

# On Optimal Cooperator Selection Policies for Multi-Hop Ad Hoc Networks

Michele Rossi, *Member, IEEE*, Cristiano Tapparello, and Stefano Tomasin, *Member, IEEE*

**Abstract**—In this paper we consider wireless cooperative multi-hop networks, where nodes that have decoded the message at the previous hop cooperate in the transmission toward the next hop, realizing a distributed space-time coding scheme. Our objective is finding optimal cooperator selection policies for arbitrary topologies with links affected by path loss and multipath fading. To this end, we model the network behavior through a suitable Markov chain and we formulate the cooperator selection process as a stochastic shortest path problem (SSP). Further, we reduce the complexity of the SSP through a novel pruning technique that, starting from the original problem, obtains a reduced Markov chain which is finally embedded into a solver based on focused real time dynamic programming (FRTDP). Our algorithm can find cooperator selection policies for large state spaces and has a bounded (and small) additional cost with respect to that of optimal solutions. Finally, for selected network topologies, we show results which are relevant to the design of practical network protocols and discuss the impact of the set of nodes that are allowed to cooperate at each hop, the optimization criterion and the maximum number of cooperating nodes.

**Index Terms**—Ad hoc wireless networks, automatic repeat request, cooperative communication, MIMO systems, multi-hop communication, optimal policies.

## I. INTRODUCTION

COOPERATION among nodes of a wireless ad hoc network has been shown to be effective in improving the efficiency of resource usage [1], e.g., increasing the network throughput or reducing the energy consumption. In recent years, cooperation has been widely studied both from an information theoretic point of view and from an implementation perspective. A significant amount of work has been done either for the case of two nodes cooperating to transmit two messages to a common destination [2], [3], or for the case of a relay network where transmission from a single source is assisted by one or more cooperative nodes [4], [5]. When multiple nodes are available for cooperation, two major policies can be adopted: a) a single cooperator is selected to aid the transmission of a target node, or b) more nodes cooperate simultaneously with some coordination. The network performance is largely dictated by the cooperator selection, both when only one node cooperates at any given time (see [6] and references therein) and when multiple nodes operate simultaneously [6], [7].

Most of the existing literature is focused on two hop transmission topologies, where the source node transmits

to the relays and then relays forward the message to the final destination. As an example of this, [8] presents a distributed routing protocol that at each hop opportunistically selects the best relay node based on instantaneous channel measurements. However, cooperation can also be applied to multihop transmissions with more than two hops, where at each hop a set of nodes forwards data to another set of nodes. A simple example of multihop transmission is provided in [9], where data are conveyed from the source to the destination of a network by couples of nodes transmitting in cascade. For the case of two transmitting nodes at any hop, in [10], [11], an Alamouti scheme is adopted for a broadband multihop transmission. The outage probability for a fixed rate transmission is analyzed in [12], for a multihop relay network where nodes are organized in clusters and perfectly know the channel within each cluster, while only path-loss and shadowing are known among clusters. Power allocation strategies for multihop multiple relay networks have been investigated in [13]. Under the assumption that relay candidates know the channel conditions, a power efficient multiple relay selection is proposed in [14], while capacity bounds are derived in [15], [16]. In the case of a single node transmitting at any time, power allocation is optimized in [17] and [18]. In [19] the minimum energy consumption is targeted for fixed nodes with no fading, an investigation which has also been conducted in [20], still with perfect channel knowledge at the transmitter and with the constraint that cooperating nodes are along the optimal non-cooperative route. [21] proposes a minimum power cooperative routing algorithm in which, at any time, either a direct transmission or a single relay-aided transmission can occur. Clustered systems are considered in [22] where both the number of nodes per cluster and the clusters are determined to minimize energy consumption in the absence of fading within the cluster, while in [23] the clusters are optimized in order to minimize the total outage probability. In [24] the choice of the number of cooperating transmitters and the cooperation strategy are investigated to exploit the diversity gain for an increase in either the range or the rate of the links or both.

The analysis of this paper extends the work in the literature as it applies to general multi-hop topologies where any number of nodes can cooperate at each hop for the delivery of the message. In detail, we consider a multihop wireless network with arbitrary topology where nodes decode the message and forward it to the next hop until it reaches the destination. In the envisioned scenario nodes cooperate by simultaneously transmitting the message (implementing a distributed space-time coding scheme with decode and forward, DF). The objective of our work is to analytically optimize multi-hop cooperative

Manuscript received October 20, 2009; revised September 16, 2010; accepted November 12, 2010. The associate editor coordinating the review of this paper and approving it for publication was R. Cheng.

The authors are with the Department of Information Engineering, DEI, University of Padova, via Gradenigo 6/B - 35131, Padova, Italy (e-mail: {Michele.Rossi, Cristiano.Tapparello, Stefano.Tomasin}@dei.unipd.it).

Digital Object Identifier 10.1109/TWC.2011.120810.091566

transmission policies (along with their performance in terms of energy expenditure and delay) in the presence of channel impairments and for general network topologies.

Transmission errors depend on path loss and multi-path fading phenomena which dictate the packet error probability for the transmission links. Note that, one may decide upon the correct reception of messages over a given link considering the instantaneous value of its fading process. This would however entail a large complexity for the communicating nodes as they should continuously exchange channel status information. In addition, since our objective in this paper is obtaining *globally optimal* transmission policies, this knowledge should be acquired for all links and for all time instants, which would be impractical. Due to this, we adopt a different model, which takes into account the average channel status for each link, i.e., path loss and fading are translated into outage probabilities. Note that this corresponds to a model with partial channel state information where large scale channel effects (i.e., path loss) are known, whereas small scale fading is modeled for each link through its statistical description.

For the cost model, each transmission has an entangled cost, which is the weighted sum of normalized consumed energy and delay. The goal of our optimization technique is determining which nodes should cooperate at each hop in order to minimize the expected cost over all possible realizations of the cooperative transmission process.

The main contributions of this paper are:

- we model the network behavior through a Markov chain and formulate the multihop cooperator selection process as a stochastic shortest path (SSP) problem. While this SSP can be solved by an iterative procedure according to the framework of real time dynamic programming (RTDP) [25], [26], the complexity of this method grows exponentially with the number of nodes in the network.
- Hence, we derive an iterative solver that operates on a reduced (pruned) Markov chain exploiting an original state pruning technique. This technique is thus integrated with a focused real time dynamic programming (FRTDP) solver [27]. We prove that, by tuning suitable parameters, the algorithm converges with a bounded (and small) additional cost with respect to that of the optimal solution, while considerably reducing the computational complexity.
- The performance bounds obtained in this paper can be useful for the design of practical protocols. In fact, we show results for selected network topologies, discussing the impact of the set of nodes that are allowed to cooperate at each hop, the optimization criterion, i.e., energy *vs* delay minimization and the maximum number of cooperating nodes.

We stress that our analytical tool is meant for centralized and off-line use and we can therefore afford higher complexities than techniques operating in real time. Nevertheless, thanks to our state pruning technique we obtain a problem solver with moderate complexity, which can find optimal policies for large networks in a reasonable time. Note that our objective is obtaining optimal policies along with their performance and not deriving fully implementable solutions. Finally, we observe that our analytical tool works with any scenario

where outage probabilities can be obtained analytically and is thus applicable as well to different network optimization problems.

The paper is organized as follows. In Section II we present the system model. In Section III we formulate the cooperator selection problem using stochastic dynamic programming and prove the technical results related to state pruning. In Section III-E we integrate our pruning technique with FRTDP. In Section IV we prove the effectiveness of our optimization approach and discuss relevant trade-offs in terms of energy, delay and complexity. Section V concludes the paper.

## II. SYSTEM MODEL

Consider a wireless network consisting of a set  $\mathcal{T}$  of static nodes spread out according to any distribution. Among the  $|\mathcal{T}|$  nodes, a source node  $s$  has a message to send to a termination node  $t$ .

### A. Network Model

We deal with the transmission of a message from a source node  $s$  to a termination node  $t$ . Transmissions are performed as follows. At the beginning, the source node  $s$  broadcasts the message, according to the DF scheme; all the nodes that now decode the message (set  $\mathcal{R}_0$ , including  $s$ ) are eligible for transmitting it in the next hop. However, only nodes in a subset  $a_1 \subseteq \mathcal{R}_0$  actually cooperate in the second hop, and they do so simultaneously transmitting the message with a distributed space-time code. The source node  $s$  may be included in  $a_1$  or not, according to the cooperation policy. Decoding and cooperative retransmissions are iterated until the termination node is reached. At the generic hop  $i$ ,  $i = 2, 3, \dots$ , nodes in the set  $a_i$  cooperate (simultaneously transmitting the message), and they are chosen from the set  $\mathcal{R}_{i-1}$  of nodes that know the message at the end of the previous hop. For a failed transmission the packet is discarded. In other words, we consider a distributed automatic repeat request (ARQ), while we leave use of hybrid ARQ (H-ARQ) for future study.

### B. Link Model

Each node is equipped with  $N_A$  antennas, and when nodes in a set  $a$  cooperatively transmit, the total number of transmit antennas is  $N_T = |a|N_A$ . As nodes decode the incoming signals separately, the number of receive antennas for each node is in any case  $N_R = N_A$ . We assume that nodes operate in half-duplex mode and that the same power is used at all transmit antennas. Furthermore, we assume no channel knowledge at the transmitter, i.e., transmit nodes are not aware of position and channel conditions of surrounding nodes.

The transmission channel from nodes in  $a$  to a generic node  $n$ , is described by the  $N_R \times N_T$  matrix  $\mathbf{H}_n(a)$ , having as entry  $[H_n(a)]_{i,j}$ ,  $i = 1, 2, \dots, N_R$ ,  $j = 1, 2, \dots, N_T$ , the channel between the  $j$ th transmit antenna and the  $i$ th receive antenna. For the statistics of  $\mathbf{H}_n(a)$  we consider two wireless propagation phenomena: path-loss and fading. According to this scenario,  $\mathbf{H}_n(a)$  is circular symmetric complex Gaussian with independent entries having zero mean. About the variance, considering a distance  $d_{i,j}^{(n)}$  between transmit and receive

antennas  $j$  and  $i$ , respectively, the power gain due to path loss is  $E[|H_n(a)_{i,j}|^2] = (d_{i,j}^{(n)}/d_0)^{-\nu}$ , where  $d_0$  is the distance at which the average gain is unitary and  $\nu$  is the path-loss exponent. For the sake of a simpler notation, we set  $d_0 = 1$  in the following. Let  $\rho$  be the average signal to noise ratio (SNR), defined as the ratio between the transmit power of a single antenna and the noise power at each receive antenna.

### C. Outage Probability

Cooperative transmission is performed by nodes through a distributed space-time code using  $|a|N_A$  transmit antennas in a synchronous manner. Moreover, in order to improve the transmission reliability, forward error correction (FEC) codes are employed. In order to allow an analysis of the proposed architecture we consider that both the space-time codes and the FEC codes are capacity-achieving, which is a reasonable assumption when advanced space-time coding techniques [28], [29] and low-density parity check codes [30] are employed. In any case, the following analysis provides a bound on the performance that can be obtained with practical systems. We assume that nodes are not aware of the instantaneous channel conditions, but only of their average gain, i.e., the path-loss component. This is realistic when we observe that channel conditions may change, e.g., due to the mobility of surrounding objects. Moreover, referring to our multi-hop route optimization, channel conditions may change as the packets go through the various hops.

