A General Method for the Design of Tree Networks Under Communication Constraints

Alla Tarighati, and Joakim Jaldén
ACCESS Linnaeus Centre, Department of Signal Processing,
KTH Royal Institute of Technology, Stockholm, Sweden
Email: {allat, jalden}@kth.se

***Abstract*—We consider a distributed detection system with communication constraints, where several nodes are arranged in an arbitrary tree topology, under the assumption of conditionally independent observations. We propose a cyclic design procedure using the minimum expected error probability as a design criterion while adopting a person-by-person methodology. We design each node jointly together with the fusion center, while other nodes are kept fixed, and show that the design of each node using the person-by-person methodology is analogous to the design of a network with two nodes, a network which we refer to as the *restricted* model. We further show how the parameters in the restricted model for the design of a node in the tree network can be found in a computationally efficient manner. The proposed numerical methodology can be applied for the design of nodes arranged in arbitrary tree topologies with arbitrary channel rates for the links between nodes and for a general M-ary hypothesis testing problem.**

***Index Terms*—Decentralized detection, Bayesian criterion, tree topology, person-by-person optimization.**

I. Introduction

We consider a distributed, or decentralized, hypothesis testing problem in a general tree network configured as a directed graph, where observations are made at spatially separated nodes. The root of the graph is the fusion center (or FC), and information from nodes propagate toward the FC. If the nodes are able to communicate all their data to the FC, there is no fundamental difference from the centralized case, where the classical solution is to use threshold tests on the likelihood ratios of the received data at the FC. However if there are communication constraints on the links between the nodes, the nodes need to carry out some processing and give a summarized, or quantized, version of their data as output.

The problem of optimal decentralized hypothesis testing has gained noticeable interest over the last 30 years, see for instance [1]–[5] and references therein. A common goal in these references is to find a strategy which optimizes a performance measure, like minimizing the error probability at the FC. However, it is difficult to derive the optimal processing strategies at the nodes in distributed networks, even for small size networks. Therefore most of the works on this topic focus on person-by-person optimization as a practical way for the design of decentralized networks. Using person-by-person optimization, it is guaranteed that the overall performance at the FC is improved (or, at least not worsened) with every iteration of the algorithm. Unfortunately person-by-person optimization gives a necessary, but in general not sufficient, condition for an optimal strategy [4].

While deriving decision function for one node in the person-by-person methodology for the design of nodes in a general tree network (including parallel and tandem networks) all other nodes and the FC are assumed to have already been designed and remain fixed. Focusing on person-by-person optimality, a typical result is that if observations at the nodes are independent conditioned on true hypothesis, likelihood ratio quantizers are person-by-person optimal, while the optimal thresholds in the quantizers are given by the solution of systems of nonlinear equations, with as many variables as the number of thresholds. This however makes the computation of optimal thresholds intractable, even for a moderate size network [4], [6].

The main contribution of this work is to show that – contrary to previous claims – it is possible to under the person-by-person methodology design a distributed detection network arranged in an arbitrary tree topology with a reasonable computational burden. In order to do that we modify the person-by-person methodology for the design of nodes in the tree topology by letting the FC update its decision function in every iteration together with the nodes. In other words, we adopt a person-by-person methodology in which at every iteration each node is designed jointly together with the FC. We further assume that the FC uses the maximum a-posteriori (MAP) rule to make the final decision, which is motivated by the optimality of the MAP rule when the performance criterion is global error probability, or error probability at the FC.

In order to obtain a tractable solution during the design of a node, let say m_0 (together with the FC), all other nodes are modeled using a Markov chain and it will be shown that the design of m_0 is analogous to the design of a special case of a network with only two nodes (which is here called the *restricted* model). Then we will show how the parameters for this restricted model can be found recursively in a computationally efficient way from the original network.

This paper is organized as follows. In Section II we present our model in detail and formulate the problem. In Section III we introduce the restricted model and describe how it can be obtained from the original tree network. We present numerical examples to illustrate the benefit of proposed approach in Section IV and Section V concludes the paper.

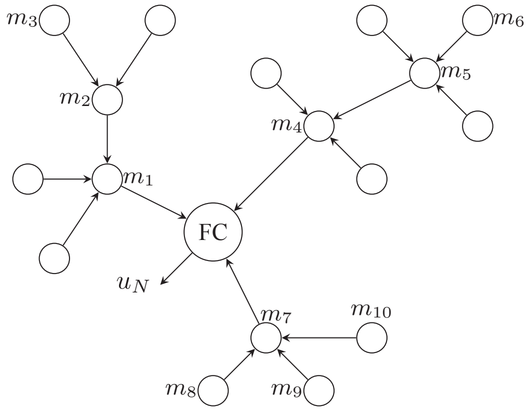

Fig. 1. An example network (observations are not shown).

