Network Delay Inference from Additive Metrics

Shankar Bhamidi
Department of Statistics
University of California, Berkeley

Ram Rajagopal
Department of EECS
University of California, Berkeley

Sébastien Roch
Department of Statistics
University of California, Berkeley

April 17, 2006

Abstract

We demonstrate the use of computational phylogenetic techniques to solve a central problem in inferential network monitoring. More precisely, we design a novel algorithm for multicast-based delay inference, i.e. the problem of reconstructing the topology and delay characteristics of a network from end-to-end delay measurements on network paths. Our inference algorithm is based on additive metric techniques widely used in phylogenetics. It runs in polynomial time and requires a sample of size only $\mathrm{poly}(\log n)$.

1 Introduction

Inferential network monitoring, also known as Network Tomography [Va96], aims at reconstructing various properties of large communication networks from indirect measurements to help oversee the performance of these networks. Network inference can be achieved by two general approaches. In the *internal* approach, one takes measurements directly at the edges and nodes of the network. However, this approach suffers from several drawbacks: direct measurements may create an extra computational burden as well as congestion in the network; the network operator may not allow access to internal devices of the network or may not make public any measurements on them; the routers may not have the technological capabilities of performing the required measurements. This has led some in the networking community to consider instead the *external* approach. In this case, one uses so-called "end-to-end" measurements, e.g. measurements of delays or rate of packet drops between nodes in the network, and seeks to infer the desired network properties from them. This gives rise to an inverse problem similar to tomographic image reconstruction, whereby the name Network Tomography.

Our aim in this paper is to propose a novel approach to this problem. We focus on multicast-based inference. Multicast routing consists in sending a packet from a source to a set of receivers through a routing tree. The packet is duplicated at each branch point and sent further down the tree. The routing tree is generally unknown to the user. The idea is to use inherent correlation of measurements between different receivers to reconstruct the topology of the routing tree as well as to estimate link properties of this tree. The main link property we consider here is the delay distribution. The multicast inference approach was introduced in [CD+99, LD+02].

A core difficulty of the problem is to devise scalable and efficient algorithms which consistently estimate the desired network properties. Several techniques have been used in the literature, notably maximum pseudo-likelihood, EM algorithms and Markov chain Monte Carlo methods. See [CC+04] for a detailed survey and bibliographic references. In this paper, we introduce a new methodology inspired by techniques from the field of phylogenetics. Our methodology has the advantage of being provably efficient and consistent. It also uses a small (asymptotic) sample size. This is crucial to reduce the burden on the network as well as to obtain a consistent "snapshot" of the network, which is intrinsically dynamic in nature. Concurrently to our work, Liang et al. [LMY06] used similar ideas to tackle the multicast packet loss inference problem.

The main problem in phylogenetics is to infer evolutionary histories from molecular data. Evolution is usually represented by a tree where branching points indicate speciation events. The root of the tree is the common ancestor to all species in the tree and the leaves are contemporary (extant) species. Molecular data is assumed to evolve according to a standard Markov model. The *phylogenetic reconstruction* problem is the following. From measurement of sequences of molecular data at the leaves, one seeks to reconstruct the topology of the evolutionary tree as well as mutation characteristics along the branches. See [Fe04] and [SS03] for an overview of the field of phylogenetics.

Various statistical and computational techniques have been used to solve the phylogenetic reconstruction problem: maximum likelihood, Markov chain Monte Carlo, parsimony, distance-based methods. In this paper, we extend this last technique to a class of models different from the Markov models used in phylogenetics and more suitable to the network monitoring problem. The main idea in distance-based methods is to define a so-called tree metric from mutation characteristics. A tree metric is a metric on the leaves of the tree which can be realized as a path metric on a corresponding weighted tree. (See Section 2 for more details.) After being estimated, the metric allows the reconstruction of the tree and its characteristics. A main advantage of this approach is that it leads to provably efficient algorithms. See, in particular, the dyadic closure method of Erdos et al. [ES+97].

1.1 Definitions and Results

A Renewal-Type Process on a Tree. We now give a more formal statement of the multicast inference problem introduced in [LD+02]. Let $T = (V, E)$ be a binary tree on $n + 1$ leaves L—representing the *routing tree*—and let $\{d_e\}_{e \in E}$ be a set of independent positive random variables on the edges—representing the *delays*. Leaf 0, the *source*, is the root of the tree. The remaining n leaves are the *receivers*.

A realization of the *multicast delay* process works as follows: the root sends a packet to the receivers through the routing tree; at every branching point, the packet is duplicated; on every link e, an independent random delay d_e is experienced by the packet. More formally, we define the multicast delay process $\{D_u\}_{u \in V}$ as follows. Let P_{ij} be the path (set of edges) between nodes i and j in T. For a node u, let

$$D_u = \sum_{e \in P_{0u}} d_e.$$

Note that D_u is the total delay at node u in the network. One can think of $\{D_u\}_{u \in V}$ as an inhomogeneous renewal-type process on a tree. That is, for all u, D_u is a sum of indepedent (but not identical) positive random variables.

The Multicast Inference Problem. The tree and delay distributions are actually unknown to us. We are only given access to k independent samples of delays at the leaves $\{D_a^1\}_{a \in L}, \ldots, \{D_a^k\}_{a \in L}$. Our goal is to reconstruct the routing tree and estimate the delay distributions using these samples. For this task to be tractable, we need to make a number of assumptions.

Assumption 1 (Basic Assumptions) *We assume that the tree T is binary and fully resolved (not necessarily complete), that is all internal vertices have degree 3. Furthermore, we assume that the delays are uniformly bounded, namely there is a constant $M > 0$ independent of n such that for all $e \in E$, $d_e \in [0, M]$.*

We need to define more precisely what we mean by "estimating" the delay distributions. In this work, we assume that each edge delay distribution (usually all different) is characterized by a constant, say

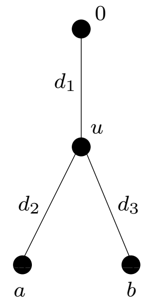

Figure 1: Unidentifiability of Location Parameter: If one were to replace d_1 with $d_1 + \mu$ and d_2, d_3 with $d_2 - \mu, d_3 - \mu$ for $\mu > 0$ (assuming μ can be chosen so that all delays remain positive) then the distribution of delays at a, b would be unchanged. This example also shows that one cannot deduce the delays on all edges given total delays at all leaves.

$J - 1 > 1$ (independent of n), number of (consecutive) central moments. That is, we assume there are *characteristic moments*

$$w_e^{(j)} = \mathbb{E}\left[(d_e - \mathbb{E}[d_e])^j\right],$$

for all $e \in E$ and $2 \leq j \leq J$. Our goal is to estimate these moments within a fixed accuracy. More formally, we make the following assumption.

