Relay Selection Schemes Relying on Adaptive Modulation and Imperfect Channel Knowledge for Cooperative Networks

Fulvio Babich*, Alessandro Crismani* and Lajos Hanzo†

* Dipartimento di Ingegneria Industriale e dell'Informazione, Università di Trieste, Italy
† School of Electronics and Computer Science, University of Southampton, UK
babich@units.it, alessandro.crismani@phd.units.it, lh@ecs.soton.ac.uk

Abstract— Cooperative MAC protocols have been shown to increase the attainable network throughput and to reduce the outage probability when perfect Channel State Information (CSI) is available either at the source station or at candidate relays. This paper investigates the realistically achievable network performance of relay selection schemes under imperfect CSI knowledge. Novel relay selection algorithms, based on a Markov Decision Process (MDP) framework, are proposed for mitigating the performance degradation introduced by imperfect channel knowledge. We demonstrate that the performance of cooperative MAC schemes crucially depends on the accuracy of the CSI knowledge. Furthermore, it is shown that the Markov framework allows the source station to predict the channel quality and hence to reduce the number of transmission errors, ultimately resulting in an improved network performance.

I. INTRODUCTION

Recently, a cooperative approach has been used for designing Medium Access Control (MAC) protocols [1]–[4], where the data transmitted from a source station are forwarded by a relay to obtain a diversity gain at the destination. Cooperative MAC schemes increase the achievable network performance by facilitating multi-hop communications that mitigate the detrimental effects of fading on the link between the source and the destination. However, the above-mentioned solutions rely on idealized simplifying assumptions for the physical layer. In particular, [1]–[4] assume that either the source station or some of the relays benefit from instantaneous and perfect Channel State Information (CSI).

Partially Observable Markov Decision Process (POMDP) techniques have been used for improving the performance of non-cooperative schemes, when the system's state is characterised by a certain degree of uncertainty [5], [6]. A cooperative system using realistic imperfect CSI knowledge was considered in [7], where the different relays employed a POMDP for selecting the cooperative partner and a transmission technique dispensing with adaptive modulation or channel coding was assumed. Similarly, a POMDP based approach was used in [8] for assessing the performance of a cooperative sensor network. Again, the authors of [8] also dispensed with adaptive modulation systems and cooperative MAC protocols. As a further advance, a Markov formulation was used in [9] for modelling the time instants of a cooperative communication

system, where the relay selection problem was treated as a "multi-armed bandit problem".

Our contribution models and compares the performance of both proactive and reactive relay selection algorithms using realistic outdated and imperfect CSI knowledge. We conceive POMDP based relay selection algorithms for mitigating the performance degradation introduced by the channel uncertainty. A sophisticated physical layer is considered, which includes adaptive modulation and incremental encoding. The Sphere Packing Bound (SPB) is used to assess the attainable performance of efficient codes [10], [11]. We demonstrate that the performance of a relay selection scheme employed for a cooperative MAC protocol is sensitive to the quality of the CSI knowledge. It is also shown that MDP based techniques allow us to mitigate the impairments imposed by the imperfect CSI, hence effectively increasing the network's performance.

The paper is organised as follows. Section II details the system under investigation. Section III describes the relay selection schemes using imperfect CSI. Our novel POMDP aided schemes are detailed in Section IV. Our comparisons between the different relaying techniques are provided in Section V. Finally, the paper is concluded in Section VI.

II. SYSTEM DESCRIPTION

Cooperative relay selection schemes may be categorised as proactive or reactive [1]. In proactive schemes cooperative relay-aided multi-hop transmissions may replace a slow-rate direct communication link for improving the network's throughput and/or integrity. In particular, proactive schemes select either the direct or the specific relay-aided link with the highest transmission rate. By contrast, reactive schemes may be viewed as an extension of Automatic Repeat reQuest (ARQ) techniques to cooperative networks and aim for reducing the error probability at the destination, thus enhancing the system's integrity. In reactive systems the Source Station (SS) firstly transmits its data directly to the Destination Station (DS). Cooperation is invoked only upon failure of the direct link. In such cases, the Relay Station (RS) retransmits the data to provide a diversity gain at the DS in order to increase the probability of recovering the transmitted information.