II. PRELIMINARIES

We describe a tree network by a directed and acyclic graph where the fusion center is the root of the network and information flows from every node on a unique path toward the root. We denote the tree network by $T \triangleq (V, E)$, where $V = \{m_1, m_2, \ldots, m_N\}$ is the set of N nodes and $E = \{e_{i,j}\}$ is the set of directed edges from node m_i to node m_j. Without loss of generality we assume that node m_N is the fusion center (labelled FC in Fig. 1).

We say that node m_i is the *predecessor* of node m_j if there is a directed path from node m_i to node m_j, and say that node m_j is a *successor* of node m_i. Accordingly we say that node m_i is an *immediate* predecessor of node m_j if $e_{i,j} \in E$ (equivalently node m_j is the immediate successor of node m_i). This is exemplified in Fig. 1 where nodes m_3 and m_2 are predecessors of node m_1, while m_2 in an immediate predecessor of node m_1. The set of all immediate predecessors to node m_i is denoted by I_i, and the set of all immediate predecessors to the fusion center is denoted by I_f. For instance, $I_f = \{m_1, m_4, m_7\}$ in Fig. 1. We also define S_i as a set consisting of node m_i and all its successors, excluding the FC. In other words, S_i is the set of all nodes the input messages of node m_i pass through to reach the FC, e.g., $S_6 = \{m_6, m_5, m_4\}$ in the example of Fig. 1. We further define the *last successor* $ls(m_i)$ of node m_i as the last node that the input of node m_i passes through before it reaches the FC, i.e., $ls(m_i) = m_l$, if $m_l \in S_i$ and $m_l \in I_f$. We let $T_i \triangleq (V_i, E_i)$ define the sub-tree of the network with node m_i as its root, where V_i and E_i are the set of nodes and directed edges in sub-tree T_i.

We assume that there are two types of nodes in the tree network: *leaves* and *relays*. A leaf is a node which makes observation and a relay is a node that only receives messages from its immediate predecessors. In Fig. 1 node m_3 and m_2 exemplify a leaf and a relay, respectively. Without loss of generality nodes which both make an observation and receive messages from their immediate predecessors are considered to be relays, since every observation can equivalently be considered as the output of a leaf with output cardinality equal to the cardinality of its observation space. Let C_l be the set of all leaves and C_r be the set of all relays in the network. In a tree network, each leaf $m_i \in C_l$ using its own observation $x_i \in \mathcal{X}_i$ makes a decision $u_i \in \mathcal{M}_i$ and sends it through a rate-constrained channel $(e_{i,j})$ to its immediate successor m_j (which is a relay or FC). Each relay $m_i \in C_r$, using input messages from all of its immediate predecessors I_i, makes a decision $u_i \in \mathcal{M}_i$ and sends it through a rate-constrained channel to its immediate successor. Eventually the fusion center makes the final decision u_N from the set $\{0, 1, \ldots, M-1\}$ in an M-ary hypothesis testing problem. In this paper, we restrict our attention to discrete observation spaces $\mathcal{X}_i$ where $m_i \in C_l$. However we wish to stress that any continuous observation space can be approximated by a discrete observation space, by representing the continuous space by a set of intervals indexed by x_i from the discrete space [7], [8].

The channel between node m_i and its successor is considered to be an error-free but rate-constrained channel with rate R_i bits. The output of node m_i is then from a discrete set $\mathcal{M}_i$ with cardinality $\|\mathcal{M}_i\| = 2^{R_i}$. Without loss of generality we assume that the output of node m_i is from the discrete set $\mathcal{M}_i = \{1, \ldots, 2^{R_i}\}$. In this setup each node is a scalar quantizer which maps its inputs to an output message using a decision function γ_i. A leaf m_l maps its observation x_l to an output message u_l using the function $\gamma_l : \mathcal{X}_l \to \mathcal{M}_l$, i.e.,

$$\gamma_l(x_l) = u_l,$$

whereas a relay m_r maps its input vector containing messages from its immediate predecessors $\{u_i : m_i \in I_r\}$ to an output message u_r using the function $\gamma_r : \mathcal{M}_{I_r} \to \mathcal{M}_r$, i.e.,

$$\gamma_r(\{u_i : m_i \in I_r\}) = u_r.$$

$\mathcal{M}_{I_r}$ is defined as the product of alphabet of immediate predecessors to node m_r. For example, relay m_7 in Fig. 1 has three immediate predecessors, $I_7 = \{m_8, m_9, m_{10}\}$, with decision spaces $\mathcal{M}_8$, $\mathcal{M}_9$ and $\mathcal{M}_{10}$, respectively, and $\mathcal{M}_{I_7} = \mathcal{M}_8 \times \mathcal{M}_9 \times \mathcal{M}_{10}$. We will use the terminology "message" and "index" interchangeably to denote the output of a node in the network.