Definition 1 (Regular Families) *Let $\varepsilon > 0$ and $J \geq 2$ be fixed. Let $\mathcal{Q} = \{Q_\theta\}_{\theta \in \Theta}$ be a family of distributions on $\mathbb{R}$ parametrized by $\theta \in \Theta$. Let $\left\{w^{(j)}(\theta)\right\}_{\{2 \leq j \leq J\}}$ be the first $J-1$ central moments of Q_θ. We say that the family $\mathcal{Q}$ is (ε, J)-regular if there exists a map Ψ from $\mathbb{R}^{J-1}$ to Θ and a $\delta > 0$ such that if the vector $\hat{\mathbf{w}} = \left\{\hat{w}^{(j)}\right\}_{\{2 \leq j \leq J\}}$ satisfies*

$$\left|\hat{w}^{(j)} - w^{(j)}(\theta)\right| \leq \delta$$

for all $2 \leq j \leq J$, then

$$\left\|Q_\theta - Q_{\Psi(\hat{\mathbf{w}})}\right\|_1 \leq \varepsilon.$$

Assumption 2 (Regular Family of Delays) *Let $\varepsilon > 0$ and $J \geq 2$ be fixed (independent of n). We assume that all edge delay distributions are from a fixed (ε, J)-regular family of distributions on $[0, M]$.*

This framework is simple enough to be tractable yet general enough to accommodate large classes of distributions: parametrized distributions, e.g. Beta distributions; and nonparametrized distributions, e.g. discretized distributions on $\{0, 1, \ldots, M\}$. Further we need the following assumption.

Assumption 3 (Lower Bound on Second Moment) *We assume that there is a $f > 0$ independent of n such that for all $e \in E$,*

$$w_e^{(2)} > f.$$

The importance of this assumption will become apparent in Section 3. We note in passing that in general the "location parameter" of the edge delay distributions is unidentifiable. See Figure 1 for details. But as noted in [LD$^+$02], in practice, one is only interested in the "variable" part of the delay. The "fixed transmission delay" is not of concern here.

To sum up, the *Multicast Inference Problem* is defined as follows.

Definition 2 (Multicast Inference Problem) *Let $\varepsilon > 0$ and $J \geq 2$ be fixed. The* Multicast Inference Problem *consists in the following. Let T and $\left\{w_e^{(j)}\right\}_{\{e \in E, 2 \leq j \leq J\}}$ be any tree and set of central moments on edges satisfying Assumptions 1 and 3. Given samples of delays at the leaves, we are required to:*

1. **Tree Reconstruction.** *Recover T.*

2. **Moment Estimation.** *Estimate all characteristic moments $\left\{w_e^{(j)}\right\}_{\{e\in E, 2\leq j\leq J\}}$ within ε.*

Our Results. Our main result is the following theorem.

Theorem 1 *Let $\varepsilon > 0$ and $J \geq 2$ be fixed. There is a polynomial-time algorithm which solves the Multicast Inference Problem with high probability using $k = O(\mathrm{poly}(\log n))$ samples.*

See Theorems 3, 4, and 5 below for more precise statements.

1.2 Organization of the Paper

The paper is organized as follows. We start in Section 2 with an overview of phylogenetic reconstruction techniques. Then, we describe our algorithms for the Multicast Inference Problem in Sections 3 and 4. More precisely, in Section 3 we tackle the tree reconstruction problem; in Section 4 we show how to estimate delay distributions. We conclude with remarks and possible extensions in Section 5.

2 Phylogenetic Reconstruction Techniques

In this section, we introduce several techniques from phylogenetics that we will use later to solve the Multicast Inference Problem. The most important notion, for our discussion, is that of a tree metric. Also, we adapt to our setting the dyadic closure method (DCM) of [ES+97], a provably efficient metric-based technique for reconstructing phylogenies.

2.1 Tree metrics

A fundamental notion in phylogenetics is that of a tree metric, which we define next.

Definition 3 (Tree Metric) *Let L be a finite set with cardinality n. A function $W : L \times L \to \mathbb{R}_{++}$ defines a (nondegenerate) tree metric if the following holds. There exist a tree $T = (V, E)$ with leaf set L and a weight function $w : E \to \mathbb{R}_{++}$ such that $W(a, b) = \sum_{e\in P_{ab}} w_e$ for all $a, b \in L$ where P_{ab} is the path between a and b in T.*

In phylogenetics, tree metrics are based on the transition probabilities of the Markov model of evolution [St94]. In our setup, we use moments of the delays instead.

Tree metrics are usually estimated from samples of the tree process at the leaves. In that context, Azuma's inequality is useful (see e.g. [MR95]).

Lemma 1 (Azuma-Hoeffding Inequality) *Suppose $X = (X_1, \ldots, X_k)$ are independent random variables taking values in a set S, and $f : S^k \to \mathbb{R}$ is any t-Lipschitz function: $|f(\mathbf{x}) - f(\mathbf{y})| \leq t$ whenever $\mathbf{x}$ and $\mathbf{y}$ differ at just one coordinate. Then,*

$$\mathbb{P}\left[f(X) - \mathbb{E}[f(X)] \geq \lambda\right] \leq \exp\left(-\frac{\lambda^2}{2t^2k}\right),$$

and

$$\mathbb{P}\left[f(X) - \mathbb{E}[f(X)] \leq -\lambda\right] \leq \exp\left(-\frac{\lambda^2}{2t^2k}\right).$$

■

2.2 Four-Point Method

A *quartet*

$$q = \{a, b, c, d\} \in \binom{L}{4},$$

is a set of four leaves. Up to leaf labeling, there is a unique (3-regular) tree topology on 4 leaves. The internal edge defines a *split*, i.e. a partition of the leaves $\{a, b, c, d\}$. This is usually denoted $ab|cd$ if the internal edge separates $\{a, b\}$ from $\{c, d\}$, that is defines the partition $\{\{a, b\}\{c, d\}\}$. A (nondegenerate) tree metric on T allows to determine uniquely which split holds over a set of leaves $\{a, b, c, d\}$. Indeed, $ab|cd$ holds if and only if

$$W(a, b) + W(c, d) \leq \min\{W(a, c) + W(b, d), W(a, d) + W(b, c)\}.$$

This is the so-called Four-Point Method (FPM). In turn, the set of all quartet splits defines the tree uniquely, and the reconstruction can be done efficiently. See [Fe04, SS03] and references therein.

2.3 Short quartet method

In [ES+97], a particularly efficient reconstruction technique, the dyadic closure method (DCM), is devised which uses only *short quartets*, that is, roughly speaking, quartets involving only leaves at a short distance from each other (see below for a precise definition). This is crucial in the phylogenetic case where "long" distances cannot be estimated with a polynomial number of samples. In the tomography setting, the DCM algorithm allows the reconstruction of the tree using as few as $\mathrm{poly}\log n$ samples.

Let $e = (u, v)$ be an edge. Let $\gamma_u(e)$ (resp. $\gamma_v(e)$) be the shortest (graph) distance from u (resp. v) to the leaves not using edge e. The depth of T is

$$\Delta(T) = \max_{e=(u,v)\in E} \max\{\gamma_u(e), \gamma_v(e)\}.$$

It is easy to show that $\mathrm{depth}(T) = O(\log n)$ if the degree of all nodes is at least 3. Let g be an upper bound on the weights, that is $w_e \leq g$ for all $e \in E$. The quartet $q = \{a, b, c, d\}$ is said to be a *short quartet* if for all $i, j \in \{a, b, c, d\}$, we have

$$W(i, j) \leq (2\,\Delta(T) + 3)g.$$

We let $Q_{\mathrm{s}}(T)$ be the set of short quartet splits. Note that this definition is different from that used in [ES+97], but it will be enough for our purposes.

