, where n is the number of taxa.*

Proof. By Lemma 3, Algorithm 3 returns the correct relation in $L_{\mathcal{B}}(P_h, P_p)$ for all pairs $P_s, P_t \in Part(v_{\mathcal{B}})$ such that $s \neq t \neq h \neq p$ provided the lowest qCF corresponds to the non-displayed quartet for all $Y \in f_4^4(S)$ such that $N^-|_Y$ contains a 4-blob. Following the same argument in the proof of Theorem 2, this occurs with probability at least $1-\epsilon$ when the number k of gene trees is at least $\frac{2}{\pi^2} \log \frac{6|f_4^4(S)|}{\epsilon}$. And, when the correct relations are returned, the (correct) circular order of $Part(\mathcal{B})$ is returned by merge sort (Algorithm 3–4).

Now consider time complexity. Merge sort makes $O(m_{\mathcal{B}} \log m_{\mathcal{B}})$ pairwise block comparisons. Each comparison requires qCF computation, which is $O(k)$ time in fast mode or $O(nk)$ time otherwise, as previously discussed for pivot scan. Repeating for all $v_{\mathcal{B}}$ gives us our result. Circular sorting makes $O(n \log n)$ quarnet queries when running in fast mode. ◀

3.4 Theoretical and Practical Considerations

Combining Theorems 1, 2, and 3 gives our final result:

▶ **Corollary 1.** *NetCS is statistically consistent under the NMSC and has time complexity $O(n \log(n)k)$ in fast mode or $O(n^2 \log(n)k)$ otherwise for k genes and n species. In fast mode, NetCS computes qCFs for $O(n \log n)$ 4-taxon subsets, referred to as quarnet queries.*

NetCS's fast mode is for the regime where the error in all qCFs is below the tolerance threshold $\pi/2$, so every quarnet can be reconstructed with perfect accuracy (i.e., we have an oracle for quarnets). Only $O(n)$ quarnets are accessed for hybrid voting and pivot scan; this increases to $O(n \log n)$ quarnets for circular sorting; thus, NetCS matches the lower bound of quarnet complexity recently established for semi-directed, level-1 network reconstruction (Proposition 6 in [13]). Moreover, when given direct access to quarnets, rather than inferring them from qCFs, NetCS has time complexity $O(n \log n)$. This improves upon the time complexity of full semi-directed level-1 network reconstruction from quarnets, which can be can be done in $O(n^2)$ using the taxon-addition-like algorithm recently presented by Frohn *et al.* (2025) (Theorem 13 in [13]). Like fast mode, the results in [13] are primarily of theoretical interest given the assumption of a perfect oracle (and there is no implementation of their algorithm to our knowledge).

Although our analysis primarily focused on fast mode, we initially designed the slower version NetCS, with the goal of increasing robustness to noisy qCFs, and only later defined the fast version to enable comparisons to [13]. Beyond the extra computation previously

mentioned, our implementation of NetCS also runs circular sorting taking each block in $Part(v_{\mathcal{B}})$, excluding the hybrid, as a candidate pivot, with the goal of increasing robustness to noisy qCFs. to increase the robustness. An error boolean is returned for each pairwise block comparison made during sorting with Algorithm 3, indicating whether the quartet with the hybrid and candidate pivot as siblings had the lowest qCF. This scenario should never occur if the candidate is the correct pivot by Theorem 2. Thus, NetCS reconstructs the blob taking the circular order with the lowest total error score, breaking ties to favor the candidate pivots originally returned from pivot scan and otherwise breaking ties arbitrarily. These additional steps do not impact statistical consistency because both pivots returned by the pivot scan will have score 0, with probability going to 1 as the number of gene trees goes to infinity, at which point, no other (incorrect) candidate can achieve a lower total error score. However, testing all candidate pivot blocks increases the total time complexity of NetCS to $O(n^3 \log(n)k)$. As we will show, even with this additional computation, NetCS is still fast in practice likely because it avoids accessing qCFs for all subsets of four taxa. Precomputation of all qCFs not only has an $O(n^4k)$ time cost but also an $\Omega(n^4)$ storage cost, which can lead to large constant factors in qCF access from the memory hierarchy (e.g., cache misses during either qCF precomputation or blob reconstruction).

NetCS is not without practical limitations, including those inherited from employing the same algorithmic framework of NANUQ+. First, an estimated TOB is required as input, which may contain varying levels of error. Second, all blobs are reconstructed independently, which can lead to an incompatible set of inferred hybrids (i.e., no valid semi-directed network exists that realizes all inferred hybrids simultaneously [18]). To address this issue, our implementation of NetCS processes blobs in decreasing order of degree and excludes blocks that conflict with previously inferred hybrids during the voting procedure. If all blocks are excluded, NetCS does not attempt to reconstruct that blob and thus fails to reconstruct the network. Likewise, NANUQ+ can fail to produce a valid output network due to conflicts among inferred hybrids.

3.5 Experimental Evaluation

To study the empirical performance of NetCS, we conducted an evaluation study using data sets simulated under the NMSC, so the true level-1 network and its tree of blobs were known. Details, including software commands and version numbers, are given in Appendix B.

Experiment 1. The goal of experiment 1 was to evaluate the accuracy of level-1 blob reconstruction with NetCS. Data sets [20] used for this evaluation were previously simulated by [10] with 12 model conditions corresponding to three numbers of taxa (50, 100, and 200) and four branch length scaling factors (2.0, 1.0, 0.5, 0.25). The scaling factors varied the level of gene tree discordance due to ILS (note that the mean discordance levels were 45%, 65%, 80%, and 90%, respectively, based on simulated level-0 networks). Twenty replicate networks were available per model condition. For each replicate, 1000 gene trees were simulated under the NMSC with independent inheritance. The number of (true) gene trees given as input to NetCS was varied from 100, 500, and 1000, yielding 36 conditions in total.

NetCS was given the true TOB as input so that hybrid detection and circular order reconstruction could be evaluated for blobs with degree at least 5. **Hybrid error** was defined as the proportion of blobs for which the predicted hybrid block differed from the true hybrid block. **Circular order error** was quantified using the Kendall tau distance. For two linear permutations σ_1, σ_2 on the set $B = \{1, 2, \ldots, m_{\mathcal{B}}\}$, the Kendall tau distance $D_{k\tau}(\sigma_1, \sigma_2)$ is the fraction of unordered pairs $(i, j) \in B$ whose relative ordering differs between σ_1 and

σ_2 [9]. We extended this notion to circular orders by considering all linearizations obtained by cutting the circle at each position as well as its reversal. Let $\Sigma(\zeta)$ denote the set of all linearizations of circular order ζ, and let $\Sigma(\zeta)^R$ denote their reversals. Then, the Kendall tau distance between two circular orders ζ_1, ζ_2 is

$$D_{k\tau}(\zeta_1, \zeta_2) = \min_{(\sigma_1,\sigma_2)\in(\Sigma(\zeta_1)\cup\Sigma(\zeta_1)^R)\times(\Sigma(\zeta_2)\cup\Sigma(\zeta_2)^R)} D_{k\tau}(\sigma_1, \sigma_2). \quad (2)$$

We reported hybrid error and circular order error averaged over all blobs of degree at least 5 in the true tree of blobs for each replicate network.