As transmit nodes are not aware of instantaneous channel conditions, messages are encoded with a capacity-achieving code having a data rate per unit frequency  $R$ . When the channel capacity, normalized with respect to the bandwidth, is below rate  $R$ , outage occurs. In this case the message is not decoded at the receiving node and is discarded. Let  $C(n, a)$  be the capacity of channel  $\mathbf{H}_n(a)$  with SNR  $\rho$ , normalized with respect to the bandwidth. Then, the outage probability can be computed from the characteristic function (cf) of capacity  $\phi_{C(n,a)}(z)$  as

$$\begin{aligned} p_{\text{out}}(n, a) &= \text{P}[C(n, a) < R] \\ &= \int_{-\infty}^{\infty} \phi_{C(n,a)}(z) \left[ \frac{1 - e^{-j2\pi z R}}{j2\pi z} \right] dz. \end{aligned} \quad (1)$$

In the following we derive the statistics of outage, that will be used to determine the cooperator selection policy in the next section. First, the normalized capacity can be written as a function of the ordered positive eigenvalues of  $\mathbf{H}_n(a)\mathbf{H}_n(a)^H$ ,  $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_{N_{\min}}]$ , with  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{N_{\min}}$  as

$$C(n, a) = \sum_{i=1}^{N_{\min}} \log_2(1 + \rho\lambda_i), \quad (2)$$

where  $N_{\min} = \min\{N_T, N_R\}$ . The cf of the capacity can be then obtained from the statistics of the ordered eigenvalues. In particular, the joint probability density function (pdf) of  $\boldsymbol{\lambda}$ ,  $f(\boldsymbol{\lambda})$  has been studied in [31] for the case  $N_T > N_R$  when the columns are independent and identically distributed while the elements within the same column are correlated. The outage capacity of the corresponding multiple input-multiple output (MIMO) system with correlation at the receive antennas

has been derived in [32]. However, in our scenario, even if we neglect the correlation due to under-spaced antennas, we still have different path-loss coefficients for each link between two nodes. Indeed, this phenomenon can be modeled as a simple correlation among transmit antennas. By indicating with  $[\mathbf{H}_n(a)]_{:,m}$  the  $m$ th column of  $\mathbf{H}_n(a)$ , the correlation matrix among transmit antennas is the diagonal  $N_{\max} \times N_{\max}$  matrix  $\boldsymbol{\Sigma}$  with entries  $\Sigma_m = E[[\mathbf{H}_n(a)]_{:,m}^H [\mathbf{H}_n(a)]_{:,m}]$ ,  $m = 1, 2, \dots, N_{\max}$ . In the general case where the nodes have multiple antennas, the characteristic function of the capacity can be derived following the analyses in [32] and [33]. For the sake of completeness, in Appendix A we derive the simplified expression of the outage probability  $p_{\text{out}}(n, a)$  for the particular case of single antenna nodes, for which we obtain the results in this paper.

### III. OPTIMAL COOPERATOR SELECTION POLICIES

The evolution of our cooperative multihop network can be described by a Markov chain, where the generic state  $x$  is identified by all nodes that have correctly decoded the message so far. The set of all states is instead denoted by  $\mathcal{S}$ . In particular, we are interested in the state in which only node  $s$  knows the message and the *termination states* in which node  $t$  knows the message. Since many states may lead to a correct decoding at node  $t$ , there are in general many termination states and we denote their set by  $\mathcal{D} = \{x : \text{node } t \in \text{state } x\}$ . In what follows, with a slight abuse of notation, we refer to  $s$  and  $t$  as the *starting* and *termination* states, respectively, where  $t$  denotes in this case any state in  $\mathcal{D}$ . We can now address the problem of finding the stochastic shortest path (SSP) from state  $s$  to state  $t$ . At each transmission hop the system is in a generic state  $x$ , representing the nodes that have decoded the message so far. If  $x \neq t$  we must select nodes in  $x$  that will cooperate in the next hop. We denote the set of cooperating nodes as the *action*  $a$ , while set  $\mathcal{A}(x)$  collects all states  $a$  being a subset of nodes of state  $x$ . The dynamics of the network is captured by transition probabilities  $p_{xy}(a)$ ,  $x, y \in \mathcal{S}$  and  $a \in \mathcal{A}(x)$ , describing the probability that nodes in state  $y$  know the message after it has been transmitted by the nodes in set  $a$  when the network was in state  $x$ . From the definition of outage probability (2), we have

$$p_{xy}(a) = \prod_{\substack{n \in \mathcal{T} \text{ st.} \\ n \in y, n \notin x}} (1 - p_{\text{out}}(n, a)) \prod_{\substack{k \in \mathcal{T} \text{ st.} \\ k \notin y}} p_{\text{out}}(k, a). \quad (3)$$

The termination state  $t$  is absorbing, i.e.,  $p_{tt}(a) = 1, \forall a \in \mathcal{A}(t)$ . Note that (3) holds in general for any outage probability, i.e. any channel/transmission model. As an important remark, note that according to our framework the transition probability  $p_{xy}(a)$  depends on starting and ending states  $x$  and  $y$ , i.e., on the nodes having the message prior to and after the transmission as well as on the nodes that transmit (i.e., action  $a$ ). Thus, the transition probabilities for the Markov chain depend on the relative positions of the transmitting nodes and on the statistical description of channel effects. This model can be extended to accommodate the cases where multiple rates and/or powers are exploited at the physical layer. This will only entail the definition of a wider action space (actions will additionally include power and/or rate values), without affecting the state space  $\mathcal{S}$ .

Each transition has also an associated cost. In formulas, a positive cost  $c(x, a, y)$  is incurred when the current state is  $x \in \mathcal{S}$ , action  $a \in \mathcal{A}(x)$  is selected and the system moves to state  $y \in \mathcal{S}$ . In detail,

$$c(x, a, y) = \alpha c_E(x, a, y) + (1 - \alpha) c_D(x, a, y), \quad (4)$$

where  $c_E = |a| + \omega(|y| - |x|)$  (*energy cost*) accounts for the energy spent in transmitting and receiving the message, i.e.,  $|a|$  is the number of cooperating nodes,  $|y| - |x|$  is the additional number of nodes that correctly receive the message and  $\omega \geq 0$  is a parameter taking into account the energy consumed for reception at these nodes.  $c_D(x, a, y) = 1$ ,  $\forall x, y \in \mathcal{S}, a \in \mathcal{A}(x)$  (*delay cost*) accounts for the delay (in number of hops) associated with a path from  $s$  to  $t$ .  $\alpha \in [0, 1]$  is a parameter that we tune to drive the optimization. Note that, since our costs are additive, computing optimal cooperation policies with the cost model of (4) by varying  $\alpha$  in  $[0, 1]$  returns the Pareto efficient frontier in terms of consumed energy *vs* delay, see [34, Section 3.2.4, p. 74]. In addition, observe that  $c_E$  and  $c_D$  are also related to other network parameters. For example, as the delay increases the effective network throughput decreases, since more transmissions are needed to convey the packet to the destination, thus reducing the efficiency of frequency reuse.

The optimization problem  $\mathcal{P} = (\mathcal{S}, \mathcal{A}, p, c, s, t)$  can then be seen as a stochastic shortest path search from state  $s$  to state  $t$  on the modified chain with states  $\mathcal{S}$ , probabilities  $\{p_{xy}(a)\}$ ,  $a \in \mathcal{A}(x)$ , and costs  $c(x, a, y)$ . Our objective is to find, for each possible state  $x \in \mathcal{S}$ , an optimal action  $a^*(x)$  so that the system will reach the termination state  $t$  following the path with minimum average cost. A generic decision policy can be written as  $\pi = \{a(x) : x \in \mathcal{S}\}$ . In general, optimal policies are guaranteed to exist under the following assumptions [35]:

- A1. for any starting state  $x \in \mathcal{S}$ , there exist at least one policy  $\pi$  that eventually reaches the termination state  $t$ , i.e.,  $\lim_{k \rightarrow +\infty} \sum_{r=1}^k p_{xt}^\pi(r) = 1$ , where  $p_{xt}^\pi(r)$  is the probability, averaged over all possible paths followed by  $\pi$ , that the message will reach state  $t$  using this policy in exactly  $r$  transmission hops;
- A2. all costs are positive.

In our scenario both assumptions hold true as costs are positive by definition and we consider strongly connected topologies, i.e., there is a positive probability that any message reaches its destination possibly through multi-hop transmissions.

### A. Optimal Solution

Let  $J(x)$  be the average cost incurred if the current state is  $x$  due to all possible paths, weighted by their probabilities, to reach the final state  $t$ . Note that  $\forall x \in \mathcal{D}$  we have  $J(x) = 0$ . Let us define  $(TJ)(x)$ ,  $x \in \mathcal{S}$ , as

$$(TJ)(x) = \min_{a \in \mathcal{A}(x)} \left[ \sum_{y \in \mathcal{N}(x)} p_{xy}(a) \left( c(x, a, y) + \gamma J(y) \right) \right], \quad (5)$$

where  $\gamma \in [0, 1)$  and  $\mathcal{N}(x)$  is the neighborhood set of  $x$ , containing states  $y \in \mathcal{S}$  such that  $p_{xy}(a) > 0$  for at least one action  $a$ . Let  $J^*(x)$  be the optimal cost-to-go, i.e., the

minimum average cost incurred if the current state is  $x$ , and the optimal policy is followed until we get to the termination state  $t$ . It is known [26] that the optimal policy  $\pi^*$  obeys the following *Bellman's optimality equation*

$$J^*(x) = (TJ^*)(x), \quad x \in \mathcal{S}. \quad (6)$$

In (5) and (6) we consider a discounted version of the SSP problem  $\mathcal{P}$ , since costs incurred in future hops are multiplied by  $\gamma \in [0, 1)$ . Note that  $\gamma = 0$  captures the behavior of a myopic decision maker which takes actions based on the cost incurred in the next hop only (immediate costs), whereas further future costs are ignored. Setting  $\gamma < 1$  is suited to a time varying network, where over a hop the status of closely located terminals remains relatively constant, whereas the status of nodes placed a few hops away will be changed by the time the message will get in their proximity.

From [26, Proposition 2.1.2, p. 91], we know that mapping  $T(\cdot)$  can be iteratively applied, i.e.,  $(T(T^{k-1}J_o))(x) = (T^k J_o)(x)$ , and the following properties hold: 1) *uniqueness*:  $J^*(x)$  is the unique solution of  $J^*(x) = (TJ^*)(x)$ ,  $\forall x \in \mathcal{S}$  and 2) *value iteration*:  $\lim_{k \rightarrow +\infty} (T^k J_o)(x) = J^*(x)$ ,  $\forall x \in \mathcal{S}$  and for any initial guess of the cost-to-go from  $x$ ,  $J_o(x)$ . We stress that these results also hold for  $\gamma = 1$ . From the above properties, iterating the optimality equation over all states in  $\mathcal{S}$  is a practical method to obtain the optimal policies. This technique, however, in our case is impractical due to the large cardinality of  $\mathcal{S}$ . Thus, we advocate the use of advanced RTDP techniques [25], [36], where we decrease the number of states to be visited through a suitable pruning strategy.

### B. Reduced Complexity Techniques

Let  $x \in \mathcal{S}$  be the system state in a generic transmission hop. Our aim is to prune the action set  $\mathcal{A}(x)$  as well as the neighborhood set  $\mathcal{N}(x)$  to the most relevant actions and system transitions in order to reduce the number of states to be visited.