In a nutshell, the DCM algorithm proceeds in the following manner. It starts by computing all short quartets. It then computes the "closure" of the set of short quartets by repeatedly applying quartet inference rules of the type

$$\text{if } ab|cd \text{ and } ab|ce \text{ are quartet splits of } T, \text{ then so is } ab|de.$$

Other similar rules exist. It can be proven that the set of all quartet splits deduced from applying DCM is in fact the set of all quartet splits of T, from which the tree can be inferred efficiently. If one has access to sufficiently many samples so that the reconstruction of all short quartets can be guaranteed with high probability, then DCM returns the correct tree with high probability. For the full details of the algorithm and the proof, see [ES+97].

We now state the main theorem concerning the performance of DCM. In particular, we reformulate the theorem in a more general setting of tree metrics rather than mutation probabilities as originally stated in [ES+97]. The proof is essentially the same and is omitted here. We first need a lemma on the proper reconstruction of quartets by the Four-Point Method.

Lemma 2 (See e.g. [ES^+97]) *Let $T = (V, E)$ be a tree on four leaves a, b, c, d. Let W be a tree metric on T and $\hat{W}$, an estimate of W. Let f be a lower bound on the weight of the internal edge of T. Then, the Four-Point Method applied to $\hat{W}$ infers the right quartet split for T if the following condition holds: for all $i, j \in \{a, b, c, d\}$, we have*

$$\left|W(i,j) - \hat{W}(i,j)\right| < \frac{f}{2}.$$

■

We also need a few definitions. We assume that we have a tree $T = (V, E)$ with a tree metric W and an estimate $\hat{W}$ of W. Let $f, g > 0$ be bounds on the weights, that is $f \le w_e \le g$ for all $e \in E$. We say that $q = \{a, b, c, d\}$ is an *almost short quartet* if for all $i, j \in \{a, b, c, d\}$, we have

$$W(i,j) \le (2\,\Delta(T) + 3)g + g.$$

We let $Q_{\mathrm{as}}(T)$ be the set of almost short quartet splits. Likewise, we say that $q = \{a, b, c, d\}$ is an *estimated short quartet* if for all $i, j \in \{a, b, c, d\}$, we have

$$\hat{W}(i,j) \le (2\,\Delta(T) + 3)g + \frac{g}{2}. \tag{1}$$

We let $Q_{\mathrm{es}}(T)$ be the set of estimated short quartet splits. A straightforward reformulation of [ES^+97, Theorem 9] in terms of metrics rather than transition probabilities gives the following.

Theorem 2 (Dyadic Closure Method [ES^+97]) *Let $T = (V, E)$ be a 3-regular tree on n leaves with a tree metric W and an estimate $\hat{W}$ of W. Let $f, g > 0$ be bounds on the true edge weights, that is $f \le w_e \le g$ for all $e \in E$. Define the following probabilities:*

$$\begin{aligned}
\mathcal{A} &= \mathbb{P}\left[\{\text{DCM correctly reconstructs } T\}^c\right],\\
\mathcal{B} &= \mathbb{P}\left[\{\text{FPM returns the correct split for all } q \in Q_{\mathrm{as}}(T)\}^c\right],\\
\mathcal{C} &= \mathbb{P}\left[\left\{Q_{\mathrm{s}}(T) \subseteq Q_{\mathrm{es}}(T) \subseteq Q_{\mathrm{as}(T)}\right\}^c\right].
\end{aligned}$$

Then,

$$\mathcal{A} \le \mathcal{B} + \mathcal{C}.$$

■

2.4 Edge Reconstruction from Additive Functions

In addition to reconstructing the topology, we will use metric ideas to estimate moments of edge delays. For this purpose, we will need to recover edge weights from appropriately defined tree metrics. In fact, we will use a notion of "generalized" tree metric which will be useful in treating odd moments. This definition allows for negative edge weights.

Definition 4 (Additive function) *A function on the leaf set of the tree $W : L \times L \to \mathbb{R}$ is called an* additive function *on the leaves if there exists weights w_e on each of the edges (not necessarily positive), such that for all leaves a, b*

$$W(a,b) = \sum_{e \in P_{ab}} w_e.$$

Similarly to the previous subsection, we are given access to an additive function W on the leaves. Our goal is now to recover the w_e's from the function W, assuming further that we are given the tree T. For this purpose, we use a standard algorithm from phylogenetics. See Figures 2, 3, and 4. We will refer to this algorithm as the Additive Function Inference (AFI) algorithm.

Algorithm Additive Function Inference
Input: function W at the leaves;
Output: edge weights w_e, for all $e \in E$;

- **Case 1: Internal Edge**
  - For all internal edges e,
    * Let $S_1, \ldots, S_4$ be the four subtrees hanging from e as in Figure 3;
    * For each S_i, compute u_i the closest (in graph distance) leaf to the root r_i of S_i;
    * Compute
$$w_e = \frac{1}{2}(W(u_1, u_3) + W(u_2, u_4) - W(u_1, u_2) - W(u_3, u_4))$$
- **Case 2: Leaf Edge**
  - For all leaf edges e,
    * Let S_1, S_2 be the two subtrees hanging from e as in Figure 4;
    * For each S_i, compute u_i the closest (in graph distance) leaf to the root r_i of S_i;
    * Compute
$$w_e = \frac{1}{2}(W(a, u_1) + W(a, u_2) - W(u_1, u_2))$$

Figure 2: Algorithm Additive Function Inference.

3 Routing Tree Reconstruction

The goal of this section is to reconstruct efficiently the topology of the routing tree.

3.1 Delay-based metrics

In phylogenetics, tree metrics are based on mutation probabilities. Here, we use moments of delays. As we explain next, for the purpose of reconstructing the topology, the variance suffices; higher moments will be used in later sections to recover delay distributions.

From Definition 3, one can define a tree metric by first choosing a tree—in our case, the (unknown) routing tree—and then defining a weight function on its edges. Any positive quantity can serve as a weight. The important point is that one must be able to estimate the resulting tree metric from samples at the leaves. This governs the choice of the weight function. Let $T = (V, E)$ be the (unknown) routing tree with leaf set L and consider the choice of weights

$$w_e^{(2)} = \mathrm{Var}[d_e],$$

for all $e \in E$ and the corresponding tree metric

$$W^{(2)}(a, b) \equiv \sum_{e \in P_{ab}} \mathrm{Var}[d_e],$$

for all $a, b \in L$. Assumption 3 implies in particular that $W^{(2)}$ is a well-defined (nondegenerate) tree metric.

Our first task is to check that this metric can be estimated from samples at the leaves. Let a, b be leaves and consider the quantity $\delta_{ab}^{(2)} \equiv \mathrm{Var}[D_a - D_b]$. The delays D_a and D_b are observed at the leaves a and b respectively and therefore the variance of $D_a - D_b$ can be easily estimated. Moreover, we claim that the equality $\delta_{ab}^{(2)} = W^{(2)}(a, b)$ holds. Denote γ_{ab} the common ancestor of a and b, i.e. the node at which the paths P_{ab}, P_{0a}, and P_{0b} intersect. Then, by independence of the edge delays, we

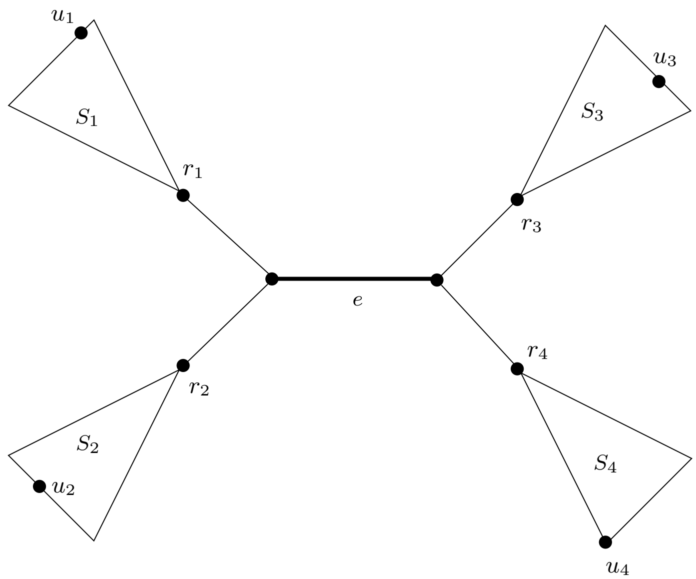

Figure 3: Internal edge case.