Experiment 2. The goal of experiment 2 was to benchmark NetCS against NANUQ+ in terms of blob reconstruction error as well as computational efficiency. This evaluation used the same data sets as experiment (1), except that true gene trees were replaced with gene tree estimated from simulated DNA sequences using IQ-TREE model selection [26, 21] (sometimes gene tree estimation from simulated DNA sequences failed but the total number of estimated gene trees was always greater than 900). Mean gene tree estimation error, quantified in terms of (normalized) Robinson-Foulds error [31], ranged from 40% to 70%. Both NetCS and NANUQ+ were given the true TOB as input so that reconstructed blobs could be compared. Computational efficiency was evaluated as wallclock time (24-hour limit) and peak resident set size (peak RSS; 256 GB limit) (note that the reported time and memory usage does not include TOB reconstruction, as the true TOB was given as input). Both methods were run on compute nodes with the same architecture (AMD EPYC 7313) when collecting these performance metrics. NANUQ+ was given access to 8 threads for qCF precomputation.

Experiment 3. The goal of the third experiment was to evaluate NetCS for end-to-end semi-directed level-1 network reconstruction in comparison with NANUQ+ and CAMUS. Data sets used for this evaluation were previously simulated for the study presenting CAMUS [39] (note that CAMUS could not be run in the previous experiments because the networks did not have a single taxon outgroup, which led to issues correctly rooting CAMUS' estimated base tree). The CAMUS data sets had four model conditions with 51, 101, 151, and 201 taxa, including a single taxon outgroup. Replicate networks with blobs of degree 4 were excluded, as the hybrid is not identifiable for those blobs; this left 47, 45, 43, and 49 replicates (out of 50) for the 51-, 101-, 151-, and 201-taxon conditions, respectively. Each replicate had 1000 gene trees. Mean gene tree estimation error was $\sim 20\%$; ILS level was not reported by [38].

Since experiment 3 was based on data sets from the study presenting CAMUS [38], CAMUS was run with the same default and recommended settings. The base tree for CAMUS was estimated with ASTRAL-IV [42] and rooted at the outgroup taxon. Unlike the other methods, CAMUS returns a *collection* of (binary) *directed*, level-1 networks, one after each edge addition; we used the directed network with the same number of edge additions as the number of non-trivial blobs (i.e., hybrid vertices) as the true network, with the goal of giving CAMUS an advantage. For running NetCS and NANUQ+, the ASTRAL-IV tree was used to estimate a TOB following the approach of [10]; see Appendix B for details. NetCS and NANUQ+ were run twice, once given the estimated TOB as input and once given the true TOB as input.

All methods were evaluated based on hybrid ancestry error and network topology error. A species has **hybrid ancestry** if it is a hybrid or is the descendant of one or more hybrids. False positives (FPs) correspond to the estimated network predicting a species has hybrid ancestry when it does not according to the true network; false negatives (FNs) are the reverse. FPR is the number of FPs divided by the number of species with hybrid ancestry according

to the estimated network. FNR is the number of FNs divided by the number of species with hybrid ancestry according to the true network. Balanced error rate, defined as the arithmetic mean of the FPR and FNR, was reported for all methods. Network topology error was evaluated as the **normalized hardwired cluster distance (HWCD)** [19]. Each edge in a directed network induces a hardwired cluster of all the taxon descendants of the edge. The hardwired clusters induced by all edges of the network form the hardwired cluster system. The HWCD is the symmetric difference of the hardware cluster systems of two directed networks. When comparing level-1 network, HWCD is a distance [25, 6, 8]. The extension to semi-directed networks takes the minimum value from all possible ways to "root" the two networks [18]. The normalized version divides HWCD by the total number of edges in both networks minus the total number of leaves in both networks. We used the implementation of HWCD in PhyloNetworks [35], treating all networks as semi-directed.

4 Experimental Results

We now present the results of the three experiments.

Experiment 1. The goal of experiment 1 was to evaluate how model condition impacted level-1 blob reconstruction with NetCS (Table 2, Fig. 2). As the number of gene trees and the branch scale factor increased, error decreased. For example, the average Kendall tau distance dropped from 0.14 to 0.0 going from the hardest model condition (branch scaling factor 0.25, 100 true gene trees) to the easiest (branch scaling factor 2.0, 1000 true gene trees), both with 50 taxa. These trends are expected as larger numbers of true gene trees correspond to lower error in observed qCFs and larger branch length scaling factors correspond to lower ILS, which, in turn, corresponds to higher error tolerance π. Similar trends were observed for model conditions with larger numbers of taxa and hybrid error. Average hybrid error was generally low at 1000 gene trees, ranging from 0.12 to 0.00 for the hardest to easiest 50-taxon model conditions (see Table 2 for details). NetCS did not have any failures in experiment 1.

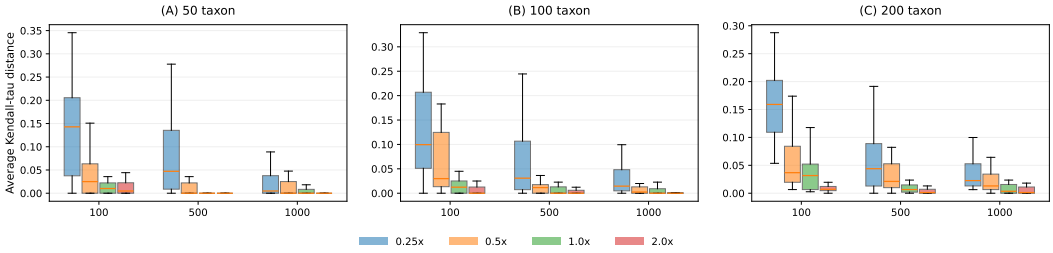

Figure 2 Experiment 1: Average Kendall-tau distance. Three subfigures (A), (B),(C) correspond to data sets with 50 taxa, 100 taxa, and 200 taxa, respectively. The x-axis is the number of true gene trees. Each colored box represents a different branch scaling factor in order of decreasing ILS. The y-axis is error in circular order reconstruction (i.e., average Kendall-tau distance across all blobs in replicate). Orange lines represent medians. Outliers are not shown.

Experiment 2. The goal of experiment 2 was to compare NetCS against NANUQ+ in terms of level-1 blob reconstruction error and computational efficiency (Table 1, Fig. 6). Again, the percent hybrids correctly predicted increased with decreasing ILS level, with NetCS making the correct call for 79%, 83%, 88%, 95% of blobs in 50-taxon data sets, respectively. NetCS outperformed NANUQ+ in hybrid detection on 5 model conditions, the reverse occurred on

1 model condition, and the two methods tied in the remaining 2 model conditions. Error in circular order reconstruction also decreased with increasing ILS level, with NetCS and NANUQ+ tied, each outperforming the other on 4 conditions. The difference between methods for both tasks was small, suggesting they achieved similar accuracy.

In contrast, NetCS and NANUQ+ differed substantially in terms of their computational efficiency. On the 50-taxon data sets, Net-CS completed in under 5 seconds on average while using less than 0.33GB memory, whereas NANUQ$^+$ required ∼7–8 minutes on average and over 0.7GB memory. The performance gap increased for the 100-taxon data sets. Net-CS completed in 20 seconds on average while using less than 0.4GB memory, whereas NANUQ$^+$ required more than one hour on average and over 50GB memory. On the 200-taxon data sets, NANUQ$^+$ either exceeded the 24-hour wall-clock limit or the 256GB memory limit, while NetCS completed within 15–215 seconds using only about 0.5GB memory. Note that NetCS did not have any failures for experiment 2, but NANUQ+ failed to return a valid network for one 50-taxon network (id #1207) for three branch scaling factors (0.25×, 0.5×, 2×) and one 100-taxon network (id #10032) for one branch scaling factor (0.25×); these data sets were excluded from error and runtime comparisons.