We assume that the observations at the leaves, conditioned on the hypotheses, are independent. Then acyclicity of the network implies that the inputs to each relay, and also to the FC, are conditionally independent. We further assume that the conditional probability masses of the observations are known and denoted $P_j(x_l) \triangleq P(x_l|H_j), \quad j = 0, 1, \ldots, M-1$, for an M-ary hypothesis testing problem.

In this paper, the objective is to arrive at a simple method for the design of the nodes decision functions, $\gamma_1, \ldots, \gamma_N$, in the tree network, in such a way that the global error probability at the FC is minimized. In order to derive the decision function at a node, we use the person-by-person methodology and assume that all other nodes have already been designed and remain fixed. However, in contrast to previous works, we treat the FC in a different way than the other nodes: the FC decision

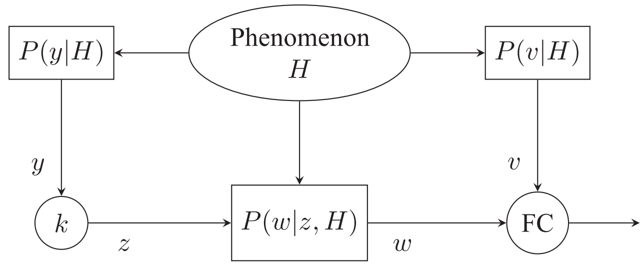

Fig. 2. Restricted model for the design of nodes in tree topology.

function γ_N is always updated together with the node decision function γ_l currently being optimized.

If the decision functions of all the leaves and all the relays are fixed, the optimal fusion center will use the maximum a-posteriori (MAP) rule in order to make the global decision u_N in favor of one of the hypotheses. Defining u_f as the vector containing the messages from the immediate predecessors of the FC, i.e., $u_f \triangleq \{u_i : m_i \in I_f\}$, the FC decides on the hypothesis $H_{\hat{m}}$ if [9]

$$\pi_{\hat{m}} P(u_f|H_{\hat{m}}) = \max_j \left\{\pi_j P(u_f|H_j)\right\}, \tag{1}$$

where $\pi_j \triangleq P(H_j)$ is the a-prior probability of hypothesis H_j, and where $j \in \{0, 1, \ldots, M-1\}$ for the M-ary hypothesis testing problem. The expected minimum error probability in estimating H given an input message vector u_f from the set $\mathcal{M}_{I_f} \triangleq \prod_{m_i \in I_f} \mathcal{M}_i$ is [10]

$$P_{\mathrm{E}} = 1 - \sum_{u_f \in \mathcal{M}_{I_f}} \max_j \left\{\pi_j P(u_f|H_j)\right\}. \tag{2}$$

Knowing the conditional probabilities of the input messages to the FC from its immediate predecessors, the error probability P_{E} at the fusion center can be uniquely computed.

In the next section we will show that the design of a node in the network is analogous to the design of a node (labeled by k) in the restricted model as shown in Fig. 2, where the FC in both networks use the MAP rule (1) as the fusion decision function. We will further show how the conditional probabilities in the restricted model can be recursively computed from the original tree network.

III. Restricted Model

Consider the distributed network with two nodes, k and FC, illustrated in Fig. 2. The fusion center FC using its input messages w and v makes a decision according to the MAP rule (1). Let w and v be from the discrete sets $\mathcal{M}_w$ and $\mathcal{M}_v$, respectively. Conditioned on hypothesis H_j, the input messages y and v are independent with known conditional probability masses $P_j(y)$ and $P_j(v)$, respectively. Node k maps its input y from a discrete set $\mathcal{M}_y$ to an output z from a discrete set $\mathcal{M}_z$ according to a decision function $\gamma_k : \mathcal{M}_y \to \mathcal{M}_z$, i.e., $\gamma_k(y) = z$. The index z then passes through a discrete channel which maps it to the index w with a known transition probability $P(w|z, H)$ that depends on the present hypothesis H. $P(y|H)$ and $P(v|H)$ are probabilistic mapping from the observation space to the discrete sets $\mathcal{M}_y$ and $\mathcal{M}_v$, respectively.