In particular, we consider a new action set  $\mathcal{A}'(x) \subseteq \mathcal{A}(x)$  ( $\mathcal{A}'(x) \neq \emptyset$ ) and a new neighborhood set  $\mathcal{N}'(x) \subseteq \mathcal{N}(x)$  ( $\mathcal{N}'(x) \neq \emptyset$ ). States pruned from  $\mathcal{N}(x)$  are those for which  $p_{xy}(a)$  is small, as detailed below. Similarly, we neglect actions which are unlikely to belong to the optimal policy. Then, indicating with  $\mathbf{J}$  the vector of the current cost estimates, according to (5) the optimal action set for state  $x$  is  $a^* = \operatorname{argmin}_{a \in \mathcal{A}'(x)} Q(x, a, \mathbf{J})$  where

$$Q(x, a, \mathbf{J}) \stackrel{\text{def}}{=} \sum_{y \in \mathcal{N}'(x)} p'_{xy}(a) \left( c(x, a, y) + \gamma J(y) \right), \quad (7)$$

$$x \in \mathcal{S}, a \in \mathcal{A}'(x),$$

$$p'_{xy}(a) = \frac{p_{xy}(a)}{\sum_{y \in \mathcal{N}'(x)} p_{xy}(a)}. \quad (8)$$

In this case (5) becomes

$$(T_p J)(x) = \min_{a \in \mathcal{A}'(x)} Q(x, a, \mathbf{J}), \quad x \in \mathcal{S}, \quad (9)$$

and (6) becomes

$$J_p^*(x) = (T_p J_p^*)(x), \quad x \in \mathcal{S}, \quad (10)$$

where  $J_p^*(x)$  is the optimal cost function for the new Markov chain. The transition probabilities of this new problem  $p'_{xy}(a)$  are normalized so that they still provide a valid probability distribution on  $\mathcal{A}'(x)$ . Note that, since the network is strongly connected, assumption A1) still holds for problem  $\mathcal{P}' = (\mathcal{S}, \mathcal{A}', p', c, s, t)$  as long as  $\mathcal{N}'(x) \neq \emptyset$ ; A2) still holds since costs are unmodified. Consequently, properties of uniqueness and value iteration still hold true for  $\mathcal{P}'$  with the new mapping  $T_p(\cdot)$ . For our optimizations, we assume that at most  $\chi_{\max}$  nodes are allowed to transmit concurrently at each hop, i.e.,  $\max_{a \in \mathcal{A}'(x)} |a| \leq \chi_{\max}, \forall x \in \mathcal{S}$ . The implications of this choice are discussed at the end of Section III-D.

### C. Performance Bounds for State Pruning

We relate  $J^*(x)$  to  $J_p^*(x)$  for arbitrary network topologies through a number of technical results. We define as a proper *upper bound* any function  $\bar{J}(x)$  such that  $\bar{J}(x) \geq J^*(x), \forall x \in \mathcal{S}$ . A valid *lower bound* is defined analogously, i.e.,  $\underline{J}(x) \leq J^*(x), \forall x \in \mathcal{S}$ . Let us also define

$$M(x) \stackrel{\text{def}}{=} \max_{a \in \mathcal{A}'(x)} \left[ \sum_{y \in \mathcal{N}(x) \setminus \mathcal{N}'(x)} p_{xy}(a) \right]. \quad (11)$$

*Lemma 3.1:* Assume that at the generic hop  $i \geq 1$  the system is in state  $x \in \mathcal{S}$ , while  $y \in \mathcal{N}(x)$  is the state at hop  $i + 1$ . Define  $c_{\max} = \alpha(\chi_{\max} + \omega(|\mathcal{T}| - 1)) + 1 - \alpha$  and  $\Delta(x) = M(x)[c_{\max} + \gamma \max_{x \in \mathcal{S}} \bar{J}(x)]$ . For any  $J(x) \leq \bar{J}(x)$ , where  $\bar{J}(x)$  is any proper upper bound for  $\mathcal{P}$ , we have:  $(TJ)(x) \leq (T_p J)(x) + \Delta(x), \forall x \in \mathcal{S}$ .

*Proof:* See the Appendix. ■

*Lemma 3.2:* Let  $x \in \mathcal{S}$  be the system state,  $\eta \in [0, 1)$  be a constant,  $M(x)$  be as defined in (11) with  $M(x) \leq \eta$  and define:

$$g(x, a) \stackrel{\text{def}}{=} \sum_{y \in \mathcal{N}(x)} p_{xy}(a) \left( c(x, a, y) + \gamma J(y) \right), \quad (12)$$

for any  $J(x)$ . If the following equality holds

$$\min_{a \in \mathcal{A}(x)} g(x, a) = \min_{a \in \mathcal{A}'(x)} g(x, a), \forall x \in \mathcal{S} \quad (13)$$

we thus have that  $(TJ)(x) \geq \delta(T_p J)(x)$ , where  $\delta = 1 - \eta$  for all  $x \in \mathcal{S}$ .

*Proof:* See the Appendix. ■

*Remark 3.3:* Lemma 3.2 proves that if, for all states  $x \in \mathcal{S}$ , we obtain set  $\mathcal{A}'(x)$  for problem  $\mathcal{P}'$  by exclusively removing non-optimal actions for problem  $\mathcal{P}$  from  $\mathcal{A}(x)$ , then we can lower bound  $(TJ)(x)$  by  $\delta(T_p J)(x)$ , where  $\delta \in (0, 1]$  depends on the transition probabilities of the pruned states in  $\mathcal{N}(x) \setminus \mathcal{N}'(x)$ .

*Theorem 3.4 (error bounds):* Let  $x \in \mathcal{S}$  be the system state, let  $\Delta \geq 0$  be a constant and assume

$$M(x) \leq \frac{\Delta}{c_{\max} + \gamma \max_{x \in \mathcal{S}} \bar{J}(x)}, \quad \forall x \in \mathcal{S} \quad (14)$$

with  $c_{\max} = \alpha(\chi_{\max} + \omega(|\mathcal{T}| - 1)) + 1 - \alpha$ . For any proper upper bound  $\bar{J}(x)$  for problem  $\mathcal{P}$  we have

(i) For all  $x \in \mathcal{S}$ ,  $J^*(x)$  can be upper bounded as

$$J^*(x) \leq J_p^*(x) + \frac{\Delta}{1 - \gamma}, \quad \forall x \in \mathcal{S}. \quad (15)$$

(ii) In addition, if for any  $x \in \mathcal{S}$  we never remove optimal actions from  $\mathcal{A}(x)$ , i.e., condition (13) of Lemma 3.2 holds and we have

$$\delta \tilde{J}_p^*(x) \leq J^*(x) \leq J_p^*(x) + \frac{\Delta}{1 - \gamma}, \quad \forall x \in \mathcal{S}, \quad (16)$$

where  $\tilde{J}_p^*(x)$  is the optimal cost function for problem  $\mathcal{P}'$  (see (10)) with the modified discount factor  $\tilde{\gamma} = \gamma\delta$  where  $\delta$  is

$$\delta = 1 - \frac{\Delta}{c_{\max} + \gamma \max_{x \in \mathcal{S}} \bar{J}(x)}. \quad (17)$$

*Proof:* See the Appendix. ■

### D. Pruning Criteria

Next, we present an efficient state pruning technique for problem  $\mathcal{P}$  where, for a given sub-optimality threshold  $\Delta/(1 - \gamma)$  and for any state  $x \in \mathcal{S}$ , set  $\mathcal{N}'(x)$  is chosen such that (14) holds, i.e., result (i) of Theorem 3.4 holds.

*Lemma 3.5 (monotonicity):* Let  $i \geq 1$  be the current transmission hop,  $x \in \mathcal{S}$  the corresponding state and  $\mathcal{T}^-(x) = \mathcal{T} \setminus x$  be the set of nodes that still have to decode the message. Let  $\mathcal{A}'(x)$  be the action set for  $\mathcal{P}'$  and state  $x$ . Define  $p_{\text{succ}}(n, a) = 1 - p_{\text{out}}(n, a)$  as the probability that a given node  $n \in \mathcal{T}^-(x)$  correctly decodes the message in hop  $i$ , conditioned on the set of nodes in  $x$  that transmit in hop  $i$ , which we refer to as  $a \in \mathcal{A}'(x)$ . This probability is also conditioned on system topology, channel model and related parameters, see (21). We define  $a_{\max} \stackrel{\text{def}}{=} \operatorname{argmax}_{a \in \mathcal{A}'(x)} |a|$ . It holds

$$p_{\text{succ}}(n, a) \leq p_{\text{succ}}(n, a_{\max}), \forall x \in \mathcal{S}, \quad \forall n \in \mathcal{T}^-(x), \forall a \in \mathcal{A}'(x). \quad (18)$$

*Proof:* The result follows as, for any  $n \in \mathcal{T}^-(x)$ , for any system topology and channel/transmission models, the decoding probability in hop  $i$ ,  $p_{\text{succ}}(n, a)$ , is non-increasing when the number of transmitting nodes goes from  $|a_{\max}|$  to  $|a| < |a_{\max}|$ . ■

Let us now introduce new notation. Given a discount factor  $\gamma$ , set the sub-optimality threshold  $\Delta/(1 - \gamma)$ , for given  $x \in \mathcal{S}$  and  $\mathcal{A}'(x)$ , for all nodes  $n \in \mathcal{T}^-(x)$ , store  $p_{\text{succ}}(n, a_{\max})$  in non-decreasing order into a vector  $\mathbf{v}$ , with entries  $v(j)$ ,  $j = 1, 2, \dots, |\mathcal{T}^-(x)|$ . Let  $m(j) \in \mathcal{T}^-(x)$  be a mapping associating  $v(j)$  to the corresponding node  $n \in \mathcal{T}^-(x)$ . For  $\kappa \geq 1$  define  $\Psi(x)$  as the set of all sequences  $(\xi(1), \xi(2), \dots, \xi(\kappa))$  such that  $1 \leq \sum_{j=1}^{\kappa} \xi(j) \leq \kappa$ , with  $\xi(j) \in \{0, 1\}$ .