Table 1 Experiment 2. The table shows the number of blobs, across all replicate networks for the same model condition, in which NetCS is better / worse / tied with NANUQ+ in terms of hybrid error and circular order error (Kendall tau distance). The Kendall-tau distance is the average across all blobs in a replicate, which is then averaged across all replicates. Runtime and peak RSS values are averages across all replicates per model condition. Results are not shown for 200 taxa, as NANUQ+ did not complete within the maximum wallclock time of 24 hours or exceeded the memory limit of 256GB. Both methods were given up to 1000 estimated gene trees.

		# of blob			# of blob			# of		Average		Runtime		Peak RSS	
# of	ILS	hybrid error			Kendall-tau			correct hybrids		Kendall-tau		(MB)		(secs)	
taxa	Level	Better	Worse	Tie	Better	Worse	Tie	Net-CS	NANUQ+	Net-CS	NANUQ+	Net-CS	NANUQ+	Net-CS	NANUQ+
50	0.25×	3	3	18	6	**9**	9	19	19	**0.043**	0.048	**330**	709	**5**	480
50	0.5×	1	**2**	21	**8**	3	13	20	**21**	0.018	**0.012**	**330**	710	**5**	466
50	1.0×	**2**	1	24	5	**6**	16	**24**	23	0.020	**0.011**	**329**	709	**4**	444
50	2.0×	**1**	0	23	**5**	3	16	**23**	22	**0.011**	0.013	**329**	708	**4**	440
100	0.25×	1	1	28	8	**9**	13	27	27	0.034	**0.031**	**396**	50582	**20**	4711
100	0.5×	**2**	1	30	9	**11**	13	**31**	30	**0.022**	0.034	**396**	50183	**17**	4415
100	1.0×	**2**	0	31	**10**	9	14	**32**	30	**0.016**	0.018	**395**	50741	**12**	4158
100	2.0×	**1**	0	32	**9**	7	17	**32**	31	**0.010**	0.026	**392**	51384	**10**	4230

Experiment 3. The goal of the final experiment was to benchmark NetCS as part of an end-to-end pipeline for network reconstruction against NANUQ+ and CAMUS. For hybrid ancestry error, NetCS (with estimated TOB), NANUQ+ (with estimated TOB), and CAMUS all exhibited poor performance (Fig. 3). CAMUS had the highest variability, sometimes achieving balanced error close to 0 and other times close to 1; however, its median balanced error rate was greater than 0.5 for all model conditions with more than 50 taxa. NetCS outperformed CAMUS and NANUQ+ in terms of hybrid ancestry predictions on these conditions, although its balanced error rate was still high. Providing NetCS and NANUQ+ with the true TOB dropped the balanced error rate to near 0 for all numbers of taxa. Thus, relatively low levels of error in the estimated TOB (0.06–0.08 on average; Table 3 in the Appendix C) can negatively impact hybrid ancestry predictions.

For network topology error, CAMUS, NetCS (with estimated TOB), and NANUQ+ (with estimated TOB) produced networks with nHWCD between 0.05–0.15 (Fig. 4). Although CAMUS slightly outperformed NetCS on 50 taxa, with median normalized HWCD (nHWCD) of 0.08 vs. 0.10. NetCS slightly outperformed CAMUS on 100 taxa (median nHWCD: 0.06 vs. 0.08). The two methods had similar performance on 150 and 200 taxa. NetCS slightly outperformed NANUQ+ in terms of topological error: median nHWCD of 0.10 vs. 0.13

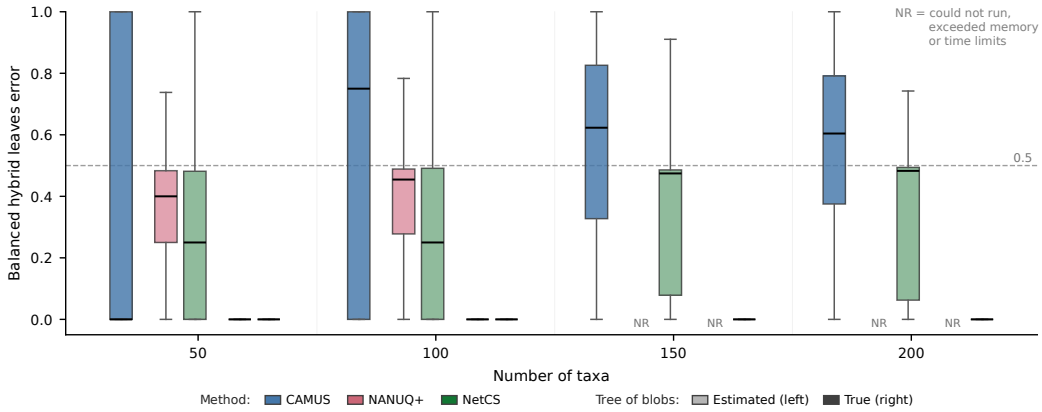

Figure 3 Experiment 3: Hybrid ancestry balanced error rate. The y-axis is the hybrid ancestry balanced error rate. The x-axis is the number of taxa. Each box represents a method, from left to right: CAMUS, NANUQ+ (given the estimated TOB), NetCS (given the estimated TOB), NANUQ+ (given true TOB), and NetCS (given true TOB). NANUQ+ failed to run on 150- and 200-taxon data sets, indicated by "NR." Bold black lines represent medians. Outliers are not shown. Dashed line represents the error rate of a random guess.

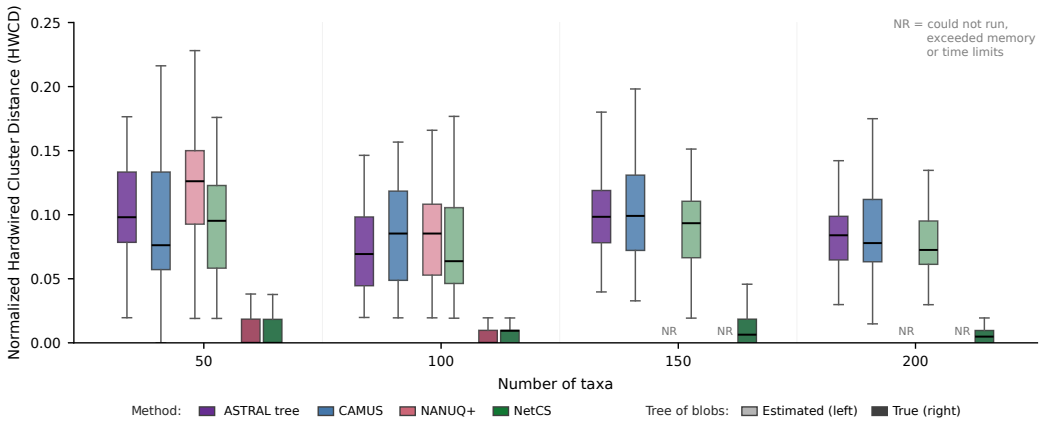

Figure 4 Experiment 3: Normalized HWCD. The y-axis is network topological error (normalized HCWD). The x-axis is the number of taxa. Each box represents a method, from left to right: NANUQ+ (given the estimated TOB), NetCS (given the estimated TOB), NANUQ+ (given true TOB), and NetCS (given true TOB). NANUQ+ failed to run on 150- and 200-taxon data sets, indicated by "NR." Bold black lines represent medians. Outliers are not shown.

and 0.06 vs. 0.08 for 50 and 100 taxa, respectively. Overall, the topological error was close to that of the ASTRAL-IV tree used by all methods. CAMUS uses the ASTRAL-IV tree as a base tree, adding edges to it. The nHWCD of the CAMUS network was better/worse than the ASTRAL-IV tree for 27/18, 19/24, 22/20, and 27/22 replicates for the 51-, 101-, 151-, and 201-taxon model conditions; thus, the edge additions to the base tree made by CAMUS decreased nHWCD slightly more often then they increased it. Providing NetCS and NANUQ+ with the true TOBs, instead of an estimated TOB reconstructed from the ASTRAL-IV tree, dropped nHWCD to near 0 for all numbers of taxa.