We show next that under the person-by-person methodology, the design of a node, let say m_0, in an arbitrary tree network is analogous to the design of node k in the restricted model for a particular instance of the parameters of the restricted model. To see this, let

$$y \triangleq \begin{cases} x_0 & \text{if } m_0 \in C_l \\ \{u_i : m_i \in I_0\} & \text{if } m_0 \in C_r\,, \end{cases} \tag{3}$$

be the complete input of node m_0 in the original network and let

$$v \triangleq \{u_i : m_i \in I_f,\, m_i \neq ls(m_0)\}, \tag{4}$$

be the complete input of the FC from its immediate predecessors I_f, excluding the node that has m_0 as its predecessor (the immediate predecessor of FC whose the path from m_0 goes through it to reach the FC), and assume that the fusion center in the restricted model uses the MAP rule (1). The conditional PMFs of the inputs to node k and FC in the restricted model are, due to the independency of observations and acyclicity of the network, then given by

$$P_j(y) = \begin{cases} P_j(x_0) & \text{if } m_0 \in C_l \\ \displaystyle\prod_{u_i : m_i \in I_0} P_j(u_i) & \text{if } m_0 \in C_r\,, \end{cases} \tag{5}$$

and

$$P_j(v) = \prod_{\substack{u_i : m_i \in I_f \\ m_i \neq ls(m_0)}} P_j(u_i)\,. \tag{6}$$

The transition probability $P(w|z, H_j)$ is then simply the transition probability from u_0 to u_K, where u_K is the output message of the last successor of m_0, i.e., $m_K \triangleq ls(m_0)$. The key point is that under the person-by-person methodology for the design of node m_0 together with the FC, all other nodes in the network remain fixed. This implies that $P_j(v)$, $P_j(y)$ and $P(w|z, H_j)$ remain fixed and together with the structure of the restricted model, they capture all the important aspects of the joint design problem of m_0 and the FC. In the rest of this section we will show how the transition probabilities $P(w|z, H_j)$ in the restricted model can be found from the original network and describe a recursive method for the computation of the conditional probability masses $P_j(y)$ and $P_j(v)$ in the restricted model from the original tree network.

First, consider an arbitrary node in a tree network, say m_i, and assume that this node has $|I_i| = L_i + 1$ immediate predecessors, containing node m_{i-1}. With a slight abuse of notation we define the set of immediate predecessors of node m_i as $I_i \triangleq \{\tilde{m}_1, \ldots, \tilde{m}_{L_i}, m_{i-1}\}$, as exemplified in Fig. 3. We also define $\tilde{u}_l \in \tilde{\mathcal{M}}_l$ as the output message of node $\tilde{m}_l$, $l = 1, \ldots, L_i$. Conditioned on H_j, each index $u_{i-1} \in \mathcal{M}_{i-1}$ at the input of node m_i is mapped to output

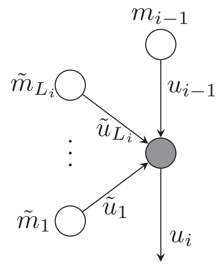

Fig. 3. Node m_i (shaded circle) and its L_i+1 immediate predecessors, I_i.

index $u_i \in \mathcal{M}_i$ according to function

$$\gamma_i(\tilde{u}_1, \ldots, \tilde{u}_{L_i}, u_{i-1}) = u_i\,,$$

with probability $P_j(u_i|u_{i-1}) \triangleq P(u_i|u_{i-1}, H_j)$ which is equal to

$$\begin{aligned} P_j(u_i|u_{i-1}) &= P_j\left(\gamma_i(\tilde{u}_1, \ldots, \tilde{u}_{L_i}, u_{i-1})|u_{i-1}\right) \\ &= \sum_{(\tilde{u}_1,\ldots,\tilde{u}_{L_i})\in\gamma_i^{-1}(u_{i-1},u_i)} P_j(\tilde{u}_1, \ldots, \tilde{u}_{L_i}) \\ &= \sum_{(\tilde{u}_1,\ldots,\tilde{u}_{L_i})\in\gamma_i^{-1}(u_{i-1},u_i)} P_j(\tilde{u}_1)\ldots P_j(\tilde{u}_{L_i})\,, \end{aligned} \tag{7}$$