*Proposition 3.6 (state pruning):* Consider the following sequential node selection procedure. Initialize set  $\mathcal{V}(x)$  as empty. Evaluate one entry of  $\mathbf{v}$  at a time, let  $\kappa \geq 1$  be the current evaluation step. If  $\kappa < |\mathcal{T}^-(x)| - 1$  and

$$\sum_{\Psi(\kappa)} \prod_{j=1}^{\kappa} v(j)^{\xi(j)} (1 - v(j))^{1 - \xi(j)} \leq \frac{\Delta}{c_{\max} + \gamma \max_{x \in \mathcal{S}} \bar{J}(x)}$$

then 1)  $\kappa \leftarrow \kappa + 1$ , 2) add  $m(\kappa)$  to  $\mathcal{V}(x)$ , i.e.,  $\mathcal{V}(x) \leftarrow \mathcal{V}(x) \cup \{m(\kappa)\}$ , stop otherwise. This procedure returns set

$\mathcal{V}(x)$ . If we prune from  $\mathcal{N}(x)$  all states  $y$  for which at least one of the nodes in set  $\mathcal{V}(x)$  is successful, it holds

$$\begin{aligned} M(x) &= \sum_{\Psi(|\mathcal{V}(x)|)} \prod_{j=1}^{|\mathcal{V}(x)|} v(j)^{\xi(j)} (1-v(j))^{1-\xi(j)} \\ &\leq \frac{\Delta}{c_{\max} + \gamma \max_{x \in \mathcal{S}} \bar{J}(x)}, \forall x \in \mathcal{S}. \end{aligned} \quad (19)$$

*Proof:* See the Appendix. ■

*Remark 3.7 (pruning in practice):* The rationale behind our pruning strategy is that, for any given  $x \in \mathcal{S}$ , there are states  $y \in \mathcal{N}(x)$  having a very small transition probability  $p_{xy}(a)$  for all possible actions  $a$ , i.e., nodes in  $\mathcal{T}^-(x)$  having a small probability of successful decoding in the next hop. Theorem 3.4 can be used as a practical tool to obtain bounds on the optimal policy when solving for  $\mathcal{P}'$  and, at the same time, to keep the error induced by state pruning negligible. Note that the complexity of the procedure in Proposition 3.6 is linear in the size of  $\mathcal{T}^-(x)$ , i.e.,  $O(|\mathcal{T}^-(x)|)$  as it suffices to sequentially evaluate nodes in  $\mathcal{T}^-(x)$ . The lower bound in Theorem 3.4 is generally very close to  $J_p^*(x)$ . This is because in general  $\Delta \ll c_{\max} + \gamma \max_{x \in \mathcal{S}} \bar{J}(x)$ , thus,  $\delta \approx 1$  and  $\tilde{\gamma} \approx \gamma$ . Lastly, we have the further approximation

$$\sum_{\Psi(\kappa)} \prod_{j=1}^{\kappa} v(j)^{\xi(j)} (1-v(j))^{1-\xi(j)} \approx \sum_{j=1}^{\kappa} v(j) \prod_{z=1, \neq j}^{\kappa} (1-v(z)) \quad (20)$$

where we neglected higher order terms, which are  $o\left(\sum_{j=1}^{\kappa} v(j) \prod_{z=1, \neq j}^{\kappa} (1-v(z))\right)$ . The above approximation is very accurate and is preferred in practice as it can be calculated in linear time.

*Remark 3.8 (characterization of set  $\mathcal{A}'(x)$ ):* for each transmission hop we assume that at most  $\chi_{\max}$  nodes are allowed to transmit concurrently. For a given  $\chi_{\max}$ ,  $\mathcal{A}'(x)$  is obtained from  $\mathcal{A}(x)$  by picking the  $\chi_{\max}$  nodes in  $x$  that are closest to  $t$ .<sup>1</sup> This, for non-pathological topologies minimizes the cost (averaged over fading) to reach the destination node  $t$ . Hence, in this way we never remove optimal actions from  $\mathcal{A}(x)$  and, in turn, (16) of Theorem 3.4 holds for the selected  $\mathcal{A}'(x)$ . Of course, optimizing for a given  $\chi_{\max}$  returns the optimal policy  $\pi^*(\chi_{\max})$  over all policies that do not exceed  $\chi_{\max}$  transmitting nodes per hop. As a last remark, observe that picking the nodes that are closest to  $t$  implies perfect knowledge of their geographical position. This is adequate for our analysis, as our objective is obtaining globally optimal policies. Also, in certain networks exploiting geographical routing, such as wireless sensor networks or vehicular networks this assumption may be realistic.

### E. Focused Real Time Dynamic Programming with State Pruning

A well established method to solve a stochastic control problem is the *value iteration* method of Section III-A. This is however infeasible when the state space is very large, as in our case. Focused real time dynamic programming (FRTDP) [27] is a heuristic search algorithm to solve stochastic Markov

<sup>1</sup> $\mathcal{A}'(x)$  coincides with  $\mathcal{A}(x)$  in case the number of nodes in this set is smaller than or equal to  $\chi_{\max}$ .

---

### Algorithm 1: Focused Real Time Dynamic Programming.

---

**Data:** Initial state of the system

**Result:** Optimal policy and relative cost

- 1  $s \leftarrow$  initial state;
  - 2  $\mathbf{D} \leftarrow \mathbf{D}_0$ ;
  - 3 **while**  $(\bar{J}(s) - \underline{J}(s)) > \epsilon$  **do**
  - 4      $(q_{prev}, n_{prev}, q_{curr}, n_{curr}) \leftarrow (0, 0, 0, 0)$ ;
  - 5     **trialRecurse**  $(s, W = I, d = 0)$ ;
  - 6     **if**  $(q_{curr}/n_{curr}) \geq (q_{prev}/n_{prev})$  **then**  $\mathbf{D} \leftarrow k_D \mathbf{D}$ ;
  - 7 **end**
- 

---

### Algorithm 2: trialRecurse(x, W, d). This function recursively implements each trial of FRTDP.

---

- 1  $(\mathcal{N}'(x), \mathcal{A}'(x)) \leftarrow$  Prune  $(x)$ ;
  - 2  $a^* \leftarrow \operatorname{argmin}_{a \in \mathcal{A}'(x)} \{Q(x, a, \underline{\mathbf{J}})\}$ ;
  - 3 **lower**  $\leftarrow Q(x, a^*, \underline{\mathbf{J}})$ ;
  - 4  $\delta \leftarrow |\underline{J}(x) - \text{lower}|$ ;
  - 5  $\underline{J}(x) \leftarrow$  lower;
  - 6  $\bar{J}(x) \leftarrow \min_{a \in \mathcal{A}'(x)} \{Q(x, a, \bar{\mathbf{J}})\}$ ;
  - 7  $y^* \leftarrow \operatorname{argmin}_{y \in \mathcal{N}'(x)} \{\gamma p'_{xy}(a^*) f(y)\}$ ;
  - 8  $f \leftarrow \min_{y \in \mathcal{N}'(x)} \{\gamma p'_{xy}(a^*) f(y)\}$ ;
  - 9  $f(x) \leftarrow \min(|\bar{J}(x) - \underline{J}(x)| - \epsilon/2, f)$ ;
  - 10 **if**  $d > D/k_D$  **then**  
      $(q_{curr}, n_{curr}) \leftarrow (q_{curr} + \delta W, n_{curr} + 1)$ ;
  - 11 **else**  $(q_{prev}, n_{prev}) \leftarrow (q_{prev} + \delta W, n_{prev} + 1)$ ;
  - 12 **if**  $(|\bar{J}(x) - \underline{J}(x)| \leq \epsilon/2)$  **or**  $(d \geq D)$  **then return**;
  - 13 **trialRecurse**  $(y^*, \gamma p'_{xy^*}(a^*) W, d+1)$ ;
  - 14  $a^* \leftarrow \operatorname{argmin}_{a \in \mathcal{A}'(x)} \{Q(x, a, \underline{\mathbf{J}})\}$ ;
  - 15  $\underline{J}(x) \leftarrow Q(x, a^*, \underline{\mathbf{J}})$ ;
  - 16  $\bar{J}(x) \leftarrow \min_{a \in \mathcal{A}'(x)} \{Q(x, a, \bar{\mathbf{J}})\}$ ;
  - 17  $f \leftarrow \min_{y \in \mathcal{N}'(x)} \{\gamma p'_{xy}(a^*) f(y)\}$ ;
  - 18  $f(x) \leftarrow \min(|\bar{J}(x) - \underline{J}(x)| - \epsilon/2, f)$ ;
- 

decision processes having a large number of states. It involves simulated greedy searches within the state space, where cost estimates are updated in a dynamic programming fashion. That is, whenever state  $x$  is reached, its new cost estimate  $J_{\text{new}}(x)$  is updated as:  $J_{\text{new}}(x) \leftarrow Q(x, a^*, \underline{\mathbf{J}})$ , where  $\underline{\mathbf{J}}$  is the vector of the current cost estimates and  $a^*$  is the optimal action based on this vector. We then integrate our pruning techniques of Section III-D into FRTDP to obtain the modified algorithms shown in Algorithms 1–3. The algorithm performs repeated walks through the state space, all starting from  $s$  and terminating in  $t$ . Upper and lower bound estimates of the costs are updated for each visited state  $x$ ; the lower bound  $\underline{J}(x)$  is used to compute optimal policies, whereas the upper bound  $\bar{J}(x)$  is used for the stopping criterion. Among other advantages, empirically, policies obtained from lower bounds tend to perform better [27]. Trials terminate whenever upper and lower bounds of the estimated policy cost from  $s \rightarrow t$  are sufficiently close. **trialRecurse** $(x, W, d)$  is the recursive function implementing each trial, starting from node  $s$  and performing actions until node  $t$  is reached.  $W$  represents the probability (updated recursively) of being in state  $x$ . We modified FRTDP adding the new function

---

**Algorithm 3:** Prune( $x$ ). This function implements the state pruning technique of Section III-D.

---

**input :**  $x \in \mathcal{S}$

**output:** sets  $\mathcal{N}'(x)$  and  $\mathcal{A}'(x)$

---

- 1  $a_{\max} \leftarrow$  take the  $\chi_{\max}$  nodes in  $x$  closest to  $t$ ;
- 2 obtain  $\mathcal{A}'(x)$  from  $a_{\max}$ ;
- 3  $\kappa \leftarrow 1$ ;  $\mathbf{v} \leftarrow \mathbf{0}$ ;  $\mathcal{V}(x) \leftarrow \emptyset$ ;
- 4 **forall** elements  $n$  in set  $\mathcal{T}^-(x)$  **do**
- 5     remove element  $n$  from  $\mathcal{T}^-(x)$ ;
- 6      $v[\kappa] \leftarrow p_{\text{succ}}(n, a_{\max})$ ;
- 7      $\kappa \leftarrow \kappa + 1$ ;
- 8 **end**
- 9 SortNonDecreasingOrder( $\mathbf{v}$ ,  $|\mathcal{T}^-(x)|$ );
- 10  $\kappa \leftarrow 1$ ;  $\overline{M} \leftarrow \frac{\Delta}{c_{\max} + \gamma \max_{x \in \mathcal{S}} \overline{J}(x)}$ ;
- 11  $M(x) \leftarrow 0$ ;
- 12 **repeat**
- 13      $M(x) = M(x)(1 - v(\kappa)) + v(\kappa) \prod_{z=1}^{\kappa-1} (1 - v(z))$ ;
- 14     **if**  $M(x) \leq \overline{M}$  **then**
- 15          $\mathcal{V}(x) \leftarrow \mathcal{V}(x) \cup \{m(\kappa)\}$ ;
- 16          $\kappa \leftarrow \kappa + 1$ ;
- 17     **end**
- 18 **until** ( $M(x) > \overline{M}$ ) or ( $\kappa == |\mathcal{T}^-(x)|$ ) ;
- 19 obtain  $\mathcal{N}'(x)$  from  $x$  and  $\mathcal{V}(x)$ ;
- 20 **return** ( $\mathcal{N}'(x), \mathcal{A}'(x)$ );