On the 150- and 200-taxon data sets, NANUQ$^+$ either exceeded the 24-hour wall-clock limit or the 256GB memory limit, while NetCS completed within 15–215 seconds using only about 0.5GB memory. NANUQ+ (given estimated TOBs) failed to return a valid network

for 2 replicates with 50 taxa and 2 with 100 taxa, whereas NetCS (given estimated TOBs) failed to return a valid network for 1 replicate with 150 taxa; these replicates were excluded to ensure fair comparison.

5 Conclusions

In this paper, we studied the problem of reconstructing level-1 blobs under the NMSC model when the tree of blobs is known. This work was motivated by recent advances in tree of blobs reconstruction [3, 17, 13, 10, 30] and by the two-step design of NANUQ+ for reconstructing semi-directed level-1 networks. Our method, NetCS, combines majority voting to identify the hybrid vertex with sorting to recover the circular order. These primitives closely reflect prior identifiability results for semi-directed level-1 networks [34, 7, 1], although, to our knowledge, they had not previously been operationalized in a fast, statistically consistent method.

A key contribution of NetCS is that it avoids the precomputation of all qCFs by (1) repurposing the 3f1a algorithm from our recent study [10] and (2) applying an efficient sorting algorithm. This gives NetCS time complexity $O(n \log(n)k)$ and only $O(n \log n)$ quarnet queries, matching the lower bound of [13]. These results indicate that consistent reconstruction of level-1 blobs under the NMSC is easy, at least from a theoretical perspective.

Notably, our implementation of NetCS performs additional computation to improve robustness to noisy qCF estimates, leading to $O(n^3 \log(n)k)$ time complexity. However, by avoiding the quartic cost of qCF precomputation, NetCS enables analyses of data sets with 200 taxa and 1000 genes in only a few minutes, substantially improving upon NANUQ+ in both runtime and memory usage. Despite its reduced access to qCFs, NetCS was still highly competitive with NANUQ+ in terms of accuracy, achieving comparable error rates for both hybrid detection and circular order reconstruction when provided with the true tree of blobs. However, when supplied with an estimated, and poor quality, tree of blobs, the accuracy of both methods deteriorated substantially. This finding suggests that the primary practical challenge may lie in fast and accurate tree of blobs reconstruction.

Direct network reconstruction circumvents this issue, motivating us to benchmark NetCS and NANUQ+ against the recently proposed method CAMUS. We did not evaluate SNaQ because we used the same data sets as the CAMUS study, which had already shown that SNaQ was less topologically accurate than CAMUS. We found that CAMUS, NetCS, and NANUQ+ had comparable network topology error and all methods failed to accurately determine whether species evolved through hybridization events, a question of importance in biological studies.

Collectively, these findings highlight the need for improved methods for level-1 network reconstruction. Given the computational efficiency and strong performance of NetCS when provided with the true tree of blobs, a promising direction for future research is to combine NetCS with improved methods for tree of blobs reconstruction. The algorithmic techniques underlying NetCS also could be explored in the context of direct network reconstruction or other approaches for level-1 blob reconstruction. For example, reducing the number of 4-taxon subsets sampled in NANUQ+ or replacing collections of 4-taxon subsets with averages. Such techniques could also be relevant for applications of the recent method Squirrel [16] to gene tree data. Squirrel is similar to NANUQ+, in that both are based on quartet distances, but was introduced in the context of level-1 blob reconstruction from DNA sequences.

References

1 Elizabeth S Allman, Hector Baños, Marina Garrote-Lopez, and John A Rhodes. Identifiability of level-1 species networks from gene tree quartets. *Bulletin of Mathematical Biology*, 86(9):110, 2024. doi:10.1007/s11538-024-01339-4.

2 Elizabeth S Allman, Hector Baños, Jonathan D Mitchell, and John A Rhodes. The tree of blobs of a species network: identifiability under the coalescent. *Journal of Mathematical Biology*, 86(1):10, 2023. doi:10.1007/s00285-022-01838-9.

3 Elizabeth S Allman, Hector Baños, Jonathan D Mitchell, and John A Rhodes. TINNiK: inference of the tree of blobs of a species network under the coalescent model. *Algorithms for Molecular Biology*, 19(1):23, 2024. doi:10.1186/s13015-024-00266-2.

4 Elizabeth S Allman, Hector Baños, John A Rhodes, and Kristina Wicke. NANUQ+: a divide-and-conquer approach to network estimation. *Algorithms for Molecular Biology*, 20(1):14, 2025. doi:10.1186/s13015-025-00274-w.

5 Cécile Ané, John Fogg, Elizabeth S Allman, Hector Baños, and John A Rhodes. Anomalous networks under the multispecies coalescent: theory and prevalence. *Journal of Mathematical Biology*, 88(3):29, 2024. doi:10.1007/s00285-024-02050-7.

6 Allan Bai, Péter L Erdős, Charles Semple, and Mike Steel. Defining phylogenetic networks using ancestral profiles. *Mathematical Biosciences*, 332:108537, 2021.

7 Hector Baños. Identifying species network features from gene tree quartets under the coalescent model. *Bulletin of Mathematical Biology*, 81(2):494–534, 2019. doi:10.1007/s11538-018-0485-4.

8 Gabriel Cardona, Joan Carles Pons, Gerard Ribas, and Tomás Martínez Coronado. Comparison of orchard networks using their extended μ-representation. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 21(3):501–507, 2024. doi:10.1109/TCBB.2024.3361390.

9 Vincent A Cicirello. Kendall tau sequence distance: Extending Kendall tau from ranks to sequences. *arXiv preprint arXiv:1905.02752*, 2019. arXiv:1905.02752.

10 Junyan Dai, Yunheng Han, and Erin K Molloy. Quartet-based species tree methods enable fast and consistent tree of blobs reconstruction under the network multispecies coalescent. *bioRxiv*, 2025. doi:10.1101/2025.11.05.686850.

11 Vu Dinh and Hector Baños. Misspecification strikes: Astral can mislead in the presence of hybridization, even for nonanomalous scenarios. *Molecular Biology and Evolution*, 42(3):msaf049, 2025. doi:10.1093/molbev/msaf049.

12 John Fogg, Elizabeth S Allman, and Cécile Ané. PhyloCoalSimulations: A simulator for network multispecies coalescent models, including a new extension for the inheritance of gene flow. *Systematic Biology*, 72(5):1171–1179, 2023. doi:10.1093/sysbio/syad030.