where $P_j(\tilde{u}_l) \triangleq P(\tilde{u}_l|H_j)$ is the conditional PMF of $\tilde{u}_l$, and where $\gamma_i^{-1}(u_{i-1}, u_i)$ is the set of all input messages $(\tilde{u}_1, \ldots, \tilde{u}_{L_i})$ that satisfy $\gamma_i(\tilde{u}_1, \ldots, \tilde{u}_{L_i}, u_{i-1}) = u_i$. It should be mentioned that, conditioned on the hypothesis, the input messages to node m_i are independent. Now we can state the first important result from (7) as following: consider node m_i and the set of its immediate predecessors, $I_i = \{\tilde{m}_1, \ldots, \tilde{m}_{L_i}, m_{i-1}\}$. Node m_i has a Markovian behavior in the sense that, conditioned on the hypothesis and the input message u_{i-1}, its output message u_i depends only on the inputs from the immediate predecessors $\{\tilde{m}_1, \ldots, \tilde{m}_{L_i}\}$ and not the sequence of messages preceding m_{i-1}. The transition probabilities for this Markov chain is found using (7). The transition probability matrix from input index u_{i-1} to output index u_i, conditioned on hypothesis H_j, is denoted $\mathbf{P}_j^{e_{i-1,i}}$ where the superscript $e_{i-1,i}$ indicates the specific edge that corresponds to the desired input u_{i-1}. $\mathbf{P}_j^{e_{i-1,i}}$ is with size $\|\mathcal{M}_i\| \times \|\mathcal{M}_{i-1}\|$, and its (m,n)th entry is defined as [11]

$$\mathbf{P}_j^{e_{i-1,i}}(m,n) \triangleq P_j(u_i = m|u_{i-1} = n)\,.$$

Consequently, each relay node m_i can be represented by $M|I_i|$ transition probability matrices $\mathbf{P}_j^{e_{k,i}}$, where $m_k \in I_i$, and $j = 0, \ldots, M-1$.

Consider again node m_0 and set $S_0 = \{m_0, m_1, \ldots, m_K\}$. There is exactly one directed path from m_0 to the FC. Assume that node m_1 is the immediate successor of node m_0 and that node m_{l+1} is the immediate successor of node m_l for $l = 1, \ldots, K-1$. The FC is the immediate successor of node m_K. After passing through m_1 to m_K, the output message u_0 of node m_0 from the discrete set $\mathcal{M}_0$ is mapped to an output message u_K of node m_K from the discrete set $\mathcal{M}_K$, which is then used as an input to the FC. The Markov property implies that the transition probability from u_0 to u_K is given by

$$\begin{aligned} P_j(u_K|u_0) &= \sum_{u_1}\ldots\sum_{u_{K-1}} P_j(u_K, u_{K-1}, \ldots, u_1|u_0) \\ &= \sum_{u_1}\ldots\sum_{u_{K-1}}\prod_{i=1}^{K} P_j(u_i|u_{i-1}, \ldots, u_0) \\ &= \sum_{u_1}\ldots\sum_{u_{K-1}}\prod_{i=1}^{K} P_j(u_i|u_{i-1})\,. \end{aligned} \tag{8}$$

Equivalently, in matrix form if we define $\mathbf{P}_j^{0\to K}(m,n) \triangleq P_j(u_K = m|u_0 = n)$, then (8) implies

$$\mathbf{P}_j^{0\to K} = \mathbf{P}_j^{e_{K-1,K}} \times \ldots \times \mathbf{P}_j^{e_{1,2}} \times \mathbf{P}_j^{e_{0,1}}\,. \tag{9}$$

As there is only one directed path from node m_0 to m_K, we omitted the corresponding edge labels in $\mathbf{P}_j^{0\to K}$. Thus, using (9) we can replace all nodes between m_0 and the FC by a single hypothesis dependent transition probability given by $\mathbf{P}_j^{0\to K}$, when designing m_0. During the design of node k in the restricted model (which is equivalent to the design of node m_0 in actual network) every channel transition probability $P_j(w = m|z = n)$ is replaced by the corresponding (m,n)th entry of $\mathbf{P}_j^{0\to K}$.

In forming the restricted model for the design of m_0 in the original network, in addition to channel transition probabilities $P(w|z,H)$, the transition probabilities $P(y|H)$ and $P(v|H)$ should be also determined. The input y to the node k is the complete input messages to node m_0 in the original tree. If node m_0 is a leaf then it only makes observation and $y = x_0$. However, if node m_0 is a relay, then y is a vector containing input messages from its immediate predecessors according to (3) and $P(y|H)$ is defined as (5). In the following we will show how $P_j(u_i)$ in a tree network [corresponding to $P_j(y)$ in the restricted model] can be found in a recursive manner. To this end, consider again node m_0 in the original tree network and its immediate predecessors $m_i \in I_0$. Suppose that node $m_i \in I_0$ receives messages from its L_i immediate predecessors $I_i \triangleq \{\hat{m}_1, \ldots, \hat{m}_{L_i}\}$ and maps its input vector (denoted by $(\hat{u}_1, \ldots, \hat{u}_{L_i})$) to an output message u_i according to a decision function $\gamma_i : \hat{\mathcal{M}}_1 \times \ldots \times \hat{\mathcal{M}}_{L_i} \to \mathcal{M}_i$, i.e.,