---

Prune( $x$ ). In detail, for each state  $x$  in a path, according to Proposition 3.6 we prune the neighborhood set. These states have a small probability of being visited and a negligible impact on the performance. Prune( $x$ ) works as follows: we select the  $\chi_{\max}$  nodes in  $x$  that are closest to node  $t$  (see Remark 3.8) and obtain the action set from these. Hence, we use the pruning algorithm of Proposition 3.6 considering all nodes  $n \in \mathcal{T}$  that have not yet decoded the message and pruning those with smaller probability of being reached at the next transmission. In particular, we add new nodes until  $M(x) > \overline{M}$ , as dictated by Proposition 3.6. In addition, for  $M(x)$  we consider the approximation of Remark 3.7.  $\mathcal{N}'(x)$ , i.e., the neighborhood set, is finally obtained from the set of selected nodes. The remainder of `trialRecurse( $s, W, d$ )` is as specified in [27]. In short, the new optimal action  $a^*$  for state  $x$  is selected according to the DP optimal equation using the latest cost estimates  $\underline{\mathbf{J}}$ . Upper and lower bounds are updated according to the optimality equation as  $(\underline{\mathbf{J}}_{\text{new}}(x), \overline{\mathbf{J}}_{\text{new}}(x)) \leftarrow (Q(x, a^*, \underline{\mathbf{J}}), \min_{a \in \mathcal{A}'(x)} Q(x, a, \overline{\mathbf{J}}))$  (lines 3 and 6). The next state to visit,  $y^*$ , is picked by maximizing the *occupancy times excess uncertainty* metric, i.e.,  $W(y)\Delta(y)$ , where  $W(y)$  is the average probability of visiting the state and  $\Delta(y) = (\overline{\mathbf{J}}(y) - \underline{\mathbf{J}}(y)) - \epsilon/2$ , represents the accuracy of its cost estimates. This is implemented as in the original algorithm [27] through a priority function  $f(y)$ , which is recursively computed for each state. The current trial terminates when the final state is reached (note that  $\overline{\mathbf{J}}(t) = \underline{\mathbf{J}}(t) = 0$ ), when a state  $x$  having estimates sufficiently close to the optimum cost is reached, i.e.,  $\overline{\mathbf{J}}(x) - \underline{\mathbf{J}}(x) \leq \epsilon/2$

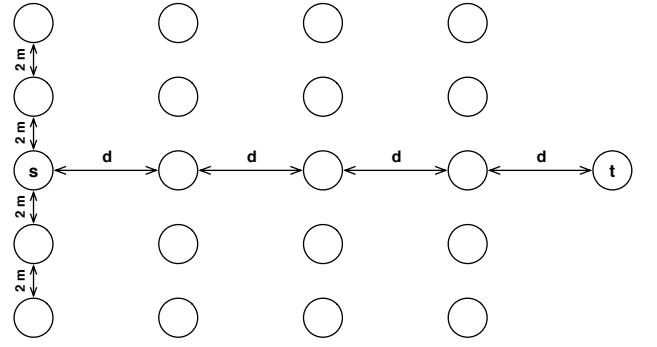


Fig. 1. Network topology for scenario A (4 columns, 21 nodes).

or when the current path is longer than  $D$ . Whenever the current trial terminates, optimal actions, lower and upper bounds and priority are updated on the way back along the traversed path from  $s \rightarrow t$  (lines 14 – 18). For the check on the path length, poor outcome selection early in a trial could lead to traversing a large number of irrelevant states, which take a long time to escape. The check on the maximum hop length implements the *adaptive maximum depth* (AMD) trial termination of [27], which solves this problem by cutting excessively long paths.

#### IV. NUMERICAL RESULTS

In this section we provide an example application of the proposed optimization techniques for cooperator selection policies, showing numerical results and obtaining insights for possible low-complexity implementations. We consider the network topology of Fig. 1, where a source node  $s$  transmits a message to a destination node  $t$  and the remaining nodes are available for cooperation. All nodes except the destination are organized in a number of columns, each comprising five nodes. The inter-column distance  $d$  is picked in the range  $45 \div 80$  m, while the distance between two adjacent nodes in a column is 2 m. The path-loss exponent is  $\nu = 3.5$  and the reference distance is  $d_0 = 1$  m. In what follows, two network scenarios are considered: scenario A) is the topology of Fig. 1 with four columns and 21 nodes and scenario B) where we extended the number of columns to eight for a total of 41 nodes. The transmit data rate  $R$  and the average SNR are set in order to obtain, for a single active link, an outage probability of 0.2 at a distance of 30 m, while transmissions among adjacent nodes in a column have average outage probability  $2 \cdot 10^{-5}$ . Each node is equipped with a single antenna (i.e.,  $N_A = 1$ ). We evaluated the performance for various values of  $\omega \geq 0$  and we observed a straightforward behavior for the optimized cost, which increases linearly with increasing  $\omega$ . Therefore, in what follows we only discuss the case  $\omega = 0$ .

Our optimization is driven by the cost model of (4), which returns the cost over a single transmission hop taking into account a weighted sum of energy ( $c_E$ ) and delay ( $c_D$ ), where  $\alpha \in [0, 1]$  is the weighting factor. Analogously, the overall cooperator selection policy is characterized by the two costs  $C_E$  and  $C_D$  that are respectively the expected normalized energy and the expected total delay of the optimal cooperator

TABLE I  
PERFORMANCE OF THE MODIFIED FRTDP OPTIMIZER AS A FUNCTION OF  $\Delta$ .

$\Delta$	Scenario A: 21 nodes network				Scenario B: 41 nodes network		
	Visited States	Failure Prob. $p_{\text{fail}}(\Delta)$	$\Delta C$ [%] Actual	$\Delta C$ [%] Predicted	Visited States	Failure Prob. $p_{\text{fail}}(\Delta)$	$\Delta C$ [%] Predicted
0	$1.3 \cdot 10^9$	0	0.00	—	—	—	—
0.001	$2.2 \cdot 10^6$	$9.0 \cdot 10^{-6}$	0.00	0.76	$3.3 \cdot 10^8$	$5.0 \cdot 10^{-6}$	0.45
0.005	$2.2 \cdot 10^6$	$9.0 \cdot 10^{-6}$	0.00	3.82	$3.3 \cdot 10^8$	$5.0 \cdot 10^{-6}$	2.24
0.01	$2.2 \cdot 10^6$	$9.0 \cdot 10^{-6}$	0.00	7.64	$3.3 \cdot 10^8$	$5.0 \cdot 10^{-6}$	4.49
0.05	$2.2 \cdot 10^6$	$9.0 \cdot 10^{-6}$	0.00	38.18	$3.3 \cdot 10^8$	$5.0 \cdot 10^{-6}$	22.45
0.1	$2.2 \cdot 10^6$	$9.0 \cdot 10^{-6}$	0.00	76.35	$3.3 \cdot 10^8$	$5.0 \cdot 10^{-6}$	44.89
0.5	$2.2 \cdot 10^6$	$9.0 \cdot 10^{-6}$	0.00	381.77	$3.3 \cdot 10^8$	$5.0 \cdot 10^{-6}$	224.46
1	$2.2 \cdot 10^6$	$9.0 \cdot 10^{-6}$	0.00	763.53	$3.3 \cdot 10^8$	$5.0 \cdot 10^{-6}$	448.92
5	$1.5 \cdot 10^6$	$5.9 \cdot 10^{-5}$	5.13	3631.35	$2.9 \cdot 10^8$	$1.3 \cdot 10^{-5}$	2175.95
10	$7.0 \cdot 10^5$	$3.6 \cdot 10^{-4}$	13.88	6704.66	$1.6 \cdot 10^8$	$2.5 \cdot 10^{-5}$	4293.26

selection policy used to route packets from  $s$  to  $t$ .<sup>2</sup> Picking  $\alpha = 1$  returns optimal policies in terms of  $C_E$ , while  $C_D$  is ignored. Conversely,  $\alpha = 0$  returns optimal policies in terms of  $C_D$ , ignoring  $C_E$ . Intermediate values of  $\alpha$  lead to suitable trade-offs between energy and delay. In what follows, optimal policies are obtained setting  $\gamma = 0.99$ , which is adequate for static networks, see Section III-A. For our FRTDP technique we set  $\epsilon = 10^{-3}$ ,  $\underline{J}(x) = 0, \forall x \in \mathcal{S}$ ,  $K_D = 1.1$  and  $D_0 = 10$ . For the upper bound  $\bar{J}(x)$  we considered  $\Delta = 0.001$  and a large initial  $\bar{J}(x) = 100, \forall x \in \mathcal{S} \setminus t$ . We obtained a first policy and the corresponding cost  $J_p^*(x), \forall x \in \mathcal{S} \setminus t$  and thus set  $\bar{J}(x) \leftarrow J_p^*(x) + \Delta/(1 - \gamma)$ .

The choice of parameter  $\Delta$  is guided by the trade-off between sub-optimality of the policy and its computational complexity. In detail, when  $\Delta = 0$  our FRTDP optimizer does not cut any state and finds optimal policies as done by RTDP [25], where  $J^*(s)$  is their cost. When  $\Delta > 0$  some states are instead pruned according to our techniques of Section III-D and our optimizer returns an approximation of the optimal policy, with cost  $J_p^*(s)$ . Note that setting  $\Delta > 0$  for any given state  $x$  reduces the number of neighboring states  $y$  and, to a lesser extent, also reduces the number of states for which the policy is computed, as states hit with small probability are not considered. As a consequence, the optimal policy is not calculated for these states. Table I shows the performance of our FRTDP algorithm as a function of  $\Delta$ , for  $d = 60$  m and  $\alpha = 1$  in terms of 1) computational complexity, expressed in terms of number of visited states, 2) estimated failure probability  $p_{\text{fail}}(\Delta)$ , i.e., the probability of hitting a state for which our optimizer did not calculate optimal actions, 3) actual cost difference with respect to RTDP, i.e.,  $100|J_p^*(s) - J^*(s)|/J^*(s)$  and 4) the maximum cost difference between  $J_p^*(s)$  and  $J^*(s)$ , as predicted by Theorem 3.4, i.e.,  $100\Delta/[J_p^*(s)(1 - \gamma)]$ . We discuss the results for scenario A first. In this case, even a small  $\Delta = 0.001$  suffices to dramatically reduce the number of visited states, which drops from  $1.3 \cdot 10^9$  to  $2.2 \cdot 10^6$ . For this  $\Delta$ , our bounds would predict a maximum additional cost that is just 0.76% larger than  $J_p^*(s)$ . We note that, for this specific network topology, the solver performance is better than that predicted by the bound. Also, there is a threshold effect on the number of pruned states

<sup>2</sup>These costs are the average of the costs obtained over all possible realizations of the cooperators selection process from  $s$  to  $t$  when the optimal policy is adopted.