13 Martin Frohn, Niels Holtgrefe, Leo van Iersel, Mark Jones, and Steven Kelk. Reconstructing semi-directed level-1 networks using few quarnets. *Journal of Computer and System Sciences*, 152:103655, 2025. doi:10.1016/j.jcss.2025.103655.

14 D. Gusfield. *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*. Cambridge University Press, Cambridge, United Kingdom, 1997.

15 Dan Gusfield, Vikas Bansal, Vineet Bafna, and Yun S. Song. A decomposition theory for phylogenetic networks and incompatible characters. *Journal of Computational Biology*, 14(10):1247–1272, 2007. doi:10.1089/cmb.2006.0137.

16 Niels Holtgrefe, Katharina T Huber, Leo van Iersel, Mark Jones, Samuel Martin, and Vincent Moulton. Squirrel: Reconstructing semi-directed phylogenetic level-1 networks from four-leaved networks or sequence alignments. *Molecular Biology and Evolution*, 42(4):msaf067, 2025. doi:10.1093/molbev/msaf067.

17 Katharina T Huber, Leo van Iersel, Mark Jones, Vincent Moulton, and Leonie Veenema-Nipius. When are quarnets sufficient to reconstruct semi-directed phylogenetic networks? *Bulletin of Mathematical Biology*, 87(10):136, 2025. doi:10.1007/s11538-025-01510-5.

18 Katharina T. Huber, Leo van Iersel, Remie Janssen, Mark Jones, Vincent Moulton, Yukihiro Murakami, and Charles Semple. Orienting undirected phylogenetic networks. *Journal of Computer and System Sciences*, 140:103480, 2024. doi:10.1016/j.jcss.2023.103480.

19 Daniel H Huson, Regula Rupp, and Celine Scornavacca. *Phylogenetic Networks*. Cambridge University Press, Cambridge, England, 2011.

20 Erin K. Molloy Junyan Dai, Yunheng Han. Data from: Quartet-based species tree methods enable fast and consistent tree of blobs reconstruction under the network multispecies coalescent, 2026. doi:10.5061/dryad.cjsxksnnp.

21 Subha Kalyaanamoorthy, Bui Quang Minh, Thomas KF Wong, Arndt Von Haeseler, and Lars S Jermiin. Modelfinder: fast model selection for accurate phylogenetic estimates. *Nature methods*, 14(6):587–589, 2017.

22 Nathan Kolbow, Sungsik Kong, Tyler Chafin, Joshua Justison, Cécile Ané, and Claudia Solís-Lemus. SNaQ.jl: Improved scalability for phylogenetic network inference. *bioRxiv*, 2025. doi:10.1101/2025.11.17.688917.

23 Nathan Kolbow, Sungsik Kong, and Claudia Solís-Lemus. A method for massively scalable phylogenetic network inference. *bioRxiv*, 2025. doi:10.1101/2025.05.05.652278.

24 Sungsik Kong, David L Swofford, and Laura S Kubatko. Inference of phylogenetic networks from sequence data using composite likelihood. *Systematic Biology*, 74(1):53–69, 2024. doi:10.1093/sysbio/syae054.

25 Michael Maxfield, Jingcheng Xu, and Cécile Ané. A dissimilarity measure for semidirected networks. *IEEE Transactions on Computational Biology and Bioinformatics*, 2025. doi:10.1109/TCBBIO.2025.3534780.

26 Bui Quang Minh, Heiko A Schmidt, Olga Chernomor, Dominik Schrempf, Michael D Woodhams, Arndt von Haeseler, and Robert Lanfear. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, 37(5):1530–1534, 2020. doi:10.1093/molbev/msaa015.

27 Erin K Molloy, Arun Durvasula, and Sriram Sankararaman. Advancing admixture graph estimation via maximum likelihood network orientation. *Bioinformatics*, 37(Supplement_1):i142–i150, 2021. doi:10.1093/bioinformatics/btab267.

28 R Alexander Pyron, Kyle A O'Connell, Edward A Myers, David A Beamer, and Hector Baños. Complex hybridization in a clade of polytypic salamanders (plethodontidae: Desmognathus) uncovered by estimating higher-level phylogenetic networks. *Systematic Biology*, 74(1):124–140, 2025. doi:10.1093/sysbio/syae060.

29 Bruce Rannala and Ziheng Yang. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, 164(4):1645–1656, 2003. doi:10.1093/genetics/164.4.1645.

30 John A. Rhodes, Elizabeth S. Allman, Cecile Ané, and Hector Baños. Statistical inference of the tree of blobs of a phylogenetic network from quartet concordance factors. *bioRxiv*, 2026. doi:10.64898/2026.05.28.728501.

31 D.F. Robinson and L.R. Foulds. Comparison of phylogenetic trees. *Mathematical Biosciences*, 53(1):131–147, 1981. doi:10.1016/0025-5564(81)90043-2.

32 Erfan Sayyari and Siavash Mirarab. Fast coalescent-based computation of local branch support from quartet frequencies. *Molecular Biology and Evolution*, 33(7):1654–1668, 2016. doi:10.1093/molbev/msw079.

33 Shubhanshu Shekhar, Sebastien Roch, and Siavash Mirarab. Species tree estimation using ASTRAL: how many genes are enough? *IEEE/ACM transactions on computational biology and bioinformatics*, 15(5):1738–1747, 2017. doi:10.1109/tcbb.2017.2757930.

34 Claudia Solís-Lemus and Cécile Ané. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. *PLOS Genetics*, 12(3):1–21, 2016. doi:10.1371/journal.pgen.1005896.

35 Claudia Solís-Lemus, Paul Bastide, and Cécile Ané. Phylonetworks: A package for phylogenetic networks. *Molecular Biology and Evolution*, 34(12):3292–3298, 2017. doi:10.1093/molbev/msx235.

36 Alexandros Stamatakis. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313, 2014. doi:10.1093/bioinformatics/btu033.

37 Mike Steel. *Phylogeny: Discrete and random processes in evolution*, volume 89 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, 2016.

38 James Willson and Tandy Warnow. CAMUS: scalable phylogenetic network estimation. *Bioinformatics*, 42(Supplement_1):btag245, 2026. doi:10.1093/bioinformatics/btag245.

39 James Willson and Tandy Warnow. Dataset for CAMUS: Scalable phylogenetic network estimation, 2026. doi:10.13012/B2IDB-6892704_V1.

40 Yun Yu, James H. Degnan, and Luay Nakhleh. The probability of a gene tree topology within a phylogenetic network with applications to hybridization detection. *PLOS Genetics*, 8(4):1–10, 2012. doi:10.1371/journal.pgen.1002660.

41 Yun Yu, Jianrong Dong, Kevin J. Liu, and Luay Nakhleh. Maximum likelihood inference of reticulate evolutionary histories. *Proceedings of the National Academy of Sciences*, 111(46):16448–16453, 2014. doi:10.1073/pnas.1407950111.

42 Chao Zhang and Siavash Mirarab. Weighting by gene tree uncertainty improves accuracy of quartet-based species trees. *Molecular Biology and Evolution*, 39(12):msac215, 2022. doi:10.1093/molbev/msac215.