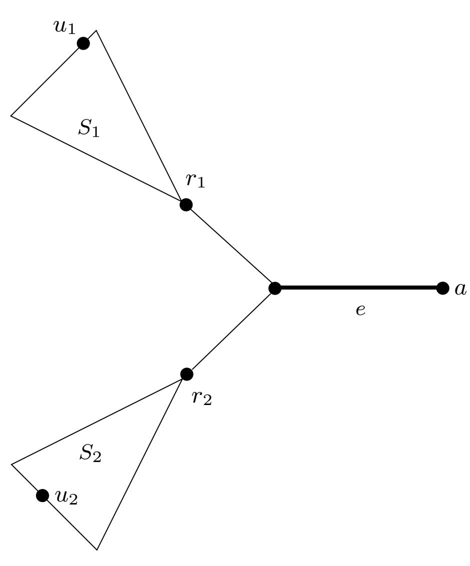

Figure 4: Leaf edge case.

have

$$\begin{aligned}
\delta_{ab}^{(2)} &= \mathrm{Var}[D_a - D_b] \\
&= \mathrm{Var}\left[\sum_{e \in P_{a\gamma_{ab}}} d_e - \sum_{e \in P_{\gamma_{ab}b}} d_e\right] \\
&= \sum_{e \in P_{a\gamma_{ab}}} \mathrm{Var}[d_e] + \sum_{e \in P_{\gamma_{ab}b}} \mathrm{Var}[d_e] \\
&= W^{(2)}(a,b).
\end{aligned}$$

To estimate $\delta_{ab}^{(2)}$, we use the standard unbiased estimator for the variance of $D_a - D_b$

$$\hat{\delta}_{ab}^{(2)} = \frac{1}{k-1} \sum_{i=1}^{k} \left[(D_a^i - D_b^i) - \hat{\delta}_{ab}^{(1)}\right]^2,$$

where

$$\hat{\delta}_{ab}^{(1)} = \frac{1}{k} \sum_{i=1}^{k} (D_a^i - D_b^i).$$

Below, we will need to show that $\hat{\delta}_{ab}^{(2)}$ is well concentrated around $\delta_{ab}^{(2)}$, which follows from the Azuma-Hoeffding inequality (see Lemma 1). The next lemmas provide the necessary Lipschitz condition.

Lemma 3 *Suppose $X = (X_1, \ldots, X_k)$ are independent random variables taking values in $[0, M]$ with $k \geq 2$. Then, the variance estimator*

$$s_X^2 = \frac{1}{k-1} \sum_{i=1}^{k} (X_i - \overline{X})^2 = \frac{1}{k(k-1)} \sum_{i<j} (X_i - X_j)^2,$$

where $\overline{X}$ is the sample average, is $\frac{M^2}{k}$-Lipschitz.

Proof: Let X and Y as above differ in one coordinate. Then

$$\left|s_X^2 - s_Y^2\right| \leq \frac{1}{k(k-1)} \sum_{i<j} \left|(X_i - X_j)^2 - (Y_i - Y_j)^2\right| \leq \frac{1}{k} M^2.$$

■

From Assumption 1, we get immediately the following.

Lemma 4 (Lipschitz Constant for Delay-based Metric) *Say $\hat{\delta}_{ab}^{(2)}$ is computed with k samples. Under Assumption 1, $\hat{\delta}_{ab}^{(2)}$ is then $\frac{|P_{ab}|^2 M^2}{k}$-Lipschitz.*

■

3.2 Inferring the routing tree

Equipped with a legitimate tree metric, we use the DCM algorithm to infer the topology. Here, we use Theorem 2 to prove that the routing tree can be inferred with $\mathrm{poly}\log n$ samples at the leaves. The full algorithm is detailed in Figure 5. This is our main result for this section.

Algorithm Routing Tree Inference
Input: delay samples at the receivers;
Output: estimated routing tree;

- **1) Distance Estimation:**
  - For all $(a,b) \in L \times L$, compute $\hat{\delta}^{(2)}_{ab}$;
- **2) Tree Reconstruction:**
  - Construct all estimated short quartets $Q_{\mathrm{es}}(T)$ (see (1));
  - Apply the DCM algorithm (see [ES+97]) to $Q_{\mathrm{es}}(T)$;
  - Deduce T from the set of quartets computed at the previous step.

Figure 5: Algorithm Routing Tree Inference.

Theorem 3 (Efficient Network Inference) *Let $T = (V, E)$ be the (unknown) routing tree where edge delays satisfy assumptions 1 and 3 (recall that we assume $J \geq 2$). Consider the tree metric $W^{(2)} = \delta^{(2)}$ and assume that the estimate $\hat{W}^{(2)} = \hat{\delta}^{(2)}$ is computed using k samples at the leaves. Then, DCM returns the correct topology for T with probability $1 - o(1)$ if $k = \Omega(\log^3 n)$ (where the constant factor depends only on f, g), as n tends to $+\infty$.*

Proof: Assume k is as stated above. We apply Theorem 2 and therefore only need to bound $\mathcal{B}$ and $\mathcal{C}$.

For $\mathcal{B}$, from Lemma 2, it suffices to prove that all distances in almost short quartets are approximated within $f/2$ with high probability. Let i, j be a pair of leaves at distance at most $(2\,\Delta(T)+3)g+g$. By lemmas 1 and 4, we have

$$\mathbb{P}\left[\left|\delta^{(2)}_{ij} - \hat{\delta}^{(2)}_{ij}\right| \geq \frac{f}{2}\right] \quad \leq \quad 2\exp\left(-\frac{f^2 k}{8(2\,\Delta(T)+4)^2 \frac{g^2}{f^2} M^2}\right) \leq \frac{1}{\mathrm{poly}(n)},$$

from $\Delta(T) = O(\log n)$, $k = O(\log^3 n)$, and the fact that f, g, M are constants. The notation $\mathrm{poly}(n)$ means $O(n^K)$ for a K *as a large as we need* as long as the constant factor in k is large enough. Since there are at most n^2 such pairs of leaves, we get $\mathcal{B} \leq \frac{1}{\mathrm{poly}(n)}$.

We now bound $\mathcal{C}$. From the definitions of $Q_{\mathrm{s}}(T)$, $Q_{\mathrm{es}}(T)$ and $Q_{\mathrm{as}}(T)$, it suffices to prove that all distances smaller than $3(2\,\Delta(T)+4)g$ are approximated within $g/2$ and that greater distances are approximated within $(2\,\Delta(T)+4)g$. This follows from an argument identical to that in the previous paragraph. Details are omitted. ■

4 Edge Delay Inference

In this section, we show how to estimate the characteristic moments of edge delays. Along with Assumptions 1 and 3, we make the following assumption which is warranted by Theorem 3.

