



**HAL**  
open science

# Sum-rate Optimization and Performance Analysis of Distributed Coding for Sensor Networks: Application to a Distributed Product Code

Karine Amis Cavalec, Yizhi Yin, Ramesh Pyndiah

► **To cite this version:**

Karine Amis Cavalec, Yizhi Yin, Ramesh Pyndiah. Sum-rate Optimization and Performance Analysis of Distributed Coding for Sensor Networks: Application to a Distributed Product Code. IEEE Transactions on Wireless Communications, 2014, 13 (7), pp.3677-3689. 10.1109/TWC.2014.2314095 . hal-01057249

**HAL Id: hal-01057249**

**<https://hal.science/hal-01057249v1>**

Submitted on 27 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Sum-Rate Optimization and Performance Analysis of Distributed Coding for Sensor Networks: Application to a Distributed Product Code

Karine Amis, Yizhi Yin, and Ramesh Pyndiah

**Abstract**—This paper deals with distributed coding for sensor networks. We consider energy-and-complexity-constrained sensors that communicate with a distant sink at a relatively low data rate with the help of a relay with higher-power and better-processing capabilities. We assume an orthogonal multiple access to the resources. Sources transmit independently and use simple short block codes to limit delay and complexity. The relay performs network encoding of the detected codewords and forwards either partial or complete coded information. The sink jointly decodes observations coming both from sources and relay to simultaneously estimate the information sent by the source nodes. We first analyse the cooperative scheme from an information theoretical point of view in order to determine the resource allocation strategy between the sources and the relay, which maximizes the overall network sum-rate. When a single relay is used and the source-to-relay link is not error-free, occasional detection errors introduce additional errors at the relay output, which, in turn, critically affect decoding performance and yield error-floors at the destination. In the case of a distributed coding scheme based on product codes, we obtain a closed-form asymptotically-tight approximation of the error probability under a minimum-distance decoding of the product code at the destination. The proposed analysis takes explicitly into account the error propagation at the relay output and may be extended to other forms of distributed coding schemes.

**Index Terms**—Channel capacity, cooperative systems, forward error correction, relays, sensor network.

## I. INTRODUCTION

SINCE its introduction by van der Meulen in 1971 [1] and the analysis of its capacity by Cover and El Gamal in 1979 [2], the relay channel has found a renewed interest with the emergence of cooperative communications a decade ago. The generalization of the original (source, relay, destination) model to more than three nodes is particularly attractive in the context of sensor networks. Thanks to technological breakthroughs, sensors with improved capabilities, low cost and reduced energy consumption can be mass-deployed [3].

The authors are with the Signal and Communications Department, Telecom Bretagne, 29238 Brest Cedex, France.

Typical applications are remote sensing, smart grid and, more generally, machine-to-machine communications [4], where a high number of devices transmit sensory data to a remote gateway, which, in turn, forwards them to a remote back-end server. Low-power consumption, small burst transmission and low latency are key features of machine-to-machine communications. Neighbouring sensors can be grouped in clusters, with a cluster head acting as a relay of the information transmitted by the other sensors. The available throughput [5], the hierarchical architecture of the sensor network [6], as well as the joint MAC and physical layers optimization have been and are still intensively studied.

The cooperative communication principle applied to the sensor network assumes sensors are grouped into clusters. Different configurations are available. Either all sensors have similar computing resources and share the same roles, alternating between source and relay, or some of them have higher capabilities and are used as relays. In all cases, the source-to-relay distance is supposed to be small relative to the source-to-destination distance. The first processing techniques proposed for the relay were the store-and-forward, the decode-and-forward and the compress-and-forward techniques [2], [7]. Introduced by [8] in 2000, network coding consists in decoding information coming from the sources, storing it and combining it before sending it to the destination. Network coding can be either random (suitable for wireline networks) or structured (suitable for wireless networks). It is particularly attractive for cooperative communications, since the cooperation is carried out in the relay, and thus does not require any information exchange between sources, nor any additional power consumption at the sources. Adaptive switching between cooperation and no cooperation depending on the source-to-destination channel conditions is also feasible and allows to trade control information throughput for data information throughput.

In this paper, we consider a sensor network, scheduled with a perfectly orthogonal division multiple access, where the sensors are sources and the relays are specific devices with higher storage, computing and energy resources. Sources use short forward error correcting codes in order to reduce delays. The relay demodulates, decodes and stores source codewords as they come along. After the source transmissions, the relay jointly encodes the source codewords and transmits partial or complete codewords to the destination. A distributed, long concatenated code with high error correction capability then arises at the destination, and benefits to all sources. This is the

key idea of this approach. The destination jointly decodes the data coming from both the sources and the relay to retrieve the source codewords. Different schemes have already been proposed based on convolutional turbo codes, low-density parity check codes (LDPC) [9], product codes [10]–[13] or other forms of concatenation [14].