Diverse MAC layer solutions have been proposed for enabling cooperative wireless networks. Some schemes rely on contention mechanisms between the candidate RSs [1], [4]. In other contributions [2], [3] the RS is chosen by the SS before

initiating its transmissions using information gleaned from past communications. In this paper the RS is chosen by the SS based on imperfect channel estimates that might be outdated. Selecting the RS at the SS avoids time-consuming backoff procedures and collisions at RSs. Furthermore, this regime reduces the power consumption of RSs in reactive schemes, where the candidate partners have to decode the SS's message even if they will not be the preferred helper. Since the scope of the paper is to compare the performance of diverse RS selection techniques, the MAC overhead involved in setting up cooperation (e.g. the RTS / HTS / CTS frame exchange in the CoopMAC protocol [3]) is not taken into account.

Consider a scenario having a SS, N RSs $\{H_n\}_{n=1}^N$ and a DS. Denote the distance between stations "a" and "b" normalized to the distance between the SS and the DS by $d(a,b)$. Time is discrete and divided into slots. Consider the lth time slot t_l. Denote the instantaneous Signal to Noise Ratio (SNR) of the Source to Destination (SD) link by $\gamma_{SD}^{t_l}$. Similarly, denote the instantaneous SNRs of the link between the SS and the nth RS and of the link between the nth RS and the DS by $\gamma_{SH_n}^{t_l}$ and $\gamma_{H_nD}^{t_l}$, respectively. Let us assume that the SNRs pertaining to different time slots are identically distributed. We denote the average quantities on the corresponding links by $\Gamma_{SD} = \mathbb{E}[\gamma_{SD}^{(\cdot)}]$, $\Gamma_{SH_n} = \mathbb{E}[\gamma_{SH_n}^{(\cdot)}] = \Gamma_{SD}/d(S,H_n)^2$ and $\Gamma_{H_nD} = \mathbb{E}[\gamma_{H_nD}^{(\cdot)}] = \Gamma_{SD}/d(H_n,D)^2$. Assuming Rayleigh fading, $\gamma_{SD}^{(\cdot)}$, $\{\gamma_{SH_n}^{(\cdot)}\}_{n=1}^N$ and $\{\gamma_{H_nD}^{(\cdot)}\}_{n=1}^N$ are exponentially distributed with means equal to Γ_{SD}, $\{\Gamma_{SH_n}\}_{n=1}^N$ and $\{\Gamma_{H_nD}\}_{n=1}^N$, respectively. The values of SNRs on the same link in two time slots t_l and $t_{l+\tau}$ are correlated according to $\mathbb{R}(\gamma_{(\cdot)}^{t_l}, \gamma_{(\cdot)}^{t_{l+\tau}}) = J_0(2\pi f_d \tau)$, where J_0 is the zero-order unmodified Bessel function of the first kind and f_d is the Doppler frequency. The channel qualities of the different links are independent.

Consider an adaptive scenario, where $\{R_i\}_{i=1}^K$ are the K available transmission rates corresponding to K different transmission modes. We consider the BPSK, QPSK, 16-QAM and 64-QAM modulation modes, that provide rates of 1, 2, 4, 6 Mbits/s, respectively. Denote the switching threshold used by the SS to select the ith transmission mode by $\overline{\Gamma}_i$. The adaptive mode switching thresholds are analytically derived by using the Sphere Packing Bound, as described in [10] for a non-cooperative scenario and extended in [11] to a cooperative scheme. The approach detailed in [11] is used for deriving the switching thresholds used by the RS. More particularly, we note that in a cooperative scenario the DS combines the signal received from the SS and from the activated RS, performing a joint decoding. Hence, as further detailed in [12], the new success threshold $\overline{\Gamma}_k^{R_i, \gamma_{SD}}$ of the kth mode on the relay to destination link depends both on the mode R_i used by the SS and on the SNR value γ_{SD} of the SD link. We consider an incremental encoding technique, where the SS and RS transmissions encode 1600 bits using, as an example, a code-rate of $4/5$. Hence, the DS operates at a code-rate of $2/5$, when the two signals are jointly decoded. An idealised ON-OFF assumption is used for modelling the behavior of efficient channel codes. More particularly, a transmission using the kth modulation mode on a link with quality γ is considered to

be successful, if γ is higher than the SPB-based switching threshold and unsuccessful otherwise [10]. Naturally, the SPB approach is accurate only in the cases where near-capacity codes are used in the physical layer. We do not consider a specific encoding technique, rather we assume that both the codes used by the SS and the RS are near-capacity schemes and hence the SPB-based predicted performance is accurate. Further details on the code design conceived for incremental encoding aided systems might be found in [12].