43 Chao Zhang, Rasmus Nielsen, and Siavash Mirarab. Aster: A package for large-scale phylogenomic reconstructions. *Molecular Biology and Evolution*, page msaf172, 2025. doi:10.1093/molbev/msaf172.

A Supplemental Methods

▶ **Lemma 2.** *Let $Part(\mathcal{B}) = P_1|\dots|P_{m_\mathcal{B}}$ be a partition induced by a non-trivial blob $\mathcal{B}$ in a binary, semi-directed, level-1 network N^-, and let P_h be the hybrid block of $Part(\mathcal{B})$. Then, P_p is a pivot block of $Part(\mathcal{B})$ if and only if for all $s,t \in \{1,\dots,m_\mathcal{B}\}$ such that $s \neq t \neq p \neq h$ and all $x_h \in P_h, x_p \in P_p, x_s \in P_s, x_t \in P_t$, $N^-|_{\{x_h,x_p,x_s,x_t\}}$ displays quartet $x_h, x_p|x_s, x_t$.*

Proof. Because x_h, x_p, x_s, x_t are drawn from distinct blocks of $Part(\mathcal{B})$ and x_h, in particular, is drawn from the hybrid block, $N^-|_{\{x_h,x_p,x_s,x_t\}}$ contains a 4-blob by Lemma 5 in [10]. Let u_j denote the vertex in $\mathcal{B}$ such that the deletion of the one cut edge incident to u_j induces the bipartition: $P_j|S \setminus P_j$ for all $j \in \{1,\dots,m_\mathcal{B}\}$.

$\Rightarrow$ Assume that P_p is a pivot block of $Part(\mathcal{B})$. By definition, there exists a reticulation edge (u_p, u_h) in N^-. After deleting the partner reticulation edge of (u_p, u_h), any path of distinct edges between $x_h \in P_h$ and $x_p \in P_p$ in $N^-|_{\{x_h,x_p,x_s,x_t\}}$ contains edge (u_p, u_h) but no other edges in $\mathcal{B}$. In contrast, any path of distinct edges between x_h and $x_s \in P_s$ or $x_t \in P_t$ in $N^-|_{\{x_h,x_p,x_s,x_t\}}$ contains edge (u_p, u_h) as well as the (non-reticulation) edge in $\mathcal{B}$ that is incident to (u_p, u_h), which induces bipartition $x_h, x_p|x_s, x_t$ giving us our result.

$\Leftarrow$ For the sake of obtaining a contradiction, assume the condition holds but P_p is not a pivot block. Then, there exists two pivot blocks P_v, P_w where $w \neq v \neq p \neq h$. For all $x_h \in P_h, x_p \in P_p, x_v \in P_v, x_w \in P_w$, $N^-|_{\{x_h,x_p,x_v,x_w\}}$ contains a 4-blob that displays two quartets: $x_h, x_v|x_w, x_p$ from deleting reticulation edge (u_w, u_h) and $x_h, x_w|x_v, x_p$ from deleting reticulation edge (u_v, u_h), as described in the forward direction of the proof. As $N^-|_{\{x_h,x_p,x_s,x_t\}}$ displays two quartets given our assumption that N^- is level-1 and binary [7], quartet $x_h, x_p|x_v, x_w$ is not displayed, which is a contradiction. ◀

Algorithm 2 Pivot Scan Algorithm.

input : Partition of species set S induced by vertex $v_{\mathcal{B}}$, denoted $Part(v_{\mathcal{B}}) = P_1|P_2|\ldots|P_{m_{\mathcal{B}}}$, Hybrid block of $Part(v_{\mathcal{B}})$, denoted P_h, Set of k unrooted, binary gene trees each on (a subset of) S

output : Two pivot blocks of $Part(v_{\mathcal{B}})$

if *fast mode* **then**
  Select one taxon arbitrarily from P_i to use for all qCF calculations, repeating for all $i \in \{1, 2, \ldots, m_{\mathcal{B}}\}$;
  $\mathcal{M}_k \leftarrow$ algorithm for computing qCF on fly based on binary lifting algorithm
else
  $\mathcal{M}_k \leftarrow$ algorithm for computing average qCF from Fig. S14 in [32]
$\mathcal{K} \leftarrow \{P_1, P_2, ..., P_m\} \setminus \{P_h\}$;
$P_s, P_t \leftarrow$ 2 arbitrary elements of $\mathcal{K}$; $\mathcal{K} \leftarrow \mathcal{K} \setminus \{P_s, P_t\}$;
repeat
  $P_v \leftarrow$ an arbitrary element in $\mathcal{K}$; $\mathcal{K} \leftarrow \mathcal{K} \setminus \{P_v\}$;
  $q_v \leftarrow$ qCF for $P_h, P_v|P_s, P_t$ computed via $\mathcal{M}_k$;
  $q_s \leftarrow$ qCF for $P_h, P_s|P_v, P_t$ computed via $\mathcal{M}_k$;
  $q_t \leftarrow$ qCF for $P_h, P_t|P_s, P_v$ computed via $\mathcal{M}_k$;
  if $q_s < q_v$ *and* $q_s < q_t$ **then** $P_s \leftarrow P_v$;
  else if $q_t < q_v$ *and* $q_t < q_s$ **then** $P_t \leftarrow P_v$;
until $\mathcal{K} = \emptyset$;
return P_s, P_t;

▶ **Lemma 3.** *Let $Part(\mathcal{B}) = P_1|\ldots|P_{m_{\mathcal{B}}}$ be a partition induced by a non-trivial blob $\mathcal{B}$ in a binary, semi-directed, level-1 network N^-. Let P_h be the hybrid block and P_p be a pivot block of $Part(\mathcal{B})$. Then, for all pairs of blocks $P_s, P_t \in Part(\mathcal{B})$ such that $s \neq t \neq h \neq p$, $P_s < P_t$ is in $L_{\mathcal{B}}(P_h, P_p)$ if and only if for all $x_h \in P_h, x_p \in P_p, x_s \in P_s, x_t \in P_t$, $N^-|_{\{x_h,x_p,x_s,x_t\}}$ displays quartet $x_t, x_h|x_p, x_s$; otherwise, $P_t < P_s$ is in $L_{\mathcal{B}}(P_h, P_p)$.*

Proof. Because x_h, x_p, x_s, x_t are all drawn from distinct blocks of $Part(\mathcal{B})$ and x_h, in particular, is drawn from the hybrid block, $N^-|_{\{x_h,x_p,x_s,x_t\}}$ contains a 4-blob by Lemma 5 in [10]. Let u_j denote the vertex in $\mathcal{B}$ such that the deletion of the one cut edge incident to u_j induces the bipartition: $P_j|S \setminus P_j$ for all $j \in \{1, \ldots, m_{\mathcal{B}}\}$.

$\Rightarrow$ Assume $P_s < P_t$ is in $L_{\mathcal{B}}(P_h, P_p)$. Then, traversing $\mathcal{B}$ from the hybrid vertex u_h toward pivot u_p encounters u_s before u_t. It follows that the restriction of N^- to $\{x_h, x_p, x_s, x_t\}$ has two reticulation edges (u_p, u_h) and (u_t, u_h), and deleting reticulation edge (u_p, u_h) gives the displayed quartet $x_t, x_h|x_p, x_s$.