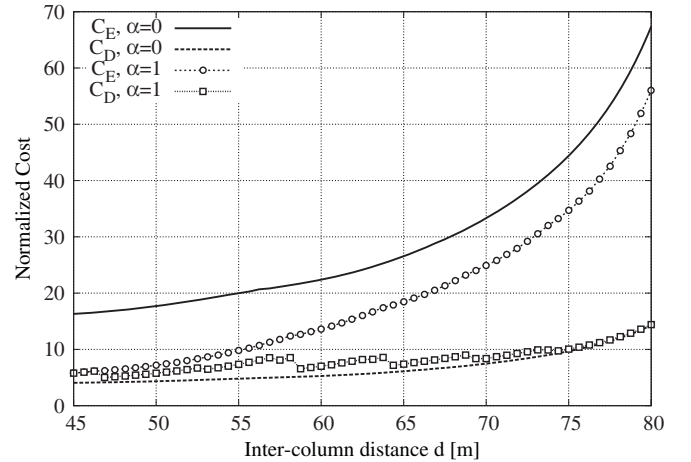


Fig. 2. Normalized costs  $C_E$  and  $C_D$  as a function of  $d$  for  $\alpha = 0$  and  $\alpha = 1$ .  $C_E$  and  $C_D$  are normalized with respect to the energy spent to transmit a single packet and the message transmission delay, respectively. Other optimization parameters are:  $\omega = 0$ ,  $\gamma = 0.99$ ,  $\Delta = 0.001$  and  $\chi_{\text{max}} = 5$ .

for increasing  $\Delta$ , which is due to the specific topology under consideration. For scenario B (41 nodes) the solver fails to obtain policies for  $\Delta = 0$ , due to the excessively large number of states. However,  $\Delta = 0.001$  already provides cooperation policies having a small bounded additional cost with respect to the unknown optimal performance. We shall observe that the bounds of Theorem 3.4 are asymptotically tight, i.e., they become more accurate as the path length increases. Finally, we note that  $p_{\text{fail}}(\Delta)$  is very small in all cases. These results show the effectiveness of our technique, which makes it possible to find quasi-optimal policies for large networks at a reduced complexity.

Fig. 2 shows  $C_E$  and  $C_D$  as a function of the inter-column distance  $d$  for  $\alpha = 1$  (minimum energy) and  $\alpha = 0$  (minimum delay). Costs are normalized with respect to the cost incurred for a single packet transmission. We observe that for  $\alpha = 1$  the energy cost  $C_E$  increases smoothly with  $d$ , while for  $\alpha = 0$  the delay cost  $C_D$  increases smoothly with time, since a larger distance  $d$  between columns yields higher outage probabilities which, in turn, lead to longer transmission delays over single hops. In the figure, we also show non targeted costs, i.e.,  $C_D$  when the optimization objective corresponds to



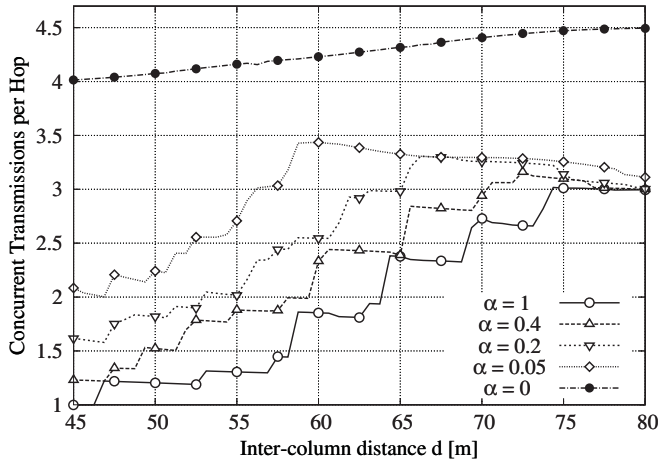


Fig. 3. Average number of cooperating nodes as a function of  $d$  for different values of  $\alpha$ . Other optimization parameters are:  $\omega = 0$ ,  $\gamma = 0.99$ ,  $\Delta = 0.001$  and  $\chi_{\max} = 5$ .

minimizing the energy consumption ( $\alpha = 1$ ) and  $C_E$  when the objective is the minimization of the delay ( $\alpha = 0$ ). Non targeted costs generally increase with increasing  $d$ . However, the corresponding curves have an irregular behavior as in some cases non-targeted costs decrease with the inter-column distance. This is due to the fact that the optimization is performed on a discrete set of policies. For example, when the target cost is  $C_E$  and  $d$  is slightly increased, to counteract the increased outage probability cooperation may start earlier and involve a larger number of nodes. The effect of this is twofold: 1)  $C_E$  is kept as small as possible and 2) the delay is decreased as more nodes transmit at each hop. Overall, the result is a slight increase in  $C_E$  (thus the smooth curve for  $C_E$ ) together with a sudden drop of  $C_D$  due to the reduced number of hops (thus the irregular curve for  $C_D$ ).

To better understand the impact of cooperation in a multi-hop scenario with optimized cooperator selection policies in Fig. 3 we show the average number of nodes that transmit simultaneously, as a function of  $d$  and for various values of  $\alpha$ . Note that, when the objective is to minimize the delay, optimal policies tend to maximize the number of cooperating nodes per hop as the cost in this case is solely given by the number of hops traveled by the message, irrespective of the number of transmitting nodes within each hop. When minimizing energy, the cost also depends on the number of cooperating nodes within each hop and, as a consequence, the optimal number of cooperating nodes per hop is smaller. Also in this case we observe an irregular behavior of the curves, which can be explained considering the discrete nature of the problem. In general, the average number of simultaneous transmissions decreases with increasing  $d$ , as outages occur more often and, in such cases, fewer nodes are available for transmission. However, this is true until the cooperation policy changes, at which point cooperation is forced among a larger number of nodes in order to minimize the targeted cost. Notably, we can see a close relationship between Fig. 3 and the non-targeted costs of Fig. 2: for example, when  $\alpha = 1$  at 58.75 m the average number of simultaneous transmissions increases from 1.45 to 1.86 (Fig. 3) and, at the same time,  $C_D$  drops from 8.53 to 6.56 (Fig. 2). This corresponds to a forced earlier

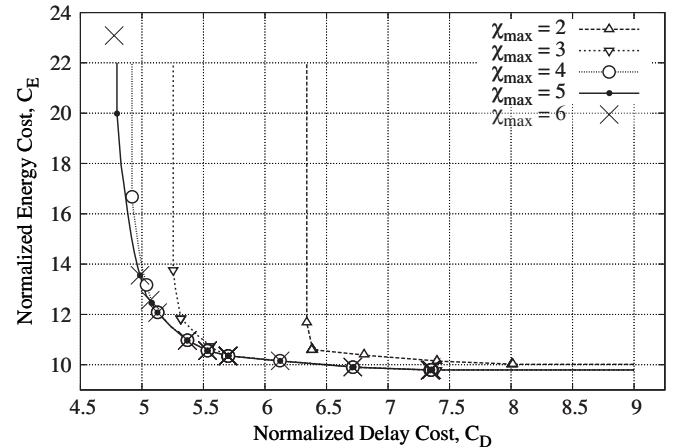


Fig. 4.  $C_E$  vs  $C_D$  for several values of  $\chi_{\max}$ . The curves are obtained for  $d = 55$  m, varying  $\alpha \in [0, 1]$ . Other optimization parameters are:  $\omega = 0$ ,  $\gamma = 0.99$ ,  $\Delta = 0.001$  and  $\chi_{\max} \in \{1, 2, 3, 4, 5, 6\}$ .

cooperation among nodes which causes an increase in  $C_E$  as well as a subsequent reduction in the number of hops. Fig. 3 also confirms that cooperation is advantageous when multihop is considered and minimization of energy consumption rather than delay or rate are targeted. As an example, for  $\alpha = 1$  the average number of simultaneous transmissions goes from 20% (i.e., 1 transmitting node for  $d = 45$  m over a maximum of  $\chi_{\max} = 5$  cooperating nodes) to 60% (i.e., 3 cooperating nodes over  $\chi_{\max} = 5$ ). In addition, we observe that the average number of cooperating nodes is small with respect to the total number of nodes in the network, thus it is meaningful to impose a maximum  $\chi_{\max} \ll |\mathcal{T}|$  on the number of cooperating nodes, as discussed in Section III-D.

Fig. 4 shows the trade-off between  $C_E$  and  $C_D$  as a function of  $\chi_{\max}$  for an inter-column distance of  $d = 55$  m. The curves are obtained by varying the weighting factor  $\alpha$  in  $[0, 1]$  and provide the delay-energy achievable regions, as for a given  $\chi_{\max}$  no policy can obtain a trade-off point situated below the corresponding optimal curve, while any point above the optimal curve is achieved by a suitable suboptimal policy. However, this figure provides even further insights on possible implementations of optimal policies. In fact, for  $\alpha > 0$  setting  $\chi_{\max} = 5$  already provides most of the benefits of optimal policies in the unconstrained optimization case ( $\chi_{\max} = +\infty$ ). This means that complexity of both policy optimization and network coordination can be reduced at almost no expense in terms of performance. On the other side, being too restrictive on the number of cooperators yields some performance loss, as for example allowing at most 2 cooperating nodes leads to a delay increase of about 20% and to an increase of energy consumption of about 10%. Note that if cooperation is not allowed (i.e.,  $\chi_{\max} = 1$ ) delay and energy consumption are centered around point  $(x, y) = (11.8, 11.8)$  (out of range in the figure). Therefore, even a minimum level of cooperation, i.e., between two nodes ( $\chi_{\max} = 2$ ), provides a substantial performance advantage. We finally observe that, if at every hop the maximum admissible number of nodes cooperate, we obtain the delay optimal policy ( $\alpha = 0$ ). This however comes at the expense of a high energy consumption. A more judicious choice leads to considerable advantages, e.g., a delay just 4%

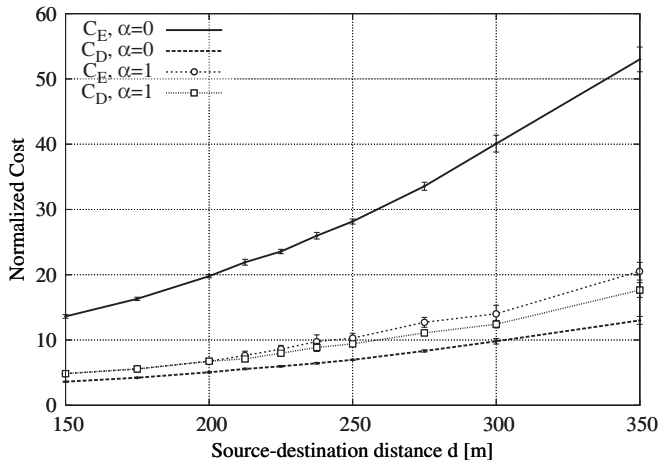


Fig. 5. Random network: Normalized costs  $C_E$  and  $C_D$  as a function of  $d$  for  $\alpha = 0$  and  $\alpha = 1$ .

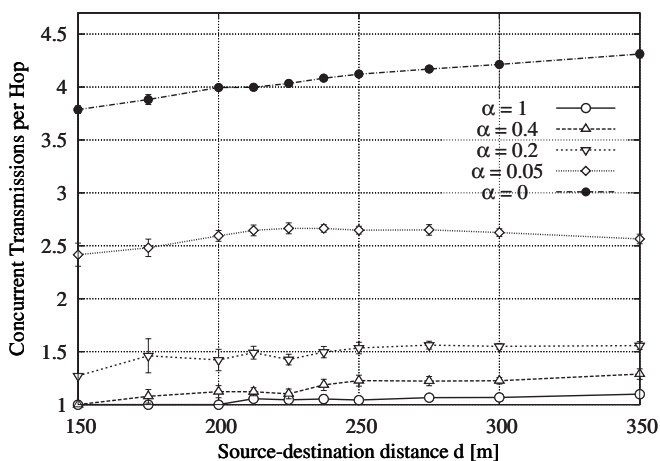


Fig. 6. Random network: average number of cooperating nodes *vs*  $d$  for different values of  $\alpha$ .

over the minimum provides a drop of consumed energy by about 30%.