We assume that the SNR values are quantised by the SS and opt for the SNR quantisation thresholds of $(L_0 = 0) < (L_k = \overline{\Gamma}_k) < (L_{K+1} = \infty)$. The fading envelope at a time instant t_l is quantised to the value c_k $(0 \le k \le K)$ if we have $L_k \le \gamma_{(\cdot)}^{t_l} < L_{k+1}$. Let us denote the SNR value obtained after quantising the quantity $\gamma_{(\cdot)}^{t_l}$ by $\hat{\gamma}_{(\cdot)}^{t_l}$. Furthermore, the quantised SNR values may be corrupted by noise. More particularly, the SS observes potentially corrupted estimates $\tilde{\gamma}_{(\cdot)}$ of the quantised SNR values $\hat{\gamma}_{(\cdot)}$. We denote the probabilities of under-estimating and over-estimating the SNR value by σ_1 and σ_2, respectively. The probability of observing a channel SNR value $\tilde{\gamma}_{(\cdot)}$, when the link quality is $\hat{\gamma}_{(\cdot)}$ is given by:

$$P\left(\tilde{\gamma}_{(\cdot)}^{t_{l-\tau}} = c_i \big| \hat{\gamma}_{(\cdot)}^{t_{l-\tau}} = c_j\right) = \begin{cases} \sigma_1 & i = j-1 \\ 1-\sigma_1 & i = j = K \\ \sigma_2 & i = j+1 \\ 1-\sigma_2 & i = j = 0 \\ 1-\sigma_1-\sigma_2 & 0 < i = j < K \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

III. RELAY SELECTION SCHEMES USING IMPERFECT CSI

In a proactive RS selection scheme the SS chooses the RS providing the highest-rate transmission. Consider the lth time slot and the nth relay. Let us assume that the SS knows the values $\tilde{\gamma}_{SD}^{t_{l-\tau}}$, $\{\tilde{\gamma}_{SH_n}^{t_{l-\tau}}\}_{n=1}^N$ and $\{\tilde{\gamma}_{H_nD}^{t_{l-\tau}}\}_{n=1}^N$. Denote the highest-rate transmission mode available on the SR link by $R_{SH_n}^* = \max\{R_j \in K : \overline{\Gamma}_j < \tilde{\gamma}_{SH_n}^{t_{l-\tau}}\}$ and the highest-rate transmission mode available on the RD link by $R_{H_nD}^* = \max\{R_j \in K : \overline{\Gamma}_j^{R_{SH_n}^*, \tilde{\gamma}_{SD}^{t_{l-\tau}}} < \tilde{\gamma}_{H_nD}^{t_{l-\tau}}\}$. The RS that achieves the highest transmission rate is selected according to:

$$\texttt{choose relay } n^* = \underset{1 \le n \le N}{\arg\max} \left\{ \left(\frac{1}{R_{SH_n}^*} + \frac{1}{R_{H_nD}^*} \right)^{-1} \right\}. \tag{2}$$

Finally, cooperation is activated if we have:

$$\left(\frac{1}{R_{SH_{n^*}}^*} + \frac{1}{R_{H_{n^*}D}^*} \right)^{-1} > R_{SD}^*, \tag{3}$$

where $R_{SD}^* = \max\{R_j \in K : \overline{\Gamma}_j < \tilde{\gamma}_{SD}^{t_{l-\tau}}\}$ is the highest attainable rate on the SD link. Otherwise direct communication provides a higher-rate and hence is to be preferred.

Note that, since the CSI known at the SS is not the instantaneous one, the specific transmission mode activated may result in an unsuccessful decoding at the DS, especially when the CSI fluctuates rapidly. Let us consider a direct communication and assume that the SS chooses the ith transmission mode based on the SNR value $\tilde{\gamma}_{SD}^{t_{l-\tau}}$. The transmission is unsuccessful if we have $\gamma_{SD}^{t_l} < \overline{\Gamma}_i$. Similarly, if the SS opts for

choosing the nth RS and the transmission modes R_j and R_k on the SR and RD links, respectively, the cooperative transmission is unsuccessful if we have $\gamma_{SH_n}^{t_l} < \overline{\Gamma}_j \vee \gamma_{H_nD}^{t_l} < \overline{\Gamma}_k^{R_j, \gamma_{SD}^{t_l}}$.