$\Leftarrow$ For the sake of obtaining a contradiction, assume the condition holds but $P_t < P_s$ is in $L_{\mathcal{B}}(P_h, P_p)$. The restriction of N^- to $\{x_h, x_p, x_s, x_t\}$ displays quartets $x_s, x_h|x_p, x_t$ and $x_h, x_p|x_s, x_t$ by the forward direction above and by Lemma 2, respectively. Only two quartets can be displayed given our assumption that N^- is level-1 and binary [7]; thus, quartet $x_t, x_h|x_p, x_s$ is not displayed, which is a contradiction. ◀

Algorithm 3 Linearized Circular Order Comparison Algorithm.

input : Distinct blocks P_h, P_p, P_s, P_t from the partition of species set S induced by vertex $v_{\mathcal{B}}$, where P_h is the hybrid block and P_p is the pivot block, Set of k unrooted, binary gene trees each on (a subset of) S, A method $\mathcal{M}_k$ for computing qCFs from gene trees given 4 pairwise disjoint subsets of S

output : Relation of P_s and P_t in $L_{\mathcal{B}}(P_h, P_p)$ and an error boolean

$error \leftarrow 0$;
$q_p \leftarrow$ qCF for $P_p, P_h | P_s, P_t$ computed via $\mathcal{M}_k$;
$q_s \leftarrow$ qCF for $P_s, P_h | P_p, P_t$ computed via $\mathcal{M}_k$;
$q_t \leftarrow$ qCF for $P_t, P_h | P_p, P_s$ computed via $\mathcal{M}_k$;
if $q_p < q_s$ *and* $q_p < q_t$ **then**
    $error \leftarrow 1$;
if $q_s < q_t$ **then**
    return $P_s < P_t, error$;
else
    return $P_t < P_s, error$;

Algorithm 4 Linearized Circular Sorting Algorithm.

input : Partition of species set S induced by vertex $v_{\mathcal{B}}$, denoted $Part(v_{\mathcal{B}})$, Hybrid block of $Part(v_{\mathcal{B}})$, denoted P_h, Pivot block of $Part(v_{\mathcal{B}})$, denoted P_p, Set of k unrooted, binary gene trees each on (a subset of) S

output : Linearized circular order of $Part(v_{\mathcal{B}})$, denoted $V_{\mathcal{B}}$, and its error score

if *fast mode* **then**
    Select one taxon arbitrarily from P_i to use for all qCF calculations, repeating for all $i \in \{1, 2, \ldots, m_{\mathcal{B}}\}$;
    $\mathcal{M}_k \leftarrow$ algorithm for computing qCF on fly based on binary lifting algorithm;
else
    $\mathcal{M}_k \leftarrow$ algorithm for computing average qCF from Fig. S14 in [32]
$V_{\mathcal{B}}, error \leftarrow$ Run merge sort on $Part(v_{\mathcal{B}}) \setminus \{P_h, P_p\}$ using Algorithm 3 to compare block pairs, summing the error booleans across all comparisons;
$V_{\mathcal{B}} \leftarrow$ Append $[P_h, P_p]$ to the front of $V_{\mathcal{B}}$;
return $V_{\mathcal{B}}, error$

B Software Commands

B.1 NetCS Command

NetCS (available at `https://github.com/molloy-lab/TREE-QMC`) was run with command:

```
tree-qmc --network -i [input gene tree] --at [input TOB] -o [ouput]
```

Commit e577f8c was used for experiments 1 and 2 and commit eb3e140 was used for experiment 3.

B.2 NANUQ+ Commands

NANUQ+, implemented in MSCQuartets (v3.2), was run with commands:

```
library(ape)
library(MSCquartets)
gttree <- read.tree("[input gene trees]")
QT <- quartetTableParallel(gttree, numCores = 8)
RQT <- quartetTableToRQT(QT)
pT <- quartetTreeTestInd(RQT, model=t3_model)
pT <- quartetStarTestInd(pT)
ToB <- read.tree("[input TOB]")
ToB <- labelIntNodes(ToB, plot=FALSE)
nets <- resolveLevel1(ToB, pT, alpha=[alpha], beta=[beta], plot=FALSE)
final_nwk <- nets[[1]][[1]]
```

For experiments 1–2, NANUQ+ was given the true TOB and significance thresholds: $\alpha = 10^{-7}, \beta = 0.99$ based on hyperparameter tuning the study simulating the data sets [10]. For experiment 3, NANUQ+ was given a TOB estimated with TOB-QMC and significance thresholds: $\alpha = 10^{-4}, \beta = 0.99$, as described in Section B.4.

B.3 ASTRAL-IV Command

ASTRAL-IV (v1.23.4.6) was run with command:

```
astral4 -t 32 -o [output] [input gene trees]
```

The resulting tree was rooted at the outgroup using the scripts provided by the CAMUS study [38] with the following command:

```
python3 root-outgroup.py [ASTRAL tree]
```

B.4 TOB-QMC commands

TOB-QMC (commit eb3e140, available at `https://github.com/molloy-lab/TREE-QMC`) was run with the command:

```
tree-qmc --blobsearchonly [input ASTRAL tree] \
         --store_pvalue --3f1a --override \
         -i [input gene trees] \
         -o [TOB refinement tree]
```

to annotate the refinement tree computed with ASTRAL-IV with p-values using NANUQ+'s hypothesis test (T3 option). TOBs were constructed by contracting branches based on different significance thresholds with the following command:

```
tree-qmc --load_pvalue \
         --alpha [alpha] --beta 0.99 \
         -i [annotated  tree] -o [output TOB]
```

The results of this tuning experiment are shown in Fig. 5. The estimated TOBs given as input to NetCS and NANUQ+ used significance threshold $\alpha = 10^{-4}$.

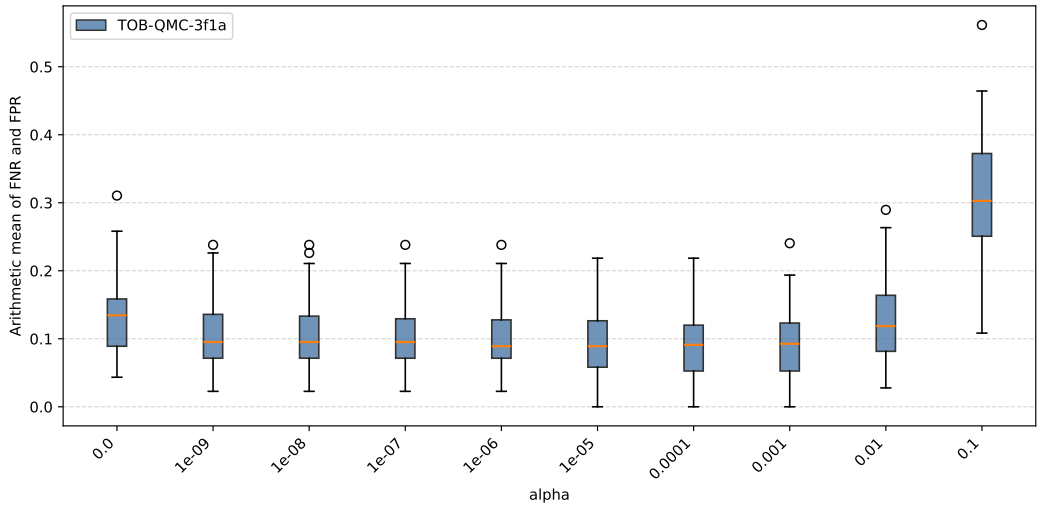

Figure 5 Experiment 3: Hypertuning parameters for TOB estimation. The x-axis is the significance threshold α required by TOB-QMC-3f1a. The y-axis is topological error between the true and estimated TOB, specifically its the arithmetic mean of FNR and FPR. FNR is the number of branches in the true TOB missing from the estimated TOB divided by the number of internal branches in the true TOB. FPR is the number of branches in the estimated TOB missing from the true TOB divided by the number of internal branches in the estimated TOB. Each boxplot has 50 replicate datasets for the 26-taxon model condition from the CAMUS study; these data sets are not used for method evaluation and benchmarking. Orange lines are medians.