$$\gamma_i(\hat{u}_1, \ldots, \hat{u}_{L_i}) = u_i\,.$$

Then the probability masses $P_j(u_i)$ at the output of node m_i are given by

$$\begin{aligned} P_j(u_i) &= P_j\left(\gamma_i(\hat{u}_1, \ldots, \hat{u}_{Li})\right) \\ &= \sum_{(\hat{u}_1,\ldots,\hat{u}_{L_i})\in\gamma_i^{-1}(u_i)} P_j(\hat{u}_1, \ldots, \hat{u}_{L_i}) \\ &= \sum_{(\hat{u}_1,\ldots,\hat{u}_{L_i})\in\gamma_i^{-1}(u_i)} P_j(\hat{u}_1)\ldots P_j(\hat{u}_{L_i})\,, \end{aligned} \tag{10}$$

where $\gamma_i^{-1}(u_i)$ is the set of all input vectors $(\hat{u}_1, \ldots, \hat{u}_{L_i})$ that satisfy $\gamma_i(\hat{u}_1, \ldots, \hat{u}_{L_i}) = u_i$. The last equation is the result of the fact that the inputs to each node in the tree network, conditioned on the hypothesis, are independent.

Equation (10) shows how the probability masses of the output of node u_i can be found based on the probability masses of its inputs and its decision function γ_i. Consider sub-tree T_0 in the network with node m_0 as its root. In the person-by-person methodology used for the design of node m_0 we assume that all other nodes (except the FC) are kept fixed, including all the predecessors of node m_0 in its sub-tree T_0. Starting from the immediate predecessors of m_0 and going backward in the sub-tree, the probability masses of the output of each node can consequently be found based on the probability masses of its input and its decision function [cf. (10)]. Eventually, for a leaf m_l in T_0 the probability masses at the output are given by

$$P_j(u_l) = \sum_{x_l \in \gamma_l^{-1}(u_l)} P_j(x_l)\,. \tag{11}$$

Thus, the required PMFs at m_0 (represented by $P_j(y)$ in the restricted model) can be found by going forward from the leaves in T_0 toward node m_0. Using the same approach, $P_j(v)$ in the restricted model can be found in a recursive manner.

The minimum error probability at the FC in the restricted model (Fig. 2) is a function of the parameters γ_k, $P(y|H)$, $P(v,|H)$ and $P(w|z,H)$, i.e.,

$$P_{\mathrm{E,min}} = \mathcal{F}\big(\gamma_k, P(y|H), P(v,|H), P(w|z,H)\big)\,. \tag{12}$$

Once the parameters in the restricted model are found, the error probability at the FC, when using the MAP criterion, is given by

$$P_{\mathrm{E}} = 1 - \sum_v \sum_w \max_j \left\{\pi_j P_j(v) P_j(w)\right\}, \tag{13}$$

where $P_j(w) \triangleq P(w|H_j)$ is

$$\begin{aligned} P_j(w) &= \sum_{z\in\mathcal{M}_z} P_j(z)P_j(w|z) \\ &= \sum_{z\in\mathcal{M}_z} \sum_{y\in\gamma_k^{-1}(z)} P_j(y)P_j(w|z)\,, \end{aligned} \tag{14}$$

and where $\gamma_k^{-1}(z)$ is the set of all input messages (vectors) y that satisfiy $\gamma_k(y) = z$. Equations (13) and (14) show how the error probability at the FC is affected by the parameters in the restricted channel (especially γ_k).

The goal of this paper is however not to show how γ_k can be designed (together with the FC) in the restricted model. Rather, the goal is just to show how the optimization problem for γ_k can be formulated compactly. However, in [12] a clear-cut guideline for the design of γ_k with a reasonable computational burden is proposed. It is in [12] shown that the design of node k in the restricted model can also be done in a person-by-person manner in terms of the input set; an output index z is assigned to a specific input y, while the assigned indices to other inputs are fixed.

In closing, we emphasize that the proposed method for the design of nodes in the general tree topology (like other methods which use person-by-person methodology) leads only to locally optimum solutions, which also depends on the initialization of the nodes. However, in the next section we show by a numerical example that good performance can nevertheless be obtained.

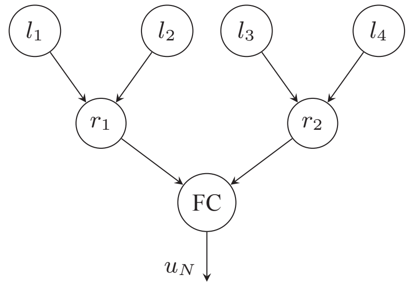

Fig. 4. 2-symmetric 2-uniform tree network.