In reactive schemes, the SS opts for the specific RS that provides the most prompt retransmission to the DS, given that it should be able to decode the SS's signal. More explicitly, the SS transmits a packet to the DS choosing the highest attainable rate $R_{SD}^* = \max\{R_j \in K : \overline{\Gamma}_j < \tilde{\gamma}_{SD}^{t_l - \tau}\}$. The SS considers as candidate partners all the RSs capable of decoding the direct transmission. Hence the nth RS is a candidate partner if $\overline{\Gamma}_{R_{SD}^*} < \tilde{\gamma}_{SH_n}^{t_l - \tau}$. Let us denote the set containing the candidate RSs at the time slot t_l by $\mathcal{H} \subset \{1, \ldots, N\}$. Denote the highest-rate attainable on the RD link for a candidate relay $h \in \mathcal{H}$ by $R_{H_hD}^* = \max\{R_j \in K : \overline{\Gamma}_j^{R_{SD}^*, \tilde{\gamma}_{SD}^{t_l - \tau}} < \tilde{\gamma}_{H_hD}^{t_l - \tau}\}$. The RS selection algorithm chooses the highest-rate RS from the candidate partners and proceeds according to:

$$\texttt{choose relay} \quad h^* = \underset{h \in \mathcal{H}}{\operatorname{argmax}} \left\{ R_{H_hD}^* \right\}. \tag{4}$$

Naturally, transmission errors are imposed by the imperfect CSI knowledge. Let us assume that the SS chooses to transmit at a rate of $R_{SD}^* = R_i$ and invokes cooperation by relying on the nth relay using a rate of $R_{H_nD}^* = R_j$ in case of the SD link's failure. The direct transmission fails if we have $\gamma_{SD}^{t_l} < \overline{\Gamma}_i$ and hence cooperation is invoked. Then, the RS's transmission may also fail, if $\gamma_{H_nD}^{t_l} < \overline{\Gamma}_j^{R_i, \gamma_{SD}^{t_l}}$, in which case the packet has to be retransmitted.

IV. POMDP AIDED RELAY SELECTION

A. Partially Observable Markov Decision Processes

Markov Decision Processes (MDP) provide a mathematical framework for modelling decisions, when the underlying system evolves according to a Markov probability law. A MDP may be represented by a tuple (S, A, T, r), where S is the set enumerating the possible system states. Different actions obeying $a(s_i^{t_i}) \in A$ may be chosen when the system is in a state $s_i^{t_l} \in S$ at the time instant t_l. The law $T(s_i^{t_l}, a, s_j^{t_{l+1}}), T : S \times A \times S \to \mathbb{R}$ represents the probability that the system moves from the state $s_i^{t_l}$ to the state $s_j^{t_{l+1}}$ upon taking action a. Finally, an immediate reward $r(s_i^{t_l}, a), r : S \times A \to \mathbb{R}$ is earned upon choosing action a when in the state $s_i^{t_l}$.

A stationary policy $\pi : S \to A$ is a function that selects an action based on the current state of the system. The optimization problem aims to maximize the discounted reward [5]:

$$J(\pi, s_i) = \mathbb{E}_\pi^{s_i} \left[\sum_{k=0}^{\infty} \gamma^k r(s_i^k, a_\pi(s_i^k)) \right], \tag{5}$$

where γ is the discount factor. Define the expected reward earned by choosing the policy π when commencing from state s_i by $V(\pi, s_i)$. The Bellman equation of [13] may be used to find the optimal expected reward as:

$$V^*(s_i) = \max_{a \in A} \left[r(s_i, a) + \gamma \sum_{s_j \in S} T(s_i, a, s_j) V^*(s_j) \right]. \tag{6}$$

The policy that achieves $V^*(s_i)$ is the optimal policy $\pi^*(s_i)$ to be adopted, when in state s_i.

MDPs require a perfect knowledge of the current state. When the state is not perfectly known, the decision process may be aided by using POMDPs, which rely on observations of the system gleaned after performing an action. Denote the set enumerating all the possible observations by O. The law $\Omega(s_i^{t_l})$, $\Omega : S \times A \to \Pi(O)$, where Π is a probability distribution, represents the probability of observing $o_j(t_l) \in O$ upon taking action a when the system is in the state $s_i^{t_l}$.

Let us denote the probability of being in a state $s_i^{t_l}$ by $b(s_i^{t_l})$, which is referred to as the state belief, and the vector enumerating the belief of every possible state by $\mathbf{b}(t_l)$. A stationary policy selects an action based on the probabilities $\mathbf{b}(t_l)$ of the different states. Diverse algorithms have been proposed for finding the optimal policy of a POMDP [14]. However, these algorithms are typically suitable for problems associated with a small number of states. When aiming for finding suboptimal policies that guarantee satisfying a given performance in more complex problems, typically heuristic algorithms have been proposed. This paper considers the so-called Q-MDP [14] heuristic, since it achieved better results compared to other heuristics, such as the Most Likely State [15] and Action Voting [16] heuristics.