Finally, we considered random networks with 21 nodes placed within a rectangular simulation area of  $50 \times d$  square meters as follows: source and destination are respectively positioned in the middle of the two opposite 50 m long sides, whereas the remaining nodes are randomly placed within the area. Optimization parameters are:  $\omega = 0$ ,  $\gamma = 0.99$ ,  $\Delta = 0.001$  and  $\chi_{\max} = 5$ . In the graphs, vertical bars are used to show 95% confidence intervals. For the random scenario, Fig. 5 shows the normalized costs  $C_E$  and  $C_D$  as a function of  $d$  for  $\alpha \in \{0, 1\}$ . Fig. 6 instead shows the average number of cooperating nodes for different values of  $\alpha$ . The considerations for these graphs are similar to those made for the previous plots. As in the previous results, optimal delay strategies ( $\alpha = 0$ ) entail the largest number of cooperating nodes. However, differently from the previous results, cooperation is almost absent when the objective is energy minimization ( $\alpha = 1$ ). Also, we note that  $d$  has a smaller impact on the optimal number of cooperating nodes, which is almost constant (compare Figs. 3 and 6).

## V. CONCLUSIONS

In this paper we found optimal cooperator selection policies for multihop networks with MIMO transmissions. The cooperator selection process was modeled for arbitrary topologies through a suitable Markov chain. Hence, this chain was reduced according to an original pruning technique which cuts states with negligible impact on the optimal solution. Thus, we integrated this pruning technique into an advanced solver based on real time dynamic programming and we showed the effectiveness of this approach in terms of goodness of the policy and computational complexity. Our solver finds policies with an additional cost bounded with respect to the optimal and allows to derive the Pareto efficient frontier in terms of transmission cost *vs* delay for arbitrary networks. Finally, through selected application examples we discussed the impact of: 1) the set of nodes that cooperate at each transmission opportunity, 2) the selection of the optimization criteria, i.e., energy *vs* delay minimization and 3) the maximum number of nodes that are allowed to cooperate.

## APPENDIX A

### OUTAGE PROBABILITY COMPUTATION, SINGLE ANTENNA NODES ( $N_A = N_R = 1$ )

When  $N_A = N_R = 1$ , the capacity turns out to be the logarithm of a linear combination of central chi square random variables, i.e.,  $C = \log_2(1 + \rho y)$ , where  $y$  is the sum of  $N_T$  exponential random variables with means  $\Sigma_k$ ,  $k = 1, 2, \dots, N_T$ . For the general case where some of the means  $\Sigma_m$  are equal, i.e.  $\Sigma_k = \Sigma_m$  for some  $k$  and  $m$ , the outage probability can be obtained using the result in [37]. By letting  $\sigma_k$ ,  $r_k$  and  $N_\sigma$  be the unique means, their multiplicity and the number of equality classes, respectively, with  $k = 1, 2, \dots, N_\sigma$  and  $\sum_{k=1}^{N_\sigma} r_k = N_T$ , the outage probability for node  $n$  when nodes in set  $a$  transmit is found using (21) at the top of the next page, where  $f_1(a, b)$  is the cumulative distribution function of a Poisson variable of parameter  $a$ , whereas the set  $\Omega(N_\sigma, k, \ell)$  defines partitions of  $\ell - 1$  through the positive integer indices  $i_j$ , such that  $\sum_{j=1, \neq k}^{N_\sigma} i_j = \ell - 1$  and  $\tau_j(x) = (\sigma_j^{-1} + x)^{-(r_j + i_j)}$ . Simpler expressions for the outage probability hold when all the means are equal or when all the means are different, i.e.,  $r_k = 1$ ,  $k = 1, 2, \dots, N_T$ , see [38, Section 3.3.1, p. 47] and [39].

## APPENDIX B

### PROOF OF LEMMAS AND THEOREMS

#### A. Proof of Lemma 3.1

*Proof:* Let  $f(x, a, y) \stackrel{\text{def}}{=} c(x, a, y) + \gamma J(y)$ . For mapping  $T(\cdot)$  ( $\mathcal{P}$ ) we obtain (22) at the top of the next page, where inequality (1) follows as the minimum taken over a subset  $\mathcal{A}'(x) \subseteq \mathcal{A}(x)$  cannot be smaller than the minimum taken over the original set  $\mathcal{A}(x)$ . Inequality (2) follows from (8) as  $\sum_{y \in \mathcal{N}'(x)} p_{xy}(a) \leq 1$ . Inequality (3) follows from the definition of upper bound  $\bar{J}(x)$  and noting that  $c(x, a, y) \leq c_{\max}$  for all states  $x, y$  and actions  $a$ . Inequality (4) follows from the definition of  $M(x)$ . ■

$$\begin{aligned}
p_{\text{out}}(n, a) &= 1 - \left( \prod_{j=1}^{N_\sigma} \sigma_j^{-r_j} \right) \sum_{k=1}^{N_\sigma} \sum_{\ell=1}^{r_k} \frac{\phi_{k,\ell}(-\sigma_k^{-1})}{\sigma_k^{-r_k+\ell-1}} f_1 \left( \sigma_k^{-1} \frac{2^R - 1}{\rho}, r_k - \ell \right) \\
\phi_{k,\ell}(x) &= (-1)^{\ell-1} \sum_{\Omega(N_\sigma, k, \ell)} \prod_j \binom{i_j + r_j - 1}{i_j} \tau_j(x).
\end{aligned} \tag{21}$$

$$\begin{aligned}
(TJ)(x) &= \min_{a \in \mathcal{A}(x)} \left[ \sum_{y \in \mathcal{N}(x)} p_{xy}(a) f(x, a, y) \right] \\
&\stackrel{(1)}{\leq} \min_{a \in \mathcal{A}'(x)} \left[ \sum_{y \in \mathcal{N}(x)} p_{xy}(a) f(x, a, y) \right] \\
&= \min_{a \in \mathcal{A}'(x)} \left[ \sum_{y \in \mathcal{N}'(x)} p_{xy}(a) f(x, a, y) + \sum_{y \in \mathcal{N}(x) \setminus \mathcal{N}'(x)} p_{xy}(a) f(x, a, y) \right] \\
&\stackrel{(2)}{\leq} \min_{a \in \mathcal{A}'(x)} \left[ \sum_{y \in \mathcal{N}'(x)} p'_{xy}(a) f(x, a, y) + \sum_{y \in \mathcal{N}(x) \setminus \mathcal{N}'(x)} p_{xy}(a) f(x, a, y) \right] \\
&\stackrel{(3)}{\leq} \min_{a \in \mathcal{A}'(x)} \left[ \sum_{y \in \mathcal{N}'(x)} p'_{xy}(a) f(x, a, y) + (c_{\max} + \gamma \bar{J}(x)) \sum_{y \in \mathcal{N}(x) \setminus \mathcal{N}'(x)} p_{xy}(a) \right] \\
&\stackrel{(4)}{\leq} \min_{a \in \mathcal{A}'(x)} \left[ \sum_{y \in \mathcal{N}'(x)} p'_{xy}(a) f(x, a, y) + M(x) \left( c_{\max} + \gamma \max_{x \in \mathcal{S}} \bar{J}(x) \right) \right] \\
&= (T_p J)(x) + \Delta(x).
\end{aligned} \tag{22}$$

$$\begin{aligned}
(TJ)(x) &= \min_{a \in \mathcal{A}(x)} g(x, a) \\
&\stackrel{(1)}{=} \min_{a \in \mathcal{A}'(x)} \left[ \sum_{y \in \mathcal{N}'(x)} p_{xy}(a) f(x, a, y) + \sum_{y \in \mathcal{N}(x) \setminus \mathcal{N}'(x)} p_{xy}(a) f(x, a, y) \right] \\
&\stackrel{(2)}{\geq} \min_{a \in \mathcal{A}'(x)} \left[ \sum_{y \in \mathcal{N}'(x)} p'_{xy}(a) \left( \sum_{y \in \mathcal{N}'(x)} p_{xy}(a) \right) f(x, a, y) \right] \\
&\stackrel{(3)}{\geq} (1 - \eta)(T_p J)(x), \forall x \in \mathcal{S}.
\end{aligned} \tag{23}$$

### B. Proof of Lemma 3.2

*Proof:* Let  $f(x, a, y) \stackrel{\text{def}}{=} c(x, a, y) + \gamma J(y)$ . By definition of mapping  $T(\cdot)$  we obtain (23) at the top of this page, where equality (1) follows from the assumption made in the lemma, inequality (2) follows as the second sum is greater than or equal to zero, and by the definition of  $p'_{xy}(a)$  (Eq. (8)). For inequality (3) consider the following

$$\begin{aligned}
1 &= \sum_{y \in \mathcal{N}'(x)} p_{xy}(a) + \sum_{y \in \mathcal{N}(x) \setminus \mathcal{N}'(x)} p_{xy}(a) \\
&\leq \sum_{y \in \mathcal{N}'(x)} p_{xy}(a) + M(x) \leq \sum_{y \in \mathcal{N}'(x)} p_{xy}(a) + \eta.
\end{aligned} \tag{24}$$

Hence, we can further write  $\sum_{y \in \mathcal{N}'(x)} p_{xy}(a) \geq 1 - \eta$ , which proves the lemma. ■

### C. Proof of Theorem 3.4

*Proof:* From Lemma 3.1 we have:

$$\begin{aligned}
(TJ)(x) &\leq (T_p J)(x) + M(x) [c_{\max} + \gamma \max_{x \in \mathcal{S}} \bar{J}(x)] \\
&\stackrel{(1)}{\leq} (T_p J)(x) + \Delta, \forall x \in \mathcal{S},
\end{aligned} \tag{25}$$

where inequality (1) follows from the assumption made for  $M(x)$  (see Eq. (14)). Hence:

$$(TJ)(x) \leq (T_p J)(x) + \Delta, \forall x \in \mathcal{S}. \tag{26}$$

Now, applying mapping  $T_p(\cdot)$  to both sides gives:

$$(T_p(TJ))(x) \leq (T_p^2 J)(x) + \gamma \Delta, \forall x \in \mathcal{S}, \tag{27}$$

where the expression on the RHS follows from Lemma 1.1.2 of [26]. Re-applying (26) (LHS):

$$\begin{aligned}
(T^2 J)(x) - \Delta &= (T(TJ))(x) - \Delta \\
&\leq (T_p^2 J)(x) + \gamma \Delta, \forall x \in \mathcal{S}
\end{aligned} \tag{28}$$

and hence  $(T^2J)(x) \leq (T_p^2J)(x) + \gamma\Delta + \Delta$ . Repeated iterations of this procedure lead to:

$$(T^k J)(x) \leq (T_p^k J)(x) + \sum_{j=0}^{k-1} \gamma^j \Delta, \forall x \in S. \quad (29)$$

Now, taking the limit as  $k \rightarrow +\infty$  to both sides of (29) leads to:

$$J^*(x) \leq J_p^*(x) + \frac{\Delta}{1-\gamma}, \forall x \in S, \quad (30)$$

which proves (i). For (ii), from Lemma 3.2 we have  $(TJ)(x) \geq \delta(T_p J)(x)$ , where  $\delta$  is as in (17). Applying  $T_p(\cdot)$  to both sides of this last inequality gives

$$(T_p(TJ))(x) \geq (T_p\delta(T_p J))(x) \stackrel{def}{=} (\tilde{T}_p(T_p J))(x), \quad (31)$$

where  $\tilde{T}_p(\cdot)$  is mapping  $T_p(\cdot)$  (Eq. (10)) with discount factor  $\tilde{\gamma} = \gamma\delta$ . Moreover, application of Lemma 3.2 to the LHS of the above inequality returns

$$(T^2J)(x)\delta^{-1} \geq (T_p(TJ))(x) \geq (\tilde{T}_p(T_p J))(x). \quad (32)$$