Assumption 4 *We assume that the routing tree was* correctly *estimated. (This is true with high probability by Theorem 3.)*

The general idea is to define additive functions whose edge weights are moments of delays. Then we use the AFI algorithm of Section 2.4 to recover the moments efficiently from the data at the leaves. As it turns out, even moments are rather straightforward to estimate inductively while odd moments are trickier. Also, note that as for the DCM algorithm, the AFI algorithm uses only *short* paths during the estimation process which allows a significant reduction in the sample size.

4.1 Algorithm for Moment Inference

To explain the idea behind our moment inference algorithm, we start with a simple example using the variance. As in Section 3, let $w_e^{(2)} = \mathrm{Var}[d_e]$ for all $e \in E$. Then recall that, for all $a, b \in L$,

$$\delta_{ab}^{(2)} = \mathrm{Var}[D_a - D_b] = \sum_{e \in P_{ab}} \mathrm{Var}[d_e] = W^{(2)}(a,b).$$

Therefore, by estimating $\mathrm{Var}[D_a - D_b]$ from the data at the leaves, we get an estimate of the additive function $W^{(2)}$. And from the AFI algorithm, we recover the $w_e^{(2)}$'s.

More generally, we let

$$w_e^{(j)} = \mathbb{E}\left[\left(\bar{d}_e\right)^j\right],$$

for all $e \in E$ where

$$\bar{d}_e = d_e - \mathbb{E}\left[d_e\right].$$

Also, let

$$W^{(j)}(a,b) = \sum_{e \in P_{ab}} w_e^{(j)},$$

for all $a, b \in L$. Let

$$\overline{D}_a = D_a - \mathbb{E}\left[D_a\right],$$

for all $a \in L$. Again, to obtain $W^{(j)}(a,b)$, we seek to use the quantity

$$\delta_{ab}^{(j)} = \mathbb{E}\left[\left(\overline{D}_a - \overline{D}_b\right)^j\right],$$

which can be estimated from the samples using

$$\hat{\delta}_{ab}^{(j)} = \frac{1}{k}\sum_{i=1}^{k}\left(\left(D_a^i - D_b^i\right) - \hat{\delta}_{ab}^{(1)}\right)^j$$

where

$$\hat{\delta}_{ab}^{(1)} = \frac{1}{k}\sum_{i=1}^{k}\left(D_a^i - D_b^i\right).$$

As Lemma 5 below shows, this can be done inductively. However, the lemma also shows that odd moments have to be treated more carefully.

We first need the following definitions. Let a, b be leaves and $j \in \mathbb{N}$. Recall that γ_{ab} is the most recent common ancestor of a and b in the tree. Suppose we denote $\nu = |P_{ab}|$, $\alpha = |P_{a\gamma_{ab}}|$, and $\beta = |P_{\gamma_{ab}b}|$. Let $[h] = \{0, \ldots, h\}$ for $h \in \mathbb{N}$. Then we let

$$\mathcal{D}_j(a,b) = \left\{(\mathbf{x},\mathbf{y}) \in [j-1]^\alpha \times [j-1]^\beta \;:\; \sum_{i=1}^{\alpha} x_i + \sum_{i=1}^{\beta} y_i = j\right\}.$$

Also, for $(\mathbf{x},\mathbf{y}) \in \mathcal{D}_j(a,b)$, we let

$$\binom{j}{\mathbf{x},\mathbf{y}} = \frac{j!}{\prod_{i=1}^{\alpha} x_i! \prod_{i=1}^{\beta} y_i!}.$$

Finally, we let

$$\mathcal{F}_j(a,b) = \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{D}_j(a,b)} \binom{j}{\mathbf{x},\mathbf{y}} \prod_{i=1}^{\alpha} w_{e_i}^{(x_i)} \prod_{i=1}^{\beta} (-1)^{y_i} w_{f_i}^{(y_i)},$$

where $P_{a\gamma_{ab}} = (e_1, \ldots, e_\alpha)$ and $P_{\gamma_{ab}b} = (f_1, \ldots, f_\beta)$.

Lemma 5 *Let $j \in \mathbb{N}$ and define the function $\mathcal{F}_j : L \times L \to \mathbb{R}$ as above. Then,*

1. *we have for all $a, b \in L$*

$$\delta_{ab}^{(j)} - \mathcal{F}_j(a,b) = \sum_{i=1}^{\alpha} w_{e_i}^{(j)} + (-1)^j \sum_{i=1}^{\beta} w_{f_i}^{(j)}, \tag{2}$$

2. *in particular, if j is even, we have for all $a, b \in L$*

$$\delta_{ab}^{(j)} - \mathcal{F}_j(a,b) = W^{(j)}(a,b). \tag{3}$$

Proof: This follows immediately from a multinomial expansion. ■

Note that $\mathcal{F}_j(a,b)$ depends only on the central moments *of order strictly less than* j of the delays on the edges on P_{ab}. Therefore, if j is even and if we have estimates of all moments of order up to $j-1$, we can estimate $W^{(j)}(a,b)$ by (3). Using the AFI algorithm, we can then get an estimate of moments of order j. However, if j is odd, the coefficient $(-1)^j$ in (2) precludes the use of this procedure. Lemma 6 below shows how to handle this case. But first we note that Lemma 5 above is sufficient for delay distributions symmetric about their mean. Indeed, in that case, all odd central moments are zero and one can use (3) recursively to estimate all (even) characteristic moments. See Figure 6.

Algorithm Symmetric Edge Reconstruction
Input: data $\{D_a^1\}_{a\in L}, \ldots, \{D_a^k\}_{a\in L}$ at the leaves; topology T;
Output: estimated characteristic (even) moments $\hat{w}_e^{(j)}$ for all $e \in E$ and $2 \le j \le J$ even;

- Initialization: set all estimates of odd moments to 0;
- Main Loop: For all $2 \le j \le J$ even,
  - For all $a, b \in L$,
    - Estimate $\hat{\delta}_{ab}^{(j)}$;
    - Estimate $\mathcal{F}_j(a,b)$ with

$$\widehat{\mathcal{F}}_j(a,b) = \sum_{(\mathbf{x},\mathbf{y})\in\mathcal{D}_j(a,b)} \binom{j}{\mathbf{x},\mathbf{y}} \prod_{i=1}^{\alpha} \hat{w}_{e_i}^{(x_i)} \prod_{i=1}^{\beta} (-1)^{y_i} \hat{w}_{f_i}^{(y_i)}.$$

    - Compute

$$\widehat{W}^{(j)}(a,b) = \hat{\delta}_{ab}^{(j)} - \widehat{\mathcal{F}}_j(a,b)$$

  - Use the AFI algorithm on $\widehat{W}^{(j)}(a,b)$ to recover all $\hat{w}_e^{(j)}$'s.

Figure 6: Algorithm Symmetric Edge Reconstruction.