B. POMDP aided formulation of relay selection schemes

Consider the uth link of the cooperative network. The evolution of the quantised fading process on a specific link may be modelled using a Finite State Markov Chain (FSMC), where transitions are governed by a first order Markov law $P(\hat{\gamma}_{(\cdot)}^{t_{l+1}} = c_j | \hat{\gamma}_{(\cdot)}^{t_l} = c_i) = t_u^{i,j}$. The transition process may be described by a matrix $T_u = [t_u^{i,j}]$. The values $t_u^{i,j}$ depend both on the Doppler spread as well as on the average SNR and may be obtained for a Rayleigh fading process as detailed in [17]. The link is in state $s_i^{t_l}$ at a time t_l, if $\hat{\gamma}_{(\cdot)}^{t_l} = c_i$. The $(K+1)$ legitimate states on the uth link are enumerated by the set S_u.

The cooperative system may be modelled by using $(1+2N)$ FSMCs. The states of the cooperative network are obtained by enumerating all the combinations between the single states of each of the $(1 + 2N)$ FSMC. Let us represent the set enumerating all the states of the cooperative system by $\mathbf{S} = S_1 \times S_2 \times \cdots \times S_{1+2N}$ and the state at time t_l by $\mathbf{s}^{t_l}$. The transition matrix $\mathbf{T}$ may be obtained as:

$$\mathbf{T}(\mathbf{s}^{t_l}, \mathbf{s}^{t_{l+1}}) = \prod_{u=1}^{1+2N} t_u^{s_i^{t_l}, s_j^{t_{l+1}}}. \tag{9}$$

The set of actions differs between the proactive as well as the reactive scheme and characterizes the SS's behaviour. In proactive systems the SS may opt for remaining silent, transmitting directly or transmitting with the aid of cooperation. In the case of a direct communication the action further divides into K actions corresponding to the K transmission modes. In the case of a cooperative session, the action further divides into K^2 actions given by the combination of the two transmission modes used on the SR and RD links. Hence, $(1 + K + NK^2)$ actions are considered. The reward is zero if the SS remains silent. In the case of a direct transmission, the reward is:

$$r(\hat{\gamma}_{SD}^{t_l} = c_i, a = R_k) = \begin{cases} R_k & c_i \geq \overline{\Gamma}_k \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

$$r\left([\hat{\gamma}_{SD}^{t_l}=c_i,\hat{\gamma}_{SH_n}^{t_l}=c_j,\hat{\gamma}_{H_nD}^{t_l}=c_h],a=\{R_l^{srn},R_m^{rnd}\}\right)=\begin{cases}\left(\frac{1}{R_l^{srn}}+\frac{1}{R_m^{rnd}}\right)^{-1} & c_j\geq\overline{\Gamma}_l\wedge c_h\geq\overline{\Gamma}_m^{R_l^{srn},c_i}\\0 & \text{otherwise}\end{cases}\tag{9}$$

$$r\left([\hat{\gamma}_{SD}^{t_l}=c_i,\hat{\gamma}_{SH_n}^{t_l}=c_j,\hat{\gamma}_{H_nD}^{t_l}=c_h],a=\{R_l^{sd},R_m^{rnd}\}\right)=\begin{cases}\left(\frac{1}{\overline{R^{sd}}}+\frac{1}{R_m^{rnd}}\right)^{-1} & c_i<\overline{\Gamma}_l\wedge c_j\geq\overline{\Gamma}_l\wedge c_h\geq\overline{\Gamma}_m^{R_l^{sd},c_i}\\R_l^{sd}+\alpha R_m^{rnd} & c_i\geq\overline{\Gamma}_l\wedge c_j\geq\overline{\Gamma}_l\wedge c_h\geq\overline{\Gamma}_m^{R_l^{sd},c_i}\\R_l^{sd} & c_i\geq\overline{\Gamma}_l\wedge(c_j<\overline{\Gamma}_l\vee c_h<\overline{\Gamma}_m^{R_l^{sd},c_i})\\0 & \text{otherwise}\end{cases}\tag{10}$$
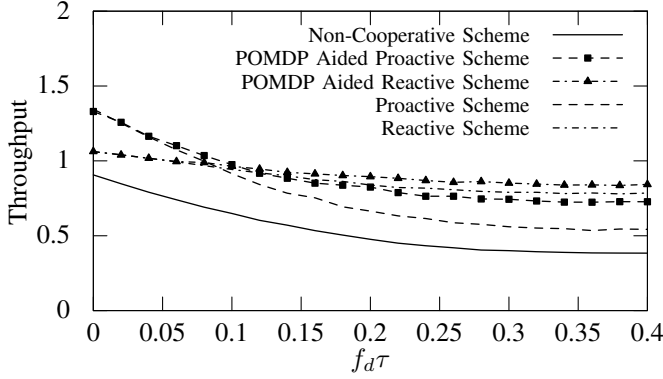