Repeated iterations of this procedure lead to  $(T^k J)(x) \geq \delta(\tilde{T}_p^{k-1}(T_p J))(x)$ . Now, taking the limit  $k \rightarrow +\infty$  to both sides of the previous inequality gives  $J^*(x) \geq \delta\tilde{J}_p^*(x)$ , where  $\tilde{J}_p^*(x)$  is the optimal cost function for problem  $\mathcal{P}'$  with discount factor  $\tilde{\gamma} = \gamma\delta$ . ■

#### D. Proof of Proposition 3.6

*Proof:* Set  $\mathcal{N}(x) \setminus \mathcal{N}'(x)$  contains the pruned states. These are states  $y$  containing nodes with small probability of receiving the message at the next hop  $i+1$ , given  $x$ . From Lemma 3.5 the maximizing action  $a'_{\max} = \operatorname{argmax}_{a \in \mathcal{A}'(x)} [\sum_{y \in \mathcal{N}(x) \setminus \mathcal{N}'(x)} p_{xy}(a)]$  corresponds to the case where the maximum number of nodes allowed by  $\mathcal{A}'(x)$  transmit, as all receiving nodes  $n \in \mathcal{T}^-(x)$  maximize their reception probability, namely  $p_{\text{succ}}(n, a)$ , for this action. Thus,  $a'_{\max} = a_{\max}$ , where  $a_{\max}$  was defined in Lemma 3.5. This implies that  $M(x) = \sum_{y \in \mathcal{N}(x) \setminus \mathcal{N}'(x)} p_{xy}(a_{\max})$  which is, by definition, the probability that the system in hop  $i+1$  will move to state  $y$  when action  $a_{\max}$  is selected. If we define  $y \in \mathcal{N}(x) \setminus \mathcal{N}'(x)$  as any state for which: 1) all nodes that were successful in  $x$  are still successful in  $y$  and 2) at least one node in  $\mathcal{V}(x)$  is successful, then by the way we constructed  $\mathcal{V}(x)$  we have

$$\begin{aligned} M(x) &= \sum_{y \in \mathcal{N}(x) \setminus \mathcal{N}'(x)} p_{xy}(a_{\max}) \\ &= \sum_{\Psi(|\mathcal{V}(x)|)} \prod_{j=1}^{|\mathcal{V}(x)|} v(j)^{\xi(j)} (1-v(j))^{1-\xi(j)} \end{aligned} \quad (33)$$

and the inequality in (19) is granted by the construction algorithm. ■

#### REFERENCES

- [1] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.
- [2] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—part I: system description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.
- [3] —, "User cooperation diversity—part II: implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.
- [4] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [5] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.
- [6] M. Chen, S. Serbetli, and A. Yener, "Distributed power allocation strategies for parallel relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 552–561, Feb. 2008.
- [7] Y. Fan and J. Thompson, "MIMO configurations for relay channels: theory and practice," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1774–1780, May 2007.
- [8] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 659–672, 2006.
- [9] L. Liu and H. Ge, "Space-time coding for wireless sensor networks with cooperative routing diversity," in *Proc. Asilomar Conf. Signals, Systems and Computers*, vol. 1, Pacific Grove, CA, Nov. 2004, pp. 1271–1275.
- [10] T. Miyano, H. Murata, and K. Araki, "Cooperative relaying scheme with space time code for multihop communications among single antenna terminals," in *Proc. IEEE GLOBECOM*, vol. 6, Dallas, TX, Nov.-Dec. 2004.
- [11] —, "Space time coded cooperative relaying technique for multihop communications," in *Proc. IEEE VTC Fall*, Los Angeles, CA, Sep. 2004.
- [12] A. Del Coso, S. Savazzi, U. Spagnolini, and C. Ibars, "Virtual MIMO channels in cooperative multi-hop wireless sensor networks," in *Proc. Annual Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 2006, pp. 75–80.
- [13] S. Savazzi and U. Spagnolini, "Energy aware power allocation strategies for multihop-cooperative transmission schemes," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 318–327, Feb. 2007.
- [14] L. Zhang and L. J. Cimini Jr., "Power-efficient relay selection in cooperative networks using decentralized distributed space-time block coding," *EURASIP J. Advances in Signal Proc.*, vol. 2008, no. 1, 2008.
- [15] L. Ong and M. Motani, "Optimal routing for the Gaussian multiple-relay channel with decode-and-forward," in *Proc. Int. Conf. Information Theory (ISIT)*, Nice, France, June 2007.
- [16] L. Ong, W. Wang, and M. Motani, "Achievable rates and optimal schedules for half duplex multiple-relay networks," in *Proc. Allerton Conf.*, Monticello, IL, Sep. 2008.
- [17] J. Si, Z. Li, Z. Liu, and X. Lu, "Joint route and power allocation in cooperative-multihop networks," in *Proc. IEEE Int. Conf. Circuits and Systems for Commun.*, Shanghai, China, May 2008, pp. 114–118.
- [18] Z. Yang and A. Høst-Madsen, "Routing and power allocation in asynchronous Gaussian multiple-relay channels," *EURASIP J. Wireless Commun. and Networking*, vol. 2006, no. 2, pp. 1–11, Apr. 2006.
- [19] F. Li, A. Lippman, and K. Wu, "Minimum energy cooperative path routing in wireless networks: an integer programming formulation," in *Proc. IEEE VTC Spring*, Melbourne, Australia, 2006.
- [20] A. E. Khandani, J. Abounadi, E. Modiano, and L. Zheng, "Cooperative routing in static wireless networks," *IEEE Trans. Commun.*, vol. 55, no. 11, pp. 2185–2192, Nov. 2007.
- [21] A. S. Ibrahim, Z. Han, and K. J. Liu, "Distributed energy-efficient cooperative routing in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3930–3941, Oct. 2008.
- [22] Y. Yuan, M. Chen, and T. Kwon, "A novel cluster-based cooperative MIMO scheme for multi-hop wireless sensor networks," *EURASIP J. Advances in Signal Proc.*, vol. 2006, no. 2, pp. 38–46, Apr. 2006.
- [23] A. Del Coso, U. Spagnolini, and C. Ibars, "Cooperative distributed MIMO channels in wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 402–414, Feb. 2007.
- [24] S. Lakshmanan and R. Sivakumar, "Diversity routing for multi-hop wireless networks with cooperative transmissions," in *Proc. IEEE SECON*, Rome, Italy, 2009.
- [25] A. G. Barto, S. J. Bradtke, and S. P. Singh, "Learning to act using real time dynamic programming," *Artificial Intelligence*, vol. 72, no. 1–2, pp. 81–138, Jan. 1995.

- [26] D. P. Bertsekas, *Dynamic Programming and Optimal Control: Vol. II*, 2nd edition. Athena Scientific, 2001.
- [27] T. Smith and R. G. Simmons, "Focused real-time dynamic programming for MDPs: squeezing more out of a heuristic," in *Proc. Nat. Conf. on Artificial Intelligence*, Boston, MA, July 2006.
- [28] K. Lingkun, X. N. Soon, R. G. Maunder, and L. Hanzo, "Maximum-throughput irregular distributed space-time code for near-capacity cooperative communications," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1511–1517, Mar. 2010.
- [29] A. Sezgin and E. A. Jorswieck, "Capacity achieving high rate space-time block codes," *IEEE Commun. Lett.*, vol. 9, no. 5, pp. 435–437, May 2005.
- [30] M. Franceschini, G. Ferrari, and R. Raheli, *LDPC Coded Modulations*. Springer Verlag, 2009.
- [31] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, pp. 475–501, 1964.
- [32] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.
- [33] T. Ratnarajah and R. Vaillancourt, "Complex singular Wishart matrices and applications," *Computers & Mathematics with Applications*, vol. 50, no. 3-4, pp. 399–411, Aug. 2005.
- [34] R. L. Keeney and H. Raiffa, *Decision with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge University Press, 1993.
- [35] D. P. Bertsekas and J. N. Tsitsiklis, "An analysis of stochastic shortest path problems," *Mathematics of Operations Research*, vol. 16, pp. 580–595, Oct. 1991.
- [36] H. B. McMahan, M. Likhachev, and G. J. Gordon, "Bounded real-time dynamic programming: RTDP with monotone upper bounds and performance guarantees," in *Proc. International Conference on Machine Learning*, Bonn, Germany, Oct. 2005.
- [37] S. V. Amari and R. B. Misra, "Closed-form expressions for distribution of sum of exponential random variables," *IEEE Trans. Reliability*, vol. 46, no. 4, pp. 519–522, Dec. 1997.
- [38] P. V. Mieghem, *Performance Analysis of Communications Networks and Systems*. Cambridge University Press, 2006.
- [39] M. Biswal, N. Biswal, and D. Li, "Probabilistic linear programming problems with exponential random variables: a technical note," *European J. Operational Research*, vol. 111, pp. 589–597, 1998.



**Michele Rossi** received the Laurea degree in Electrical Engineering (with honors) and the Ph.D. degree in Information Engineering from the University of Ferrara in 2000 and 2004, respectively. From March 2000 to October 2005 he has been a Research Fellow at the Department of Engineering of the University of Ferrara. During 2003 he was on leave at the Center for Wireless Communications (CWC) at the University of California San Diego (UCSD), where he performed research on wireless sensor networks. In November 2005 he joined the Department of Information Engineering of the University of Padova, Italy, where he is an Assistant Professor. Dr. Rossi is currently part of the EU-funded projects SENSEI, IOT-A and SWAP, all on wireless sensor networks and Internet of Things. Broadly, his research interests are centered around the following topics: data dissemination in distributed ad hoc and wireless sensor networks, integrated MAC/routing schemes, application of compressive sensing for the reconstruction of signals in wireless sensor networks and cooperative routing policies for ad hoc networks. His current research activity also includes energy scavenging solutions for wireless sensor nodes and the utilization of sensing technology within smart buildings.



**Cristiano Tapparello** received the B.Sc. degree and the M.Sc. degree (with honors) in Computer Engineering from University of Padova, Italy, in 2005 and 2008, respectively. Since January 2009 he has been a Ph.D. student at the same University. His current research interests include cooperative routing policies for ad hoc networks, interference-aware optimization of routing and scheduling and information-theoretic aspects of cognitive radio techniques in ad hoc networks.



**Stefano Tomasin** (S'99, M'03) received the Laurea degree and the Ph.D. degree in Telecommunications Engineering from the University of Padova, Italy, in 1999 and 2003, respectively. In the Academic year 1999–2000 he was on leave at the IBM Research Laboratory, Zurich, Switzerland, doing research on signal processing for magnetic recording systems. In the Academic year 2001–2002 he was on leave at Philips Research, Eindhoven, the Netherlands, studying multicarrier transmission for mobile applications. In the second half of 2004 he was visiting at Qualcomm, San Diego (CA) doing research on receiver design for mobile cellular systems. Since 2005 he is assistant professor at University of Padova, Italy. In 2007 he has been visiting faculty at Polytechnic University of Brooklyn, NY, working on cooperative communications. His current research interests include signal processing for wireless communications, access technologies for multiuser/multi-antenna systems and cross-layer protocol design and evaluation.