We now tackle odd moments. A proper estimation procedure follows from the next lemma. We first need a definition. For $a, b \in L$, and $1 \le i^* \le \alpha$, we let

$$\mathcal{E}_j^{(1)}(a,b;i^*) = \left\{ (\mathbf{x},\mathbf{y}) \in [j-1]^\alpha \times [j-1]^\beta \,:\, \sum_{i=1}^{\alpha} x_i + \sum_{i=1}^{\beta} y_i = j,\ x_{i^*} \ge 1 \right\}.$$

and

$$\mathcal{G}_j^{(1)}(a,b) = \sum_{i^*=1}^{\alpha} \sum_{(\mathbf{x},\mathbf{y})\in\mathcal{E}_j^{(1)}(a,b;i^*)} x_{i^*} \binom{j-1}{\mathbf{x},\mathbf{y}} \prod_{i=1}^{\alpha} w_{e_i}^{(x_i)} \prod_{i=1}^{\beta} (-1)^{y_i} w_{f_i}^{(y_i)},$$

where we use the notations above Lemma 5. Similarly, for $1 \leq i^* \leq \beta$, we define $\mathcal{E}_j^{(2)}(a, b; i^*)$ and $\mathcal{G}_j^{(2)}(a, b)$ by interchanging the roles of $\mathbf{x}$ and $\mathbf{y}$. Our final definition requires a combinatorial notion. Let a, b, c be any leaves in a *rooted* tree T with root 0 (which is also a leaf). We write $ab|c$ if $ab|c0$ holds in T. Then, for all leaves $a, b, c \neq 0$ with $ab|c$, let

$$\phi_{ab|c}^{(j)} = \mathbb{E}\left[\left(\overline{D}_a - \overline{D}_b\right)^{j-1}\left(\overline{D}_a + \overline{D}_b - 2\overline{D}_c\right)\right].$$

Lemma 6 *Let $j \in \mathbb{N}$ and define the functions $\mathcal{G}_j^{(1)}, \mathcal{G}_j^{(2)} : L \times L \to \mathbb{R}$ as above. Then, using the notations above, we have for all $a, b, c \in L$*

$$W^{(j)}(a, b) = \phi_{ab|c}^{(j)} - \left[\mathcal{G}_j^{(1)}(a, b) + \mathcal{G}_j^{(2)}(a, b)\right]. \tag{4}$$

Proof: We write

$$\mathbb{E}\left[\left(\overline{D}_a - \overline{D}_b\right)^{j-1}\left(\overline{D}_a + \overline{D}_b - 2\overline{D}_c\right)\right] = \mathbb{E}\left[\left(\overline{D}_a - \overline{D}_b\right)^{j-1}\left(\overline{D}_a - \overline{D}_c\right)\right] + \mathbb{E}\left[\left(\overline{D}_a - \overline{D}_b\right)^{j-1}\left(\overline{D}_b - \overline{D}_c\right)\right]$$

Let (as in Figure 7)

$$H_1 = \sum_{e \in P_{a\gamma_{ab}}} \bar{d}_e \qquad H_2 = \sum_{e \in P_{b\gamma_{ab}}} \bar{d}_e \qquad H_3 = \sum_{e \in P_{\gamma_{ac}\gamma_{ab}}} \bar{d}_e \qquad H_4 = \sum_{e \in P_{c\gamma_{ac}}} \bar{d}_e.$$

Note that all these random variables are independent and have 0 mean. Then

$$\begin{aligned}
\mathbb{E}\left[\left(\overline{D}_a - \overline{D}_b\right)^{j-1}\left(\overline{D}_a - \overline{D}_c\right)\right] &= \mathbb{E}\left[(H_1 - H_2)^{j-1}(H_1 + H_3 - H_4)\right] \\
&= \mathbb{E}\left[(H_1 - H_2)^{j-1}(H_1)\right] \\
&= \mathbb{E}\left[\left(\sum_{e \in P_{a\gamma_{ab}}} \bar{d}_e - \sum_{e \in P_{b\gamma_{ab}}} \bar{d}_e\right)^{j-1}\left(\sum_{e \in P_{a\gamma_{ab}}} \bar{d}_e\right)\right] \\
&= \sum_{e \in P_{a\gamma_{ab}}} w_e^{(j)} + \mathcal{G}_j^{(1)}(a, b).
\end{aligned}$$

Similarly,

$$\mathbb{E}\left[\left(\overline{D}_a - \overline{D}_b\right)^{j-1}\left(\overline{D}_b - \overline{D}_c\right)\right] = \sum_{e \in P_{b\gamma_{ab}}} w_e^{(j)} + \mathcal{G}_j^{(2)}(a, b).$$

The result follows. ■

The algorithm for the general case is detailed in Figure 8.

4.2 Analysis of the ER Algorithm

For convenience, we analyze the symmetric case. The general case follows from a similar argument.

We begin with a concentration result for the estimate $\hat{\delta}_{ab}^{(j)}$. For convenience, we assume $M \geq 1$. This can always be obtained by rescaling. The dependence of the bounds on the depth Δ explains the importance of using short paths in the estimation procedures.

Proposition 1 *Let $a, b \in L$ at graph distance less than 2Δ where recall that Δ is the depth of T. Fix $j \in \mathbb{N}$. We have the following:*

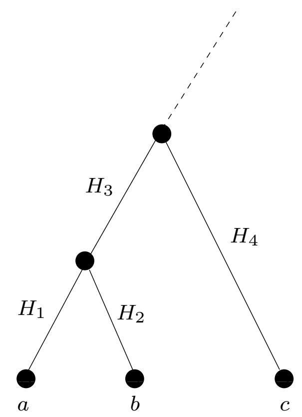

Figure 7: The H_i's are centered sums of delays on the corresponding paths.

Algorithm Edge Reconstruction
Input: data $\{D_a^1\}_{a\in L}, \ldots, \{D_a^k\}_{a\in L}$ at the leaves; topology T;
Output: estimated characteristic moments $\hat{w}_e^{(j)}$ for all $e \in E$ and $2 \leq j \leq J$;

- Initialization: set all estimates of first moments to 0;
- Main Loop: For all $2 \leq j \leq J$,
  - For all $a, b \in L$,
    * Pick the closest leaf c above γ_{ab}
    * Compute $\hat{\phi}_{ab|c}^{(j)}$, the plug-in estimator for $\phi_{ab|c}^{(j)}$;
    * Estimate $\mathcal{G}_j^{(1)}(a,b)$ with

$$\widehat{\mathcal{G}}_j^{(1)}(a,b) = \sum_{i^*=1}^{\alpha} \sum_{(\mathbf{x},\mathbf{y})\in\mathcal{E}_j^{(1)}(a,b;i^*)} x_{i^*} \binom{j-1}{\mathbf{x},\mathbf{y}} \prod_{i=1}^{\alpha} \hat{w}_{e_i}^{(x_i)} \prod_{i=1}^{\beta} (-1)^{y_i} \hat{w}_{f_i}^{(y_i)},$$

      and similarly for $\mathcal{G}_j^{(2)}(a,b)$;
    * Compute

$$\widehat{W}^{(j)}(a,b) = \hat{\phi}_{ab|c}^{(j)} - \left(\widehat{\mathcal{G}}_j^{(1)}(a,b) + \widehat{\mathcal{G}}_j^{(2)}(a,b)\right),$$

  - Use the AFI algorithm on $\widehat{W}^{(j)}(a,b)$ to recover all $\hat{w}_e^{(j)}$'s.

Figure 8: Algorithm Edge Reconstruction.