Fig. 1. Throughput of cooperative and non-cooperative relay selection schemes under outdated CSI knowledge, $\Gamma_{SD}=5$ dB.
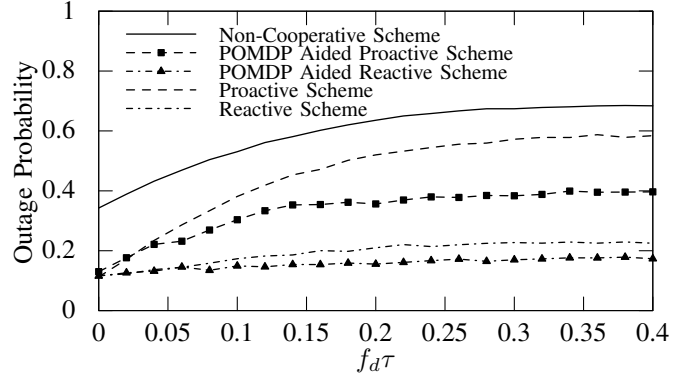

Fig. 2. Outage performance of cooperative and non-cooperative relay selection schemes under outdated CSI knowledge, $\Gamma_{SD}=5$ dB.

In the case that the nth RS is chosen for a cooperative communication, the reward is obtained according to (9).

In reactive systems, the SS can choose to remain silent or to transmit at any time instant. In the case of opting for transmission the action divides into $N\times K^2$ further actions corresponding to the transmission modes selected on the SD and RD links for each RS. The reward is zero if the SS remains silent and is calculated according to (10) if the SS transmits, where $\overline{R^{sd}}$ is the minimum between R_l^{sd} and the highest rate supported by the SNR $\hat{\gamma}_{SD}^{t_l}$, while the value $\alpha=0.1$ is used to prioritise the activation of the highest-rate RS in the case where the direct communication might fail.

Assume that in a slot t_l the SS observes an outdated and imperfect SNR estimate on each network link. More explicitly, consider one of the $(1+2N)$ FSMCs modeling the network and assume that an observation $o_j(t_{l-\tau})=\tilde{\gamma}_{(\cdot)}^{t_{l-\tau}}$ of the state $s_i^{t_{l-\tau}}$ is obtained at the time instant t_l. The probability $P(o_j(t_{l-\tau})|s_i^{t_{l-\tau}})$ of observing a state j while the system was in the state i is obtained using (1). Since observations pertain to the past, an approach similar to that of [5] is used for updating the belief, namely by introducing the initial belief b and updated belief $\tilde{b}$ concepts. The initial belief at time $t_{l-\tau}$ is obtained from the updated belief at time $t_{l-\tau-1}$ as:

$$b(s_i^{t_{l-\tau}})=P(s_i^{t_{l-\tau}})=\sum_{j=0}^{K}P\left(s_i^{t_{l-\tau}}|s_j^{t_{l-\tau-1}}\right)\tilde{b}(s_j^{t_{l-\tau-1}}).\tag{11}$$

The updated belief $\tilde{b}(s_i^{t_{l-\tau}})=P(s_i^{t_{l-\tau}}|o_j(t_{l-\tau}))$ represents the probability of being in a particular state at the time $t_{l-\tau}$ after obtaining an observation $o_j(t_{l-\tau})$ and is obtained as:

$$\tilde{b}(s_i^{t_{l-\tau}})=\frac{P(o_j(t_{l-\tau})|s_i^{t_{l-\tau}})b(s_i^{t_{l-\tau}})}{P(o_j(t_{l-\tau}))}.\tag{12}$$

The updated belief of the cooperative network $\tilde{b}(\mathbf{s}^{t_{l-\tau}})$ may be obtained as the product of the believes $\tilde{b}(s^{t_{l-\tau}})$ of the single links. Finally, the outdated belief is projected into the actual time slot and the current belief is calculated as $b(\mathbf{s}^{t_l})=P(\mathbf{s}^{t_l})=\tilde{b}(\mathbf{s}^{t_{l-\tau}})\cdot\mathbf{T}^{\tau}$.

Once we obtain the current belief the Q-MPD heuristic may be used for finding the transmission policy π.