IV. Examples

In this section we present some results from the application of the proposed method in the design of a tree network. We will consider a 2-symmetric 2-uniform tree network, as defined in [13] given in Fig. 4. The primary reason for choosing such a simple network is to be able to assess the performance of the proposed method through comparison with previous results for tree network and for specific channel rates. We assume the leaves $l_1,\ldots,l_4$ make observations $x_1,\ldots,x_4$, respectively, and the relays r_1, r_2 summarize the messages received from their corresponding immediate predecessors and the FC makes the final decision u_N in favor of one hypothesis. We consider the case of binary hypothesis testing $M = 2$, where real valued observations are, conditioned on the hypothesis, independent and identically distributed. The observation model at each leaf, where each observation consists of an antipodal signal $\pm a$ in unit-variance additive white Gaussian noise $n_i, i = 1,\ldots,4$, is given by

$$\begin{aligned} H_0 &: x_i = -a + n_i \\ H_1 &: x_i = +a + n_i\,. \end{aligned}$$

The per channel signal-to-noise ratio (SNR) is then defined as $\mathcal{E} = |a|^2$. We further assume equally likely hypotheses ($\pi_0 = \pi_1 = 0.5$). Channels between the nodes are considered error-free but rate-constrained where the rate of the leaf-to-relay links are equal to R_l bits, and the rate of the relay-to-FC links are equal to R_r bits. This implies that the leaves' output massages are from the set $\{1,\ldots,2^{R_l}\}$ and the relays' output messages are from the set $\{1,\ldots,2^{R_r}\}$. The FC using the MAP rule (1) makes final decision u_N from the set $\{H_0, H_1\}$.

In our simulations we initialized the relays with random functions, while for $R_l = 1$ we initialized the leaves in all methods with the optimal local decision functions. For $R_l > 1$ we uniformly quantized the two decision regions of the $R_l = 1$ initialization.

A performance comparison of the designed tree networks for different rate pairs (R_l, R_r) and for different per channel

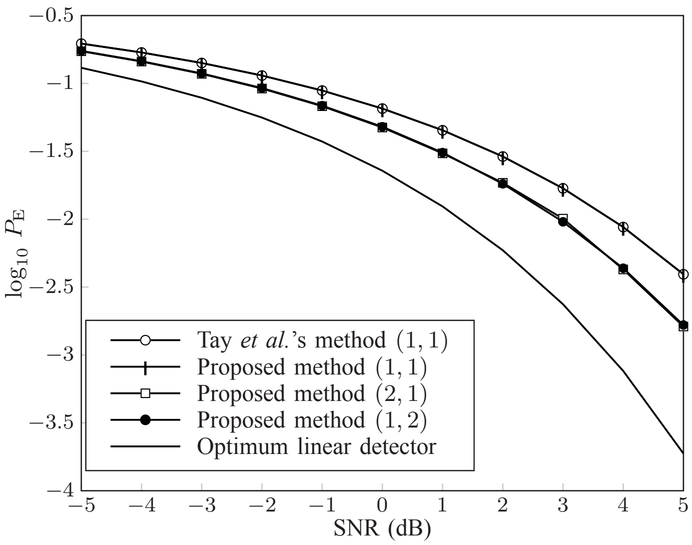

Fig. 5. Error probability performance of the designed 2-symmetric 2-uniform tree network for different rate pairs (R_l, R_r), and for different per channel SNRs.

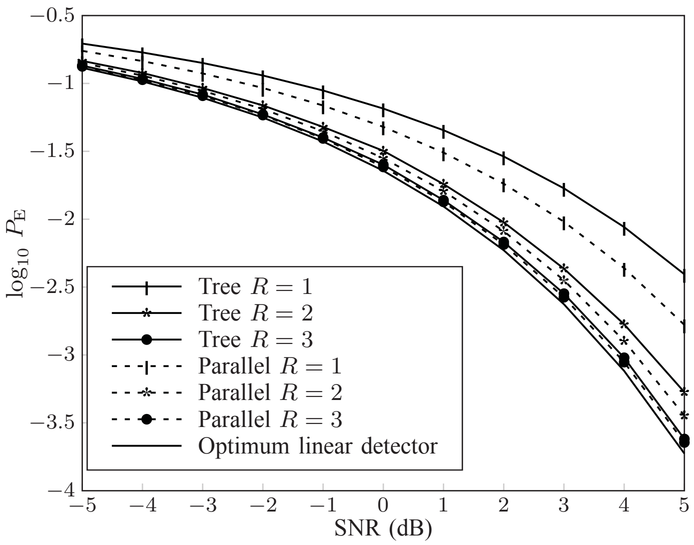

Fig. 6. Comparison of error probability performance of the designed tree network with rate pairs (R, R) and parallel network with rate R.