Of particular interest to such distributed coding schemes are the determination (and maximization) of the achievable sum-rate as well as the error rate analysis at the destination under realistic transmission conditions. Analysis of the achievable rate for the three-node relay channel was proposed in [2]. Later, [15] defined upper bounds on the achievable rate for a general network with multiple nodes acting both as transmitters and receivers, which take into account the resource allocation among the nodes. The distributed coding scheme studied in this paper is a particular instance of a multiple access relay channel (MARC). The MARC is the theoretical model for a network where multiple sources transmit towards the same destination with the help of a relay. Bounds on the achievable rate for the MARC have been introduced in [16] in the particular case of two sources and later generalized in [14] to an arbitrary number of sources. Both references considered an orthogonal access to the resources but, more importantly, assumed blocks of arbitrary length (infinite-blocklength regime). In this paper, building upon the results of [15], we extend the analysis of [16] to an arbitrary number of sources. For continuous-input MARC with additive white Gaussian noise channels (referred to as AWGN-MARC in this paper), the resulting upper bound on the achievable rate is found to coincide with the upper bound obtained in [14]. Inspired by the finite-blocklength analysis pioneered in [17] for point-to-point communication channels, we then introduce an approximate upper bound on the achievable rate for the AWGN-MARC with finite-length codes and a finite, discrete input alphabet, a scenario which is closer to practical conditions. The latter upper-bound approximation is the first major contribution of this paper. It is ultimately used to find the optimal resource allocation strategy (or equivalently, the relative amount of redundancy to be sent by the relay) that maximizes the sum-rate of the sensor network.

In realistic transmission conditions, the source-to-relay channel is noisy, involving errors at the relay encoding input. Network coding amplifies the relay input errors, thereby causing an error floor at the destination. Solutions to the error propagation problem fall into two main classes. The first method, called threshold-based relaying [18], consists in applying a threshold on the relay demodulation output, then processing and forwarding only reliable decisions. In the second class, referred to as channel-aware detection at the destination in [19], the relay processes and forwards all received samples and the destination exploits full or partial channel state information to manage the error propagation problem. Maximum-likelihood (ML) detection can be applied at the receiver considering all error events [18]–[20]. The prohibitive complexity of such a receiver limits its practical application to network coding schemes such as in [18] and [20], without forward error correcting (FEC) code at the sources and with a small number of sources and relays. When ML detection is too complex, suboptimal solutions can be considered. The probability of error at the relay can be

used to weight [21] or threshold [11] information from the relay. The use of multiple relays [13], [22] can also help to break the error correlation and thus to enhance the decoding performance at the receiver. Closed-form, approximations or bounds on the error probability of network coding schemes are useful either to switch from a non-cooperation mode to a cooperation mode or to define the threshold to be applied in the relay in the case of threshold-relaying-based cooperative schemes. Analyses of the error probability at the destination taking into account error events in the relay are available in [18]–[21] and references therein. In [20], a closed-form bit error rate approximation is given for  $M$ -ary pulse amplitude modulations (PAM) and  $M$ -ary quadrature amplitude modulations (QAM) in the case of a decode-and-forward scheme involving one source and one relay, without FEC. Reference [18] proposes to optimize the threshold in a threshold-relaying cooperative scheme in the case of a two-way relay channel. Assuming binary pulse shift keying (BPSK) modulation, they derive a closed-form expression of the bit error rate in the case of a decode-and-forward protocol as well as approximations in the case of network coding with or without FEC at the sources. In [19], a closed-form expression of the error probability is used to compare different network coding schemes in the case of BPSK. In [21], analytical upper and lower bounds on the error probability are obtained by considering a cooperative scheme with two sources and a relay. The sources use binary linear codes and the relay applies a modulo-2 addition (XOR) of the first source estimated codeword with the second source interleaved estimated codeword. The resulting distributed code is a turbo-like code and the decoding rule requires weighting the information coming from the relay. Bounds on the error probability are obtained and used to numerically optimize the weight applied to the relay-to-destination output. In this paper, we analyse the impact of errors occurring in the relay on the decoding error probability at the destination. In contrast to previous work, no restrictions are placed on either the number of sources or the FEC code used in the relay. For the sake of exposition, the performance analysis is applied to the product code-based cooperative scheme introduced in [11], assuming BPSK transmission over frequency-flat fast-fading Rayleigh channels. The proposed analysis may be however extended to other forms of distributed concatenated codes and to other signal sets with appropriate modifications. For the distributed product code of [11], the complete enumeration of error events as in [18] and [20] cannot be considered and the decoding rule used in [21] is too complex for a practical implementation. We derive a closed-form high-SNR approximation of the error probability under minimum-distance (MD) decoding of the product codeword at the destination. This is the second major contribution of this paper. The analysis shows that the error probability is dominated by single source codeword error events at the relay in the high SNR region. Simulations based on iterative decoding of the distributed product code show that the proposed closed-form expression is reasonably tight at medium to high SNR.

The paper is organized as follows. Section II introduces the notations and the system model. The upper bound on the achievable sum-rate for the AWGN-MARC is introduced in

Section III, in the infinite blocklength regime with continuous input first, and then for finite-length codes with finite, discrete input alphabet. The corresponding upper bound is used to optimize the resource allocation between the sources and the relay in order to maximize the overall throughput. Section IV focuses on the error propagation analysis in the relay and its impact on the derivation of the error probability at the destination in the case of frequency-flat fast-fading Rayleigh channels. Section V deals with simulations for illustration purposes. Section VI concludes the paper and opens the way for discussion and perspectives.

## II. NOTATIONS AND SYSTEM MODEL

We consider a sensor network consisting of  $K$  source nodes with limited power supply, low complexity and low rate, denoted by  $S_k$ , grouped within a cluster using a relay denoted by  $R$ , with better processing capabilities. We assume that the source cluster communicates with a distant destination, denoted by  $D$ , such that the source-to-destination distances are of the same order, and to simplify the setup, we consider them to be equal. The source-to-relay distances are also considered to be equal. The source-to-relay distance is small comparatively to the source-to-destination distance, and the relay-to-destination distance is taken as equal to that of the source-to-destination.

We assume an orthogonal multiple access to the resources, so that each transmission is received without multiple access interference. The relay cannot transmit and receive simultaneously. The  $k$ th source (respectively the relay) is allocated a fraction  $\tau_{S,k}$  (respectively  $\tau_R$ ) of the available resource.  $\{\tau_{S,k}\}$  and  $\tau_R$  are mutually dependent and satisfy  $\sum_{k=1}^K \tau_{S,k} + \tau_R = 1$ . As a consequence, the network has  $K + 1$  transmission nodes. Sources have the same features and use the same small block code of dimension  $k_S$  and length  $n_S$  (denoted by  $\mathcal{C}_S(n_S, k_S)$ ). We thus have  $\tau_{S,k} = \tau_S$  and  $\tau_R = (1 - K\tau_S)$ .

In the following, for the sake of simplicity and without loss of generality, we assume a time domain multiple access. Let  $T$  denote the time duration of a transmission cycle to the destination. On the time interval  $[(k-1)\tau_S T, k\tau_S T]$ , the  $k$ th source transmits a real-valued modulated sequence, denoted by  $\mathbf{x}_{S,k}$  associated to a codeword denoted by  $\mathbf{c}_{S,k} = c_{S,k}(1) \cdots c_{S,k}(n_S)$ . In order to keep notations simple and consistent with Section III, we extend the definition of  $\mathbf{x}_{S,k}$  to the whole interval  $[0, T]$ , imposing  $x_{S,k}(i) = 0$  outside the  $k$ th source transmission interval. The modulated symbols  $x_{S,k}(i)$  in the  $k$ th source transmission time slot are assumed to have unit variance. The overhearing relay stores the corresponding observation in  $\mathbf{y}_R$ . We ignore the processing time in the relay. The relay estimates  $\{\mathbf{c}_{S,k}\}$  from  $\mathbf{y}_R$ . We denote by  $\{\hat{\mathbf{c}}_{S,k}\}$  the detected codewords. In this paper, the relay is assumed to apply soft-decision decoding to detect the source codewords. Soft-decision decoding generates an estimate, which can be a valid codeword, with correlated errors if erroneous. After the detection process, the relay jointly encodes the source codewords. Let  $r$  be the number of resulting redundancy bits to be transmitted to the destination. The relay maps the redundancy binary sequence to a modulated sequence denoted by  $\mathbf{x}_R$ . As for  $x_{S,k}$ , we extend the definition of  $x_R$  to the interval  $[0, T]$ ,

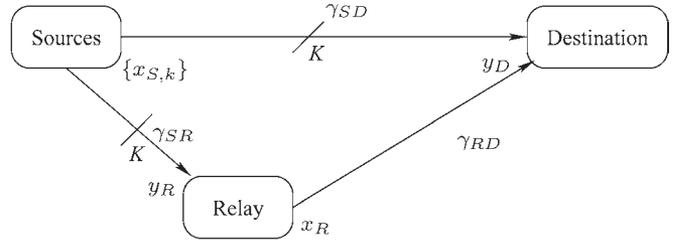


Fig. 1. MARC scheme.

imposing  $x_R(i) = 0$  outside the relay transmission interval. We assume that the variance of modulated symbols  $x_R(i)$  considered within the relay transmission interval is equal to 1. The cooperation lies in the transmission of  $\mathbf{x}_R$ , the result of source codewords combination, from the relay to the destination, on the time slot  $[K\tau_S T, T]$ , ending the transmission process.

We assume non-frequency selective channels. We also consider perfect synchronization (carrier, time). We denote by  $d_{\ell q}$  the distance between nodes  $\ell$  and  $q$  where  $(\ell, q) \in \{S_1, S_2, \dots, S_K, R, D\}$ . We assume  $d_{SR} \ll d_{SD}$ . On the time slot  $[(k-1)\tau_S T, k\tau_S T]$ , the relay receives the noisy  $k$ th source sequence denoted by  $\mathbf{y}_R$ , whose  $i$ th component is given by

$$y_R(i) = \frac{h_{SR}(i)}{(d_{SR})^{\frac{\beta}{2}}} x_{S,k}(i) + n_k(i) \quad (1)$$

where  $\beta$  stands for the path loss exponent,  $h_{SR}(i)$  is a Rayleigh-distributed random variable satisfying  $E[(h_{SR}(i))^2] = \gamma_{SR} d_{SR}^\beta$  and  $n_k(i)$  is an additive white Gaussian noise sample with zero mean and unit variance. The average signal to noise ratio on the source to relay channel is thus equal to  $\gamma_{SR}$ . The destination receives

$$y_D(i) = \frac{h_{SD}(i)}{(d_{SD})^{\frac{\beta}{2}}} x_{S,k}(i) + n(i) \quad (2)$$

in the time slot  $[(k-1)\tau_S T, k\tau_S T]$  and

$$y_D(i) = \frac{h_{RD}(i)}{(d_{RD})^{\frac{\beta}{2}}} x_R(i) + n(i) \quad (3)$$

on the time interval  $[K\tau_S T, T]$ . In (2) and (3),  $h_{SD}(i)$  and  $h_{RD}(i)$  are Rayleigh-distributed random variables satisfying  $E[(h_{SD}(i))^2] = \gamma_{SD} d_{SD}^\beta$  and  $E[(h_{RD}(i))^2] = \gamma_{RD} d_{RD}^\beta$ .  $n(i)$  is an additive white Gaussian noise sample with zero mean and unit variance. The average signal to noise ratio on the source-to-destination channel (resp. relay-to-destination channel) is thus equal to  $\gamma_{SD}$  (resp.  $\gamma_{RD}$ ). The resulting MARC scheme is illustrated in Fig. 1.

*Notations:* Bold lower-case letters denote vectors and bold upper-case letters denote matrices. Given a subset  $\mathcal{S}$  of a set  $\Omega$ , the relative complement of  $\mathcal{S}$  in  $\Omega$  is denoted by  $\mathcal{S}^c$  and is defined by  $\mathcal{S} \cap \mathcal{S}^c = \emptyset$  and  $\mathcal{S} \cup \mathcal{S}^c = \Omega$ . The Hamming distance (respectively weight) is denoted by  $d_H$  (respectively  $w_H$ ).

### III. OPTIMAL TRADE-OFF BETWEEN ACHIEVABLE RATE AND RELAY NETWORK USE PROPORTION FOR THE AWGN-MARC

We first consider a general real-valued AWGN multiple access relay channel. In this setting,  $h_{SR}(i)$  (resp.  $h_{SD}(i)$ ,  $h_{RD}(i)$ ) is constant and equal to  $\sqrt{\gamma_{SR}d_{SR}^\beta}$  (resp.  $\sqrt{\gamma_{SR}d_{SD}^\beta}$ ,  $\sqrt{\gamma_{RD}d_{RD}^\beta}$ ). We want to optimize the system throughput as a function of  $\tau_S$  defined above as the network use proportion per source. As we consider symmetric sources, we denote the rate per source by  $R_S$ . We aim at maximizing the sum rate  $\sum_{k=1}^K R_{S,k} = KR_S$ . We apply the theorems of [15] to derive the rate upper bounds. We first consider the infinite blocklength regime and then extend our results to the finite blocklength regime. Ideal, arbitrary long, capacity-achieving codes are assumed in the first case, whereas optimal (in the sense of minimum error probability for fixed length and rate [17]) finite-length codes are considered in the second scenario.

#### A. Infinite Blocklength Regime

1) *Bound on Achievable Rate in the General Case of a Network With Finite States [15]:* Let us briefly recall Theorem 2 of [15]. We consider a network consisting of  $N$  nodes. The network nodes can transmit and receive, but not within the same network use. Given a node  $k$ , we denote by  $x_k(i)$  and  $y_k(i)$  the  $i$ th transmitted variable and the  $i$ th received variable, respectively.  $x_k(i)$  depends both on all intended messages from node  $k$  towards the other nodes and on all previous received variables  $\{y_k(0), \dots, y_k(i-1)\}$ . A network state corresponds to a valid partitioning of the nodes into a subset of transmitting nodes and a disjoint subset of receiving nodes. Given such a partitioning, referred to by the network state index  $\ell$ , let  $\mathcal{S}$  be the subset of transmitting nodes. Then the relative complement of  $\mathcal{S}$  in  $\{1, \dots, N\}$ , denoted by  $\mathcal{S}^c$ , is the subset of receiving nodes. We denote by  $x_S^{(\ell)}$  (resp.  $y_S^{(\ell)}$ ) the group of transmitted (resp. received) variables from all nodes of subset  $\mathcal{S}$  over all network uses (whose number is assumed to tend to infinity) corresponding to the network state  $\ell$ . Let  $L$  stand for the number of network states. The sequence of states is fixed and known to all nodes. Let  $t_\ell$  stand for the portion of time the network has been used in state  $\ell$ , assuming the network use number tends to infinity. Let  $R_{k,j}$  denote the rate between nodes  $k$  and  $j$ .

*Theorem [15]:* If the information rates  $\{R_{k,j}\}$  are achievable, then there exist some joint probability distributions  $\{p(x_1, x_2, \dots, x_N|\ell)\}$ , where  $\ell$  refers to the network state, such that the sum rate of information transfer between a transmitting node subset  $\mathcal{S}_1$  to a disjoint receiving node subset  $\mathcal{S}_2$  is upper-bounded by

$$\sum_{(k,j) \in \mathcal{S}_1 \times \mathcal{S}_2} R_{k,j} \leq \min_{\mathcal{S}} \sum_{\ell=1}^L t_\ell I(x_S^{(\ell)}; y_{\mathcal{S}^c}^{(\ell)} | x_{\mathcal{S}^c}^{(\ell)}) \quad (4)$$

with  $\mathcal{S} \subset \{1, 2, \dots, N\}$ ,  $\mathcal{S} \cap \mathcal{S}_1 = \mathcal{S}_1$ ,  $\mathcal{S} \cap \mathcal{S}_2 = \emptyset$ .

2) *Application to the AWGN-MARC:* We now adapt the previous theorem to the MARC described in Section II, with the

node set  $\{S_1, S_2, \dots, S_K, R, D\}$ . To derive the upper bounds, two generic cases have to be considered.

(i) (*broadcast channel*) In the broadcast channel, only the  $K$  sources can transmit and the relay and destination receive. Let  $p$  be an integer such that  $1 \leq p \leq K$ . We consider  $\mathcal{S}_1 = \{S_{k_1}, \dots, S_{k_p}\}$  with  $1 \leq k_1 < k_2 < \dots < k_p \leq K$  and  $\mathcal{S}_2 = \{R, D\}$ .

For any  $\mathcal{S} = \{S_{k_1}, S_{k_2}, \dots, S_{k_q}\} \subset \{S_1, S_2, \dots, S_K\}$ , taking into account the independence between source sequences, we have

$$\sum_{\ell} t_\ell I(x_S^{(\ell)}; y_{\mathcal{S}^c}^{(\ell)} | x_{\mathcal{S}^c}^{(\ell)}) = q\tau_S I(x_{S,k}; y_D, y_R). \quad (5)$$

We consider a degraded relay channel:  $(x_{S,1}, \dots, x_{S,k}, \dots, x_{S,K}) \rightarrow (y_R, x_R) \rightarrow y_D$  is a Markov chain. As a consequence,  $I(x_{S,k}; y_D, y_R) = I(x_{S,k}; y_R)$ . Applying (4), we conclude that

$$pR_s \leq p\tau_S I(x_{S,k}; y_R). \quad (6)$$

(ii) (*cooperative multiple access channel*) In the cooperative multiple access channel, the  $K$  sources and the relay transmit towards the destination, which is the single receiving node. Let  $p$  be an integer such that  $1 \leq p \leq K$ . We consider  $\mathcal{S}_1 = \{S_{k_1}, \dots, S_{k_p}, R\}$  with  $1 \leq k_1 < k_2 < \dots < k_p \leq K$  and  $\mathcal{S}_2 = \{D\}$ .

For any  $\mathcal{S} = \{S_{k_1}, \dots, S_{k_q}, R\} \subset \{S_1, \dots, S_K, R\}$ , taking into account the independence between source sequences, we have

$$\sum_{\ell} t_\ell I(x_S^{(\ell)}, y_{\mathcal{S}^c}^{(\ell)} | x_{\mathcal{S}^c}^{(\ell)}) = q\tau_S I(x_{S,k}; y_D) + (1 - K\tau_S)I(x_R; y_D). \quad (7)$$

Applying (4), we conclude that

$$pR_s \leq p\tau_S I(x_{S,k}; y_D) + (1 - K\tau_S)I(x_R; y_D). \quad (8)$$

Taking into account cases (i) and (ii), we obtain the following upper bound:

$$R_s \leq \min \left[ \tau_S I(x_{S,k}; y_R), \tau_S I(x_{S,k}; y_D) + \frac{(1 - K\tau_S)}{K} I(x_R; y_D) \right]. \quad (9)$$

We consider independent and identically distributed (i.i.d.) Gaussian codes [23], proved to be capacity-achieving codes for many transmission scenarii. Using the mutual information expression of such Gaussian codes, we upper-bound the source rate by

$$R_s \leq \min \left[ \frac{\tau_S}{2} \log_2(1 + \gamma_{SR}), \frac{\tau_S}{2} \log_2(1 + \gamma_{SD}) + \frac{(1 - K\tau_S)}{2K} \log_2(1 + \gamma_{RD}) \right] \quad (10)$$

where  $\gamma_{ij}$  denotes the signal to noise ratio between nodes  $i$  and  $j$ .

The optimal trade-off is thus achieved for  $(\tau_S/2)(\log_2(1 + \gamma_{SD}) - \log_2(1 + \gamma_{RD})) + (1/2K)\log_2(1 + \gamma_{RD}) = (\tau_S/2)(\log_2(1 + \gamma_{SR}))$  yielding

$$\tau_S^* = \frac{1}{K} \min \left[ 1, \frac{\log_2(1 + \gamma_{RD})}{\log_2 \left( \frac{(1 + \gamma_{SR})(1 + \gamma_{RD})}{(1 + \gamma_{SD})} \right)} \right] \quad (11)$$

where  $\tau_S^* = (1/K)$  corresponds to the no-cooperation case. According to (11), when the source-to-relay channel is worse than the source-to-destination channel, that is to say  $\gamma_{SR} < \gamma_{SD}$ , the optimum choice corresponds to the no-cooperation case. Otherwise,  $\tau_S^* < (1/K)$  and cooperation is recommended.  $\tau_S^*$  decreases as  $(\gamma_{SR}/\gamma_{SD})$  increases, either when the source-to-relay channel improves or when the source-to-destination channel deteriorates. If  $\tau_S^* = (1/K)$ , we obtain

$$R_s^* = \frac{1}{2K} \log_2(1 + \gamma_{SD}) \quad (12)$$

which is a well-known result for orthogonal multiple access such as TDMA, FDMA or CDMA without cooperation. In the case of cooperation

$$R_s^* \leq \frac{1}{2K} \frac{\log_2(1 + \gamma_{SR}) \log_2(1 + \gamma_{RD})}{\log_2(1 + \gamma_{SR}) + \log_2(1 + \gamma_{RD}) - \log_2(1 + \gamma_{SD})}. \quad (13)$$

This result is consistent with [14].

### B. Approximation for the Finite Blocklength Regime

In the previous subsection, we have considered the maximum achievable rate without taking into account the target error probability and assuming that the blocklength is infinite. In practice, finite-length blocks are used and the maximum achievable rate depends on the blocklength, the target error probability and the channel characteristics. In [17], the correction to the channel capacity is established for the finite blocklength regime as a function of a channel dispersion  $V$ , in the case of binary symmetric channels (BSC), binary erasure channels (BEC), as well as for AWGN channels with unconstrained inputs. If  $C_\infty$  denotes the channel capacity (maximum rate achievable with vanishing low error probability) in the infinite blocklength regime, the maximum rate achievable with a length- $n$  block code and an error probability no larger than  $p_e$  over a real-valued AWGN channel with SNR  $\gamma$  is well-approximated by [17]

$$C_{n,p_e,\gamma} \approx C_\infty(\gamma) - \sqrt{\frac{V(\gamma)}{n}} Q^{-1}(p_e) + \frac{\log n}{2n} \quad (14)$$

where  $V(\gamma) = (\gamma(\gamma + 2)/2(\gamma + 1)^2) \log^2(e)$  is the AWGN channel dispersion and  $Q(\cdot)$  is the well-known complementary cumulative density function of a zero-mean unit-variance Gaussian random variable.

In practice, transmitted symbols belong to a discrete alphabet. To date, computation of the channel dispersion parameter for discrete-input continuous-output channels remains an open issue. In the following, for a  $D$ -ary input channel alphabet, we upper bound the channel capacity as follows:

$$C_{n,p_e,\gamma} \leq \min \left[ \log(D), C_\infty^{AWGN}(\gamma) - \sqrt{\frac{V(\gamma)}{n}} Q^{-1}(p_e) + \frac{\log n}{2n} \right]. \quad (15)$$

Let us go back to the MARC sum rate definition in the finite blocklength regime. In the following,  $C_n(\gamma)$  denotes the maximum achievable rate for a finite blocklength  $n$  with an average SNR  $\gamma$ . As for the infinite blocklength regime, we have to consider the broadcast channel and the cooperative multiple access channel. In the broadcast channel, the  $K$  sources transmit mutually independent codewords of length  $n_S$  towards the relay and the destination. The upper bound given in (15) can be used with  $n = n_S$ . In the cooperative multiple access channel, the sources and the relay cooperate and we consider a network encoding distributed among the sources and the relay. Given a cut set involving  $\ell$  active sources, the network code dimension is  $\ell k_S$  and its length is  $\ell n_S + r$ , where  $r$  is fixed and equal to  $r = (\tau_R/\tau_S)n_S$ . We assume that the sources and the relay share the knowledge of the transmitted distributed codeword. For this particular cut set, the source rate is

$$\tau_S C_{n_S \ell + r}(\gamma_{SD}) + \frac{\tau_R}{\ell} C_{n_S \ell + r}(\gamma_{RD}) \quad (16)$$

with the approximation

$$C_n(\gamma) = \min \left[ \log_2(D), \frac{1}{2} \log_2(1 + \gamma) - \sqrt{\frac{V(\gamma)}{n}} Q^{-1}(p_e) + \frac{\log_2 n}{2n} \right]. \quad (17)$$

We can upper-bound the source rate considering all cut sets

$$R_s \leq \min_{0 \leq \ell \leq K} \left[ \tau_S C_{n_S}(\gamma_{SR}), \tau_S C_{n_S \ell + r}(\gamma_{SD}) + \frac{\tau_R}{\ell} C_{n_S \ell + r}(\gamma_{RD}) \right]. \quad (18)$$

If  $\gamma_{SD} = \gamma_{RD}$ , (18) can be simplified yielding

$$R_s \leq \min_{0 \leq \ell \leq K} \left[ \tau_S C_{n_S}(\gamma_{SR}), \frac{\ell \tau_S + \tau_R}{\ell} C_{n_S \ell + r}(\gamma_{SD}) \right] = \min_{0 \leq \ell \leq K} \left[ \tau_S C_{n_S}(\gamma_{SR}), \frac{1 - (K - \ell) \tau_S}{\ell} \times C_{\frac{n_S}{\tau_S}(1 - (K - \ell) \tau_S)}(\gamma_{SD}) \right]. \quad (19)$$

We can optimize the elementary source time allocation  $\tau_S$  so as to maximize the network throughput. In Section V-A, we determine the value  $r = (\tau_R/\tau_S)n_S$  for which the upper bound on (19) is maximum in two different settings.

## IV. ERROR PROBABILITY ANALYSIS FOR DISTRIBUTED PRODUCT CODES OVER FREQUENCY-FLAT FAST-FADING RAYLEIGH CHANNELS

The source-to-relay channel is not always error-free, yielding errors at the relay encoder input that propagate to its output. The error probability at the relay encoder output depends on the SNR and the detection strategy used in the relay. Whatever the relay decoding strategy, simulations have shown a flattening of the error rate as the SNR increases [22]. In this section, we aim at analyzing this phenomenon by establishing the error probability at the destination under MD-decoding in the presence of errors on the source-to-relay channel. Without loss of generality and for the sake of simplicity, we consider  $\gamma_{SD} = \gamma_{RD}$ .

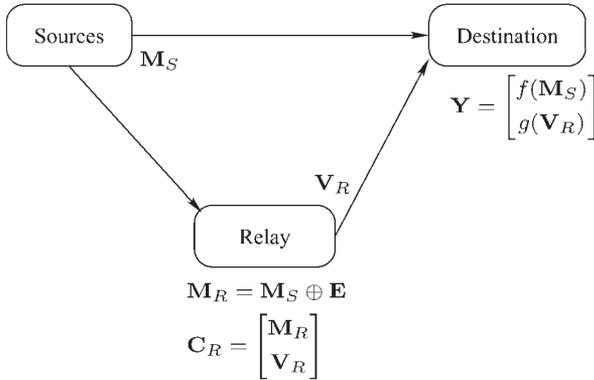


Fig. 2. Distributed coding scheme based on a product code.

### A. Product Code-Based Cooperation Strategy

Sources use the same systematic block code  $\mathcal{C}_S(n_S, k_S)$ . During one cooperative transmission, the  $k$ th source transmits a codeword belonging to the alphabet  $\mathcal{C}_S$ . The relay decodes the corresponding observation and stores the soft-decoded codeword as the  $k$ th row of a  $K \times n_S$  matrix. Once the  $K$  noisy source codewords are processed, the relay encodes the  $n_S$  columns of the matrix. The relay code, denoted by  $\mathcal{C}_R(n_R, k_R = K)$ , is systematic and produces length- $n_R$  codewords columnwise. Thanks to the product code properties, the  $(n_R - k_R)$  last length- $n_S$  rows of the resulting matrix in the relay are codewords belonging to  $\mathcal{C}_S$ . The relay transmits these  $(n_R - k_R)$  codewords to the destination. Collecting the observations coming from the sources and the relay, the destination decodes the equivalent noisy product codeword and thereby simultaneously estimates the  $K$  source messages. The resulting distributed coding scheme and the notations are illustrated in Fig. 2.

### B. Word Error Probability at the Relay Encoder Output

Let us define  $\alpha_i = (h_{SR}(i))/\sqrt{(d_{RD})^\beta \gamma_{SR}}$  and  $p(\alpha_i)$  the conditional bit error probability on the source-to-relay channel, given  $\alpha_i$ . We assume BPSK modulation over a fast fading Rayleigh channel, with perfect channel state information at the receiver.  $\alpha_i$  is a Rayleigh-distributed random variable satisfying  $E[\alpha_i^2] = 1$ . Conditioned on  $\alpha_i$ ,  $p(\alpha_i)$  is given by

$$p(\alpha_i) = \frac{1}{2} \operatorname{erfc} \left( \sqrt{\alpha_i^2 \frac{\gamma_{SR}}{2}} \right). \quad (20)$$

The relay receives  $K = k_R$  noisy source codewords, processes and stores them rowwise, before columnwise encoding. Let us denote by  $\hat{\mathbf{m}}$  the relay encoder input and by  $\mathbf{m}$  the corresponding error-free message. The probability  $\Pr(\hat{\mathbf{m}} \neq \mathbf{m})$  depends on the relay detection strategy. The components of the relay encoder input  $\hat{\mathbf{m}}$  come from independent channels. We denote  $p_{r,i}^{in} = \Pr(\hat{m}_i \neq m_i)$  and we will give its expression later on. Let us denote by  $\mathbf{c}$  (respectively  $\hat{\mathbf{c}}$ ) the codeword associated to  $\mathbf{m}$  (respectively  $\hat{\mathbf{m}}$ ). As we consider systematic block codes,  $\mathbf{c}$  can be split into the information part  $\mathbf{m}$  and the parity check part  $\mathbf{v}$ . Let  $B_R(w, d)$  be the number of codewords of  $\mathcal{C}_R$  of Hamming weight  $d$  resulting from the encoding of a

weight- $w$  message. Let  $d_R$  be the minimum Hamming distance of  $\mathcal{C}_R$ . The word error probability at the relay output,  $p_r^{out}$ , can be written as follows:

$$\begin{aligned} p_r^{out} &= \sum_{d=d_R}^{n_R} \sum_{w=1}^{k_R} \Pr(d_H(\mathbf{c}, \hat{\mathbf{c}}) = d, d_H(\mathbf{m}, \hat{\mathbf{m}}) = w) \\ &= \sum_{d=d_R}^{n_R} \sum_{w=1}^{k_R} \Pr(d_H(\mathbf{c}, \hat{\mathbf{c}}) = d | d_H(\mathbf{m}, \hat{\mathbf{m}}) = w) \\ &\quad \times \Pr(d_H(\mathbf{m}, \hat{\mathbf{m}}) = w). \end{aligned} \quad (21)$$

As the considered block code is linear,  $\Pr(d_H(\mathbf{c}, \hat{\mathbf{c}}) = d | d_H(\mathbf{m}, \hat{\mathbf{m}}) = w) = \Pr(w_H(\mathbf{c} \oplus \hat{\mathbf{c}}) = d | w_H(\mathbf{m} \oplus \hat{\mathbf{m}}) = w)$  where  $w_H$  denotes the Hamming weight. We can thus write

$$\Pr(d_H(\mathbf{c}, \hat{\mathbf{c}}) = d | d_H(\mathbf{m}, \hat{\mathbf{m}}) = w) = \frac{B_R(w, d)}{\sum_{s=d_R}^{n_R} B_R(w, s)} \quad (22)$$

where  $\sum_{s=d_R}^{n_R} B_R(w, s)$  is the number of codewords obtained by encoding weight- $w$  messages.

As the source codewords are independent and the source transmissions are orthogonal, the relay detection results are independent. It follows that  $\Pr(d_H(\mathbf{m}, \hat{\mathbf{m}}) = w) = \binom{k_R}{w} \prod_{j=1}^w p_{r,j}^{in} \prod_{j=w+1}^{k_R} (1 - p_{r,j}^{in})$ , where we assume, without any loss of generality, that the  $w$  errors are located in the first positions. Equation (21) can thus be rewritten as

$$\begin{aligned} p_r^{out} &= \sum_{d=d_R}^{n_R} \sum_{w=1}^{k_R} \binom{k_R}{w} \frac{B_R(w, d)}{\sum_{s=d_R}^{n_R} B_R(w, s)} \\ &\quad \times \prod_{j=1}^w p_{r,j}^{in} \prod_{j=w+1}^{k_R} (1 - p_{r,j}^{in}). \end{aligned} \quad (23)$$

With soft-decision decoding at the relay, denoting by  $A_S(d)$  the number of codewords of  $\mathcal{C}_S$  of Hamming weight  $d$

$$p_{r,i}^{in} \leq \sum_{d=d_S}^{n_S} A_S(d) \frac{d}{2n_S} \operatorname{erfc} \left( \sqrt{\left( \sum_{j=1}^d \alpha_j^2 \right) \frac{\gamma_{SR}}{2}} \right). \quad (24)$$

The average word error probability at the relay encoder output is given by

$$\bar{p}_r^{out} = \sum_{d=d_R}^{n_R} \sum_{w=1}^{k_R} \binom{k_R}{w} \frac{B_R(w, d)}{\sum_{s=d_R}^{n_R} B_R(w, s)} (\bar{p}_r^{in})^w (1 - \bar{p}_r^{in})^{k_R - w} \quad (25)$$

with

$$\begin{aligned} \bar{p}_{r,i}^{in} &\leq \sum_{d=d_S}^{n_S} A_S(d) \frac{d}{n_S} \left[ \frac{1}{2} (1 - \mu_{SR}) \right]^d \\ &\quad \times \sum_{\ell=0}^{d-1} \binom{d-1+\ell}{\ell} \left[ \frac{1}{2} (1 + \mu_{SR}) \right]^\ell \end{aligned} \quad (26)$$

where we have used the mathematical expectation of  $(1/2) \operatorname{erfc}(\sqrt{(\sum_{j=1}^d \alpha_j^2)(\gamma_{SR}/2)})$  available in Chapter 14 of [24]. The average word error probability at the relay decoder output is given by (25) with  $\bar{p}_{r,i}^{in}$  replaced by its expression in (26).

C. Word Error Probability Under Soft-Decision Decoding at the Destination Assuming Soft-Decision Decoding in the Relay

The relay stores the soft-decision decoded source codewords in a  $K \times n_S$  matrix. We denote by  $\mathbf{M}_R = \mathbf{M}_S \oplus \mathbf{E}$  the resulting matrix, where  $\mathbf{M}_S$  corresponds to the  $K$  source codewords stacked rowwise and  $\mathbf{E}$  is the error matrix. We denote by  $\mathbf{C}_S$  (respectively  $\mathbf{C}_R$ ) the product codeword associated to  $\mathbf{M}_S$  (respectively  $\mathbf{M}_R$ ). Let  $\hat{\mathbf{C}}$  be the product codeword estimated by the decoder at the destination. We aim to compute the word error probability under MD-decoding at the destination. It is defined as

$$P_W = \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S) \quad (27)$$

$$= \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S | \mathbf{C}_S) \quad (28)$$

which is justified by the fact that the product code is linear and all codewords are uniformly distributed. An error event occurs in the relay when the soft-decision decoder makes a wrong decision. By considering the relay error events, we can rewrite the pairwise error probability

$$\Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S | \mathbf{C}_S) = \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S, w_H(\mathbf{E}) = 0 | \mathbf{C}_S) + \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S, w_H(\mathbf{E}) > 0 | \mathbf{C}_S). \quad (29)$$

The first term in (29) (error-free soft-decision decoding at the relay) can be also expressed as

$$\Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S, w_H(\mathbf{E}) = 0 | \mathbf{C}_S) = \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S | \mathbf{C}_S, w_H(\mathbf{E}) = 0) \Pr(w_H(\mathbf{E}) = 0). \quad (30)$$

$\Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S | \mathbf{C}_S, w_H(\mathbf{E}) = 0)$  corresponds to the error probability at the soft-decision decoder output at the destination for an error-free source-to-relay channel. For a product code under MD-decoding, it may be upper-bounded by the classical union bound

$$\begin{aligned} & \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S | \mathbf{C}_S, w_H(\mathbf{E}) = 0) \\ & \leq \sum_{d=d_P}^{n_R \times n_S} A_P(d) \mathbb{E} \left[ \frac{1}{2} \operatorname{erfc} \left( \sqrt{\sum_{i=1}^d \alpha_i^2 \frac{\gamma_{SD}}{2}} \right) \right] \\ & = \sum_{d=d_P}^{n_R \times n_S} A_P(d) \left[ \frac{(1 - \mu_{SD})}{2} \right]^d \sum_{\ell=0}^{d-1} \binom{d-1+\ell}{\ell} \\ & \quad \times \left[ \frac{(1 + \mu_{SD})}{2} \right]^\ell \end{aligned} \quad (31)$$

with  $\mu_{SD} = \sqrt{(\gamma_{SD}/2 + \gamma_{SD})}$ ,  $\{\alpha_i\}$  mutually independent and Rayleigh-distributed random variables satisfying  $\mathbb{E}[\alpha_i^2] = 1$ ,  $A_P(d)$  the number of product codewords of weight  $d$  and  $d_P$  the minimum Hamming distance of the product code.  $d_P$  is equal to  $d_S \times d_R$ . An asymptotically tight approximation for high SNR values is

$$\begin{aligned} \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S | \