1. *There exists a constant C such that*

$$\mathbb{P}\left(\left|\hat{\delta}_{ab}^{(j)} - \mathbb{E}\left[\hat{\delta}_{ab}^{(j)}\right]\right| > \lambda\right) \leq 2\exp\left(-\frac{\lambda^2 k}{C\Delta^{2j+1}}\right). \tag{5}$$

2. *There exists a constant C' such that*

$$\mathbb{E}\left(\left|\delta_{ab}^{(1)} - \hat{\delta}_{ab}^{(1)}\right|^j\right) \leq C' \frac{\Delta^j}{k^{j/2}}. \tag{6}$$

3. *Assume that $k \geq \Delta^2$. Then C'' such that*

$$\left|\mathbb{E}\left[\hat{\delta}_{ab}^{(j)}\right] - \delta_{ab}^{(j)}\right| \leq C'' \frac{M^{2j}\Delta^{j+1}}{\sqrt{k}}. \tag{7}$$

4. If, further,

$$C'' \frac{M^{2j}\Delta^{j+1}}{\sqrt{k}} \le \lambda$$

then, we have

$$\mathbb{P}\left[\left|\hat{\delta}_{ab}^{(j)} - \delta_{ab}^{(j)}\right| > 2\lambda\right] \le 2\exp\left(-\frac{\lambda^2 k}{C\Delta^{2j+1}}\right). \tag{8}$$

Proof: 1. We use Azuma's inequality (see Lemma 1). Let

$$\mathcal{K}_i \equiv (D_a^i - D_b^i) - \hat{\delta}_{ab}^{(1)}.$$

Because $|P_{ab}| \le 2\Delta$ and $d_e \in [0, M]$ for all e, it follows that

$$|\mathcal{K}_i| \le 4M\Delta.$$

Then let

$$L \equiv \frac{1}{k}\sum_{i=1}^{k} \mathcal{K}_i^j,$$

and let L' be the same quantity when an arbitrary d_e^i is perturbed by δ. In particular, we assume $|\delta| \le M$. Without loss of generality, say the perturbation is in the first sample. Then,

$$L' = \frac{1}{k}\left(\left(\mathcal{K}_1 + \frac{(k-1)}{k}\delta\right)^j + \sum_{i=2}^{k}\left(\mathcal{K}_i - \frac{\delta}{k}\right)^j\right). \tag{9}$$

Now expanding (9) and noting $\Psi = \max_{1\le i\le j}\binom{j}{i}$, we get

$$\begin{aligned}
|L - L'| &\le \frac{1}{k}\left(j\Psi(4M\Delta)^{j-1}\Delta + (k-1)\left(j\Psi(4M\Delta)^{j-1}\frac{M}{k}\right)\right) \\
&\le C\frac{\Delta^j}{k}
\end{aligned}$$

for some constant C depending on M, j. Noting that L depends on at most $2\Delta k$ random variables d_e^i, we get the result by an application of Azuma's inequality (for a different C).

2. Note that

$$L = \delta_{ab}^{(1)} - \hat{\delta}_{ab}^{(1)},$$

is a $\frac{2M\Delta}{k}$ -Lipschitz function of $\{d_e^i\}_{e\in P_{ab}, i\in[k]}$ thus we have by Azuma's inequality

$$\mathbb{P}\left[\left|\delta_{ab}^{(1)} - \hat{\delta}_{ab}^{(1)}\right| > \lambda\right] \le 2\exp\left(-\frac{k\lambda^2}{8M^2\Delta^2}\right).$$

Now we use the fact that for a positive random variable Y,

$$\mathbb{E}\left[Y^j\right] = j\int_0^{\infty} \lambda^{j-1}\mathbb{P}(Y > \lambda)\mathrm{d}\lambda.$$

If $Y = |\delta_{ab}^{(1)} - \hat{\delta}_{ab}^{(1)}|$ and $\psi = \frac{k}{8M^2\Delta^2}$, we have

$$\mathbb{E}\left[Y^j\right] \le \psi^{-\frac{j}{2}}\int_0^{+\infty} y^{\frac{j}{2}-1}e^{-y}\mathrm{d}y = \left(\frac{8M^2\Delta^2}{k}\right)^{j/2} C'.$$

That proves 2 (for a different C').

3. We have

$$\hat{\delta}^{(j)}_{ab} = \frac{1}{k}\sum_{i=1}^{k}\left((D^i_a - D^i_b) - \hat{\delta}^{(1)}_{ab}\right)^j = \frac{1}{k}\left((D^i_a - D^i_b) - \delta^{(1)}_{ab} + \left(\delta^{(1)}_{ab} - \hat{\delta}^{(1)}_{ab}\right)\right)^j.$$

Now expand using the binomial theorem and take expectations to get

$$\begin{aligned}\left|\mathbb{E}\left[\hat{\delta}^{(j)}_{ab}\right] - \delta^{(j)}_{ab}\right| &\leq \frac{1}{k}\mathbb{E}\left|\sum_{i=1}^{k}\sum_{h=1}^{j}\binom{j}{h}\left(D^i_a - D^i_b - \delta^{(1)}_{ab}\right)^h(\hat{\delta}^{(1)}_{ab} - \delta^{(1)}_{ab})^{j-h}\right| \\ &\leq C''(4M\Delta)^j \max_{h\leq j}\left\{\mathbb{E}\left|\delta^{(1)}_{ab} - \hat{\delta}^{(1)}_{ab}\right|^j\right\}.\end{aligned}$$

4. Now using 1. and 3. and the fact that the number of samples satisfies $k \geq \Delta^2$ we have

$$\max_{j\leq k}\frac{\Delta^j}{k^{j/2}} = \frac{\Delta}{\sqrt{k}}$$

to get the desired result. ■

Choosing a number of sample large enough, we get the following theorem guaranteeing the correctness of SymER with high probability. Recall that $J = O(1)$ and that, in general, $\Delta = O(\log n)$ where n is the number of leaves.

Theorem 4 *Let $\varepsilon > 0$ be arbitrarily small. If $k = \omega(\Delta^{2J^2}\log n)$, then after an application of* SymER*, one has*

$$\mathbb{P}\left[\left|\hat{\delta}^{(j)}_e - \delta^{(j)}_e\right| \leq \varepsilon,\ \forall e \in E,\ \forall 1 \leq j \leq J\right] \geq 1 - o(1), \tag{10}$$

as $n \to +\infty$. The algorithm runs in time $O(\Delta^J n^2)$.