V. RESULTS

Figures 1 and 2 compare the throughput and the outage probability of both non-cooperative and cooperative systems using a single RS halfway between the SS and the DS. A perfect but outdated CSI is considered. It can be seen that the performance rapidly degrades, when the normalized delay increases (subsequent attempts become less correlated), due to the errors introduced by the outdated CSI. The degradation is much more severe for proactive protocols, where an aggressive adaptive regime is used for improving the throughput. By contrast, reactive schemes are less sensitive to the outdated CSI, hence are better suited to scenarios where the CSI uncertainty is high. We may also observe that the POMDP framework efficiently improves the attainable network performance.

Figure 3 compares the throughput of cooperative systems subjected to imperfect SNR knowledge. Two RSs are considered, both located at the same distance from the SS and the DS, respectively. It can be seen again that the POMDP aided RS selection techniques are the ones offering the best throughput and that the reactive scheme is less sensitive than the proactive one to the SNR uncertainty. Furthermore, the reactive scheme benefits less from employing POMDP based decisions than the proactive one. The erratic fluctuation of the throughput curves of the POMDP based solutions is caused by the quantised model adopted for the faded channel and by the finite set of actions. Particularly, since the policy chosen
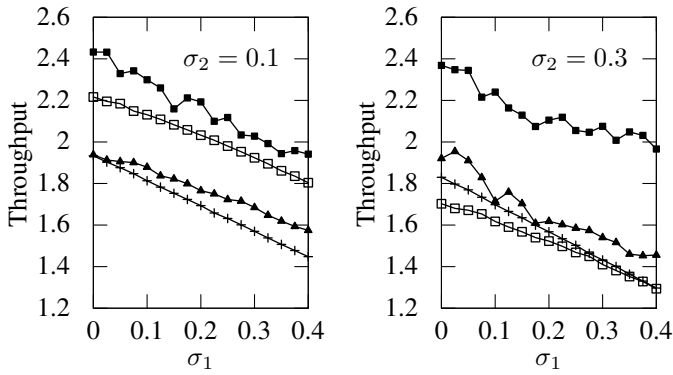
Fig. 3. Throughput comparison of the proactive (□), the POMDP aided proactive (■), the reactive (+) and the POMDP aided reactive (▲) relay selection schemes using corrupted CSI estimates $\Gamma_{SD} = 10$ dB.

Fig. 4. Throughput of cooperative relay selection schemes using a different number of relays under outdated CSI knowledge, $\Gamma_{SD} = 10$ dB.

by the POMDP aided schemes remains the same for a set of under-estimation probabilities, the throughput curve exhibits an oscillating behaviour across the entire probability domain.

Figure 4 compares the throughput of the schemes under investigation in a normalised scenario, where the RSs are placed randomly between the SS and the DS. A different number of candidate RSs selected from the set of 10 available ones is considered for the non-POMDP aided schemes and their performance is compared. A scheme always uses the specific relays closest to the SS. The POMDP schemes rely on only two RSs, since the number of states of the underlying POMDP problem grows exponentially with the number of helpers. Again, one can see that the POMDP aided RS selection schemes attain the highest performance. Furthermore, it is interesting to note that the attainable performance hinges on the RS link SNRs of those closest to the SS. More specifically, the performance of the schemes using the two RSs closest to the SS approaches that of the schemes using all the ten RSs. Finally, the POMDP aided schemes using only two RSs clearly outperform the schemes where the decision is only based on the outdated CSI and not assisted by the Markov framework, even if all the ten RSs are available for cooperation.

VI. CONCLUSIONS AND FUTURE RESEARCH

We investigated the performance of reactive and proactive relay selection schemes when the CSI is not perfectly known at the SS. It was shown that reactive schemes are less sensitive to the CSI inaccuracy, while proactive schemes exhibit a better performance, when the CSI is accurate. Novel RS selection schemes, using a POMDP aided framework, were proposed for predicting the CSI and hence for mitigating the errors introduced by imperfect CSI. It was shown that the POMDP aided schemes clearly outperform their non-POMDP based counterparts. More specifically, proactive schemes exhibit a significant performance improvement when POMDP aided RS selection is considered. Even though reactive schemes exhibit less substantial improvements, the POMDP approach may still be used for approaching the attainable performance.