SNRs is illustrated in Fig. 5. The results of the proposed method are compared to the optimum un-constrained linear detector (which is optimum for this problem) applied to the set of all inputs and results due to Tay *et al.*'s method which leads to the optimal error exponent for an r-symmetric tree [13] for rate pair $(1, 1)$. In that case, the relays use an AND strategy and the leaves have the same threshold γ on their observations. Using an exhaustive search, we found the best γ which minimizes the error probability at the FC, given that the FC uses the MAP rule. The simulation results in Fig. 5 show that for rate pair $(1, 1)$ the proposed method gives the same result as the asymptotically optimum solution, and increasing the rate of the links gives better performance. Also, note that the performance of designed tree networks for rate pair $(1, 2)$ coincides with that for rate pair $(2, 1)$ for equally probable hypotheses. It should however be mentioned that it is not a general result and for other a-prior probability assignments, the resulting curves do not show the same performance.

The proposed method for the design of general tree network can also be used for the design of parallel networks. It is a well known statement that the performance of any optimum tree network is dominated by the performance of an optimum parallel network, for an equal number of observations [6]. Fig. 6 shows the error probability performance of designed tree and parallel networks, where the rate of all the links in the tree network and the rate of all the links in the parallel network are equal to R. As is illustrated in Fig. 6, for the same channel rates, the parallel network outperforms the corresponding tree network. As the rate of the links increases, the performance of the tree network and the parallel network converges to the un-constrained ($R = \infty$) case and the performance of tree network will asymptotically be the same as parallel network.

We can also see that the simulation results of the proposed numerical method are in line with what can be expected in situations where the optimal decision functions are obvious. For example, consider a 2-symmetric 2-uniform tree network with rate pair $(1, 2)$, where the leaves send one-bit messages to the relays and the relays send two-bit messages to the FC. In this case, an optimal relay would simply put the one-bit received messages from its predecessors together and send the resulting two-bit message to the FC, which means the performance of the (optimal) $(1, 2)$ tree network is the same as the performance the (optimal) parallel network with one-bit channel rates. This is consistent with Fig. 5 and Fig. 6, where using the proposed design method yields the same performance in terms of error probability for both cases, which indicates that the proposed method is working as expected.

V. Conclusion

In this paper, we have considered the distributed hypothesis testing problem in a general tree network where the nodes make observations which are, conditioned on the true hypothesis, independent. We have shown that the design of nodes under the person-by-person methodology is analogous to the design of a two-node network, the *restricted* model, in which the decision function can be designed efficiently (cf. [12]). We also have shown how the parameters of the restricted model for the design of a node in the general tree can be formed in a recursive and computationally efficient manner.

Acknowledgment

This work has been supported in part by the ACCESS seed project DeWiNe.

References

[1] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, 1993, pp. 297–344.
[2] Z.-B. Tang, K. Pattipati, and D. Kleinman, "Optimization of detection networks. II. Tree structures," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, no. 1, pp. 211–221, 1993.

[3] A. Pete, K. Pattipati, and D. Kleinman, “Optimization of detection networks with multiple event structures,” *IEEE Trans. Autom. Control*, vol. 39, no. 8, pp. 1702–1707, Aug 1994.
[4] P. K. Varshney, *Distributed Detection and Data Fusion*. Springer-Verlag New York, Inc., 1996.
[5] R. Viswanathan and P. Varshney, “Distributed detection with multiple sensors I. Fundamentals,” *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, 1997.
[6] W. P. Tay, J. Tsitsiklis, and M. Win, “Data fusion trees for detection: Does architecture matter?” *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4155–4168, Sept 2008.
[7] M. Longo, T. Lookabaugh, and R. Gray, “Quantization for decentralized hypothesis testing under communication constraints,” *IEEE Trans. Inf. Theory*, vol. 36, no. 2, pp. 241 –255, Mar. 1990.
[8] A. Tarighati and J. Jaldén, “Bayesian design of decentralized hypothesis testing under communication constraints,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, May 2014.
[9] A. Lapidoth, *A Foundation in Digital Communication*. Cambridge University Press, 2009.
[10] M. Feder and N. Merhav, “Relations between entropy and error probability,” *IEEE Trans. Inf. Theory*, vol. 40, no. 1, pp. 259–266, 1994.
[11] J. R. Norris, *Markov chains*. Cambridge university press, 1998, no. 2008.
[12] A. Tarighati and J. Jaldén, “Bayesian design of tandem networks for distributed detection with multi-bit sensor decisions,” *IEEE Trans. Signal Process.*, submitted for publication, available online: arxiv.org/abs/1403.4155.
[13] W. P. Tay, J. Tsitsiklis, and M. Win, “Bayesian detection in bounded height tree networks,” *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4042–4051, Oct 2009.