Proof: Let $(a, b) \in L \times L$ be called a *short pair* if a, b are at graph distance at most 2Δ. Denote $\mathcal{S}$ be the set of all short pairs. Let

$$\sigma_j = \max_{(a,b)\in\mathcal{S}}\left|\widehat{W}^{(j)}(a, b) - W^{(j)}(a, b)\right|.$$

It follows immediately that the application of the AFI algorithm implies

$$\max_{e\in E}\left|\hat{w}^{(j)}_e - w^{(j)}_e\right| \leq 4\sigma_j.$$

Therefore, it suffices to prove

$$\max_{1\leq j\leq J}\sigma_j = o(1),$$

with high probability as n tends to $+\infty$. Further, let

$$\tau_j = \max_{(a,b)\in\mathcal{S}}\left|\hat{\delta}^{(j)}_{ab} - \delta^{(j)}_{ab}\right|.$$

Recall that

$$\widehat{W}^{(j)}(a, b) = \hat{\delta}^{(j)}_{ab} - \widehat{\mathcal{F}}_j(a, b)$$

where

$$\widehat{\mathcal{F}}_j(a, b) = \sum_{(\mathbf{x},\mathbf{y})\in\mathcal{D}_j(a,b)}\binom{j}{\mathbf{x}, \mathbf{y}}\prod_{i=1}^{\alpha}\hat{w}^{(x_i)}_{e_i}\prod_{i=1}^{\beta}(-1)^{y_i}\hat{w}^{(y_i)}_{f_i}.$$

Note that $\widehat{\mathcal{F}}_j(a,b)$ has at most Δ^j terms (including the multinomial factor). Therefore,

$$\sigma_j \le \tau_j + \Delta^j(2^j-1)(4\sigma_{j-1})M^{j-1} \le \tau_j + C\Delta^j\sigma_{j-2},$$

for some constant $C > 0$ (independent of j), where we have used $d_e \in [0, M]$ and the fact that odd central moments are 0. If we had a uniform bound τ^* on the τ_j's, then we would have

$$\sigma_j \le \tau^* + C\Delta^J\sigma_{j-2} \le \tau^* C^* \Delta^{J^2/2},$$

for some $C^* > 0$ depending on J, M, where we used $\sigma_2 = \tau_2$.

So it suffices to have $\tau^* = (\omega_n \Delta^{J^2/2})^{-1}$ where $\omega_n \to +\infty$ as $n \to +\infty$ arbitrarily slowly. By the last part of Proposition 1, using a union bound over the $O(n^2)$ short pairs of leaves, it follows that $k = C'\omega_n\Delta^{2J^2}\log n$ samples are enough to guarantee

$$\mathbb{P}\left[|\tau_j| \le (\omega_n\Delta^{J^2/2})^{-1},\ \forall 1 \le j \le J,\ \forall \text{ short pairs } a, b\right] \ge 1 - o(1),$$

for some C' depending on J, M.

As for the computational complexity of the algorithm, assume first that the tree is represented in such a way that finding the set of edges on the path between two leaves a, b at distance $O(\Delta)$ takes time $O(\Delta)$ (this is easy in a rooted tree). Note that for each j, a, b the sum

$$\widehat{\mathcal{F}}_j(a,b) = \sum_{(\mathbf{x},\mathbf{y})\in\mathcal{D}_j(a,b)} \binom{j}{\mathbf{x},\mathbf{y}} \prod_{i=1}^{\alpha} \hat{w}_{e_i}^{(x_i)} \prod_{i=1}^{\beta} (-1)^{y_i}\hat{w}_{f_i}^{(y_i)}.$$

can be computed in time Δ^J. Since there are $O(n^2)$ pairs of leaves and the AFI part of the algorithm takes $O(n^2)$ time, the total complexity is $O(\Delta^J n^2)$. ■

Similarly, in the general case, one can prove:

Theorem 5 *Let $\varepsilon > 0$ be arbitrarily small. If $k = \omega(\Delta^{2J^2}\log n)$, then after an application of* ER*, one has*

$$\mathbb{P}\left[\left|\hat{\delta}_e^{(j)} - \delta_e^{(j)}\right| \le \varepsilon,\ \forall e \in E,\ \forall 1 \le j \le J\right] \ge 1 - o(1), \tag{11}$$

as $n \to +\infty$. The algorithm runs in time $O(\Delta^J n^2)$.

5 Concluding Remarks

1. Throughout, the model was assumed to be static. In real-life networks, characteristics of the network change over time. One could try to adapt our algorithm to a more dynamic setting. See for example [CD+00] for a discussion of temporal issues.

2. We have assumed that delays are finitely supported. This assumption is seemingly inessential. One should be able to replace our boundedness conditions by regularity conditions on the moment generating function for example.

3. We assumed a binary topology. However, note that if the routing tree is given, this assumption is not necessary.

4. We plan to implement our inference algorithm on real datasets. It would be interesting in particular to evaluate the robustness of the algorithm to the partial lack of independence between link delays.

Acknowledgments

We thank Gang Liang, Elchanan Mossel, and Bin Yu for discussions and encouragements. S.R. gratefully acknowledges the partial support of CIPRES, NSERC, FQRNT, and a Loeve Fellowship, and he thanks Martin Nowak and the Program for Evolutionary Dynamics at Harvard University, where part of this work was done. S.B. thanks David Aldous for interesting discussions.

References

[CD+99] R. Caceres, N. Duffield, J. Horowitz, and D. Towsley, Multicast-based inference of network-internal loss characteristics, IEEE Trans. Info. Theory, 45(7):2462–2480, November 1999.

[CD+00] J. Cao, D. Davis, S. Vander Wiel, and B. Yu, Time-varying network tomography: router link data, J. Amer. Statist. Assoc, 95:1063–1075, 2000.

[CC+04] Castro, Rui and Coates, Mark and Liang, Gang and Nowak, Robert and Yu, Bin, Network tomography: recent developments, Statistical Science, 19(3):499–517, 2004.

[Ch96] J. T. Chang, Full Reconstruction of Markov Models on Evolutionary Trees: Identifiability and Consistency, Mathematical Biosciences, 137, 51–73, 1996.

[CGG02] M. Cryan, L. A. Goldberg, and P. W. Goldberg, Evolutionary Trees can be Learned in Polynomial Time in the Two-State General Markov Model, SIAM Journal on Computing, 31(2), 375–397, 2002.

[Du96] R. Durrett, *Probability: Theory and Examples*, Duxbury, 1996.

[ES+97] P. L. Erdos, M. Steel, L. Szekely, and T. Warnow, A Few Logs Suffice to Build (Almost) all Trees (I), Random Structures and Algorithms, 14, 153–184, 1997.

[Fa73] Farris, J.S. 1973. A probability model for inferring evolutionary trees. Systematic Zoology 22, 250–256.

[Fe04] Felsenstein, J. 2004. Inferring Phylogenies. Sinauer Associates, Sunderland.

[GJ76] Garey, M.R. and Johnson, D.S. 1976. Computers and Intractability. A Guide to the Theory of NP-completeness. W.H.Freeman, San Francisco.

[KM+94] M. J. Kearns, Y. Mansour, D. Ron, R. Rubinfeld, R. E. Schapire, and L. Sellie, On the learnability of discrete distributions, STOC 1994.

[LMY06] G. Liang, E. Mossel, and B. Yu, Network Topology Inference through End-to-end measurements, Preprint, 2006.

[LD+02] F. Lo Presti, N. Duffield, J. Horowitz, and D. Towsley, Multicast-based Inference of Network-internal Delay Distributions, IEEE/ACM Transactions on Networking, 10(6):761–775, 2002.

[MR05] E. Mossel and S. Roch, Learning nonsingular phylogenies and hidden Markov models, STOC 2005.

[MR95] R. Motwani and P. Raghavan, Randomized Algorithms, Cambridge University Press, 1995.

[SS03] C. Semple and M. Steel. *Phylogenetics*, volume 22 of *Mathematics and its Applications series*. Oxford University Press, 2003.

[St94] M. Steel, Recovering a tree from the leaf colourations it generates under a Markov model, Appl. Math. Lett. 7(2), 19–24, 1994.

[Va84] L. G. Valiant, A theory of the learnable, Communications of the ACM, 27(11), 1134–1142, 1984.

[Va96] Y. Vardi, Network tomography: estimating source-destination traffic intensities from link data, J. Amer. Stat. Assoc., 91(433):365-377, 1996.