Relay selection schemes require further investigations. It will be beneficial to design novel MAC schemes that fully exploit the POMDP framework for further reducing the performance degradation imposed by imperfect SNR knowledge.
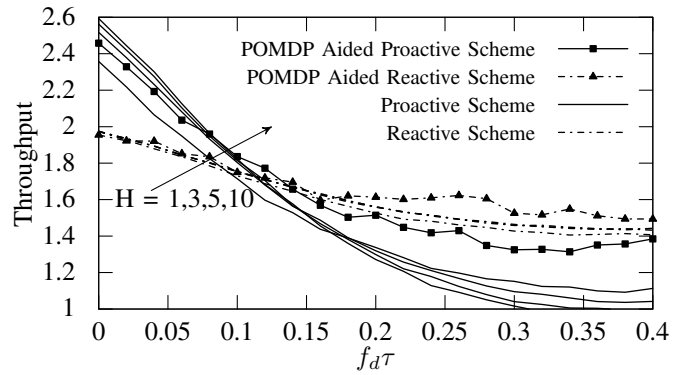
REFERENCES

[1] H. Shan, W. Zhuang, and Z. Wang, "Distributed Cooperative MAC for Multihop Wireless Networks," *IEEE Communications Magazine*, vol. 47, pp. 126–133, February 2009.
[2] S. Sayed, Y. Yang, and H. Hu, "CARD: Cooperative Access with Relay's Data for Multi-Rate Wireless Local Area Networks," in *IEEE International Conference on Communications*, pp. 1–6, June 2009. Dresden, Germany.
[3] P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S. Panwar, "CoopMAC: A Cooperative MAC for Wireless LANs," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 340–354, 2007.
[4] J. Alonso-Zarate, E. Stavrou, A. Stamou, P. Angelidis, L. Alonso, and C. Verikoukis, "Energy-Efficiency Evaluation of a Medium Access Control Protocol for Cooperative ARQ," in *IEEE International Conference on Communications*, pp. 1–5, June 2011. Kyoto, Japan.
[5] A. Karmokar, D. Djonin, and V. Bhargava, "POMDP-Based Coding Rate Adaptation for Type-I Hybrid ARQ Systems over Fading Channels with Memory," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 3512–3523, December 2006.
[6] A. T. Hoang and M. Motani, "Cross-Layer Adaptive Transmission with Incomplete System State Information," *IEEE Transactions on Communications*, vol. 56, pp. 1961–1971, November 2008.
[7] G. Shirazi, P.-Y. Kong, and C.-K. Tham, "A Cooperative Retransmission Scheme in Wireless Networks with Imperfect Channel State Information," in *IEEE Wireless Communications and Networking Conference*, pp. 1–6, April 2009.
[8] H. Li, N. Jaggi, and B. Sikdar, "Relay Scheduling for Cooperative Communications in Sensor Networks with Energy Harvesting," *IEEE Transactions on Wireless Communications*, vol. PP, no. 99, pp. 1–11, 2011.
[9] Y. Wei, F. R. Yu, and M. Song, "Distributed Optimal Relay Selection in Wireless Cooperative Networks with Finite-State Markov Channels," *IEEE Transactions on Vehicular Technology*, vol. 59, pp. 2149–2158, June 2010.
[10] F. Babich, "On the Performance of Efficient Coding Techniques over Fading Channels," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 290–299, January 2004.
[11] F. Babich, "Cooperative Systems for the Fading Channel with Adaptive Modulation and Incremental Encoding," in *IEEE Global Telecommunications Conference*, pp. 1–5, December 2010. Miami, USA.
[12] F. Babich and A. Crismani, "Incremental and Complementary Coding Techniques for Cooperative Medium Access Control Protocols," in *Vehicular Technology Conference, Spring*, pp. 1–5, May 2011. Budapest, Hungary.
[13] R. Bellman, *Dynamic Programming*. Princeton University Press, 1957.
[14] M. L. Littman, A. R. Cassandra, and L. P. Kaelbling, "Learning Policies for Partially Observable Environments: Scaling Up," in *Proceedings of the Twelfth International Conference on Machine Learning*, pp. 362–370, 1995. Tahoe City, USA.
[15] I. Nourbakhsh, R. Powers, and S. Birchfield, "DERVISH: An Office-Navigating Robot," *AI magazine*, vol. 16, no. 2, p. 53, 1995.
[16] R. Simmons and S. Koenig, "Probabilistic Robot Navigation in Partially Observable Environments," in *International Joint Conference on Artificial Intelligence*, vol. 14, pp. 1080–1087, 1995.
[17] H. S. Wang and N. Moayeri, "Finite-State Markov Channel - A Useful Model for Radio Communication Channels," *IEEE Transactions on Vehicular Technology*, vol. 44, pp. 163–171, February 1995.