B.5 CAMUS Command

CAMUS (v1.0.0-6-g3138bcf) was run with command:

```
camus -o [output prefix] [rooted base tree] [input gene trees]
```

where the rooted base tree was computed with ASTRAL-IV.

C Supplemental Results

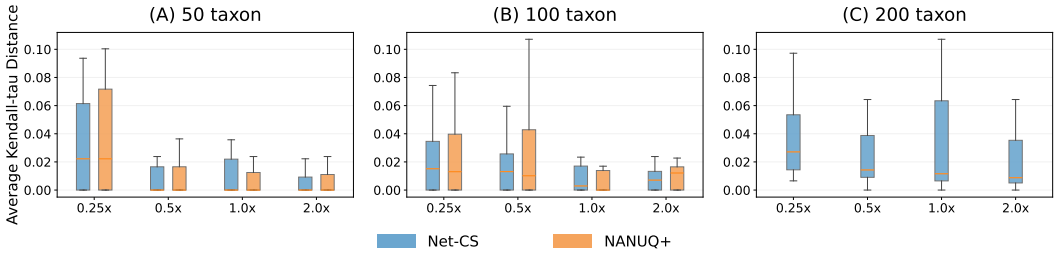

Figure 6 Experiment 2: Average Kendall-tau distance. Three subfigures (A), (B),(C) correspond to data sets with 50 taxa, 100 taxa, and 200 taxa, respectively. The y-axis is the average Kendall-tau distance, which evaluates the error in circular order reconstruction. The x-axis is the scaling factor: 0.25×, 0.5×, 1.0×, and 2.0× in order of decreasing ILS. Each colored box represents a method, either NetCS or NANUQ+. Both methods were given all estimated gene trees, up to 1000. For 200 taxon data, only Net-CS results are presented since NANUQ$^+$ did not complete within the maximum wallclock time of 25 hours. Orange lines are medians. Outliers are not shown.

Table 2 Experiment 1. Each row shows results for particular model condition defined by the number of taxa and branch scaling factor, which in turn varies ILS level. Evaluation metrics are averaged over all replicates for the same model condition. Columns indicate number of gene trees, varying from 100, 500, to 1000. Hybrid error is the fraction of incorrect hybrid across all blobs in a network. Kendall-tau distance is averaged across all blobs in a network.

# of taxa	scale factor	Hybrid error			Kendall-tau distance			Peak RSS (MB)			Runtime (secs)		
		100	500	1000	100	500	1000	100	500	1000	100	500	1000
50	0.25×	0.708	0.392	0.117	0.144	0.082	0.032	273	299	331	4	5	5
50	0.5×	0.308	0.092	0.100	0.050	0.015	0.012	273	299	330	7	4	6
50	1.0×	0.142	0.100	0.108	0.017	0.003	0.020	274	298	329	4	3	4
50	2.0×	0.100	0.050	0.000	0.022	0.004	0.003	274	298	328	4	3	4
100	0.25×	0.525	0.142	0.108	0.136	0.058	0.033	282	333	395	12	11	17
100	0.5×	0.100	0.025	0.000	0.070	0.031	0.016	283	332	395	8	11	13
100	1.0×	0.042	0.025	0.025	0.028	0.020	0.011	283	332	395	7	7	10
100	2.0×	0.017	0.000	0.000	0.008	0.004	0.005	285	334	395	5	8	9
200	0.25×	0.650	0.042	0.054	0.168	0.057	0.034	300	398	524	41	40	65
200	0.5×	0.096	0.037	0.025	0.054	0.038	0.025	299	399	525	25	31	55
200	1.0×	0.062	0.025	0.050	0.045	0.019	0.019	302	401	526	27	25	49
200	2.0×	0.013	0.025	0.025	0.015	0.010	0.011	303	403	528	19	22	46

Table 3 Experiment 3: TOB topological error. Number of taxa in model condition for CAMUS data sets. Number of replicates remaining after excluding those whose ground truth network contains one or more 4-blob. Error between the true and estimated TOB are reported as FNR and FPR, defined in Figure 5 caption. Error values are mean ± standard deviation across replicates.

Taxa	# of Replicates	FNR	FPR	(FNR+FPR)/2
50	47	0.039 ± 0.029	0.120 ± 0.053	0.079 ± 0.033
100	45	0.038 ± 0.017	0.078 ± 0.040	0.058 ± 0.024
150	43	0.039 ± 0.018	0.088 ± 0.034	0.064 ± 0.022
200	49	0.040 ± 0.014	0.077 ± 0.022	0.059 ± 0.015

Revision Notice

This is a revised version of the eponymous paper that appeared in the proceedings of WABI 2026 (LIPIcs, volume 390, https://www.dagstuhl.de/dagpub/978-3-95977-446-8, published in August, 2026). The revised version corrects an error in computational experiment 3. The authors discovered that the trees of blobs (TOB) estimated for experiment #3 did not satisfy the properties of the protocol used to reconstruct them. Although there was a minor bug in the TOB reconstruction code, this bug did not explain the issue with the files, and ultimately, the authors were unable to determine their provenance. The authors re-ran the analysis scripts for experiment 3 starting with TOB estimation. The resulting TOBs satisfied the properties of the protocol and were much more accurate. This, in turn, changed the results of experiment 3, as all methods that used the estimated TOBs as input performed considerably better. However, this correction did not change the overall trends or conclusions of the study.

List of changes for experiment 3

- Corrected Figure 3
- Corrected Figure 4
- Corrected Figure 5
- Added Table 3
- Corrected the text in the Results section
- Corrected the text in the Discussion section (last two sentences of second to last paragraph)
- Added commit numbers for TREE-QMC in Appendix B
- Corrected the text in the section on TOB reconstruction in Appendix B

To prepare the revision, the authors additionally rechecked the other two experiments; they corrected the following minor issues.

List of changes for experiment 1

- Corrected Figure 2 for experiment 1 so that no replicates are excluded (previously the authors had inadvertently excluded the failed replicates from experiment 2)
- Corrected text in the Results section so that no replicates are excluded
  - Replaced 0.15 with 0.14, 0.123 with 0.12, and 0.000 with 0.00
  - Swapped order of "hardest" and "easiest" in the second to last sentence
- Corrected Table 2 so that no replicates are excluded

List of changes for experiment 2

- Corrected reporting of failures in last sentence of the Results section (NANUQ+ failed on 4 replicates rather than 2, as previously reported)
- Corrected Figure 6 to exclude replicates with failures

List of miscellaneous changes

- Added link to analysis scripts and data files
- Corrected typo in"Assumptions and Identifiability" paragraph (changed Fig. 1A to Fig. 1D)
- Corrected typo the definition of hybrid ancestry definition given in Section 3.5
- Updated citation [38]

Dagstuhl Publishing – September 28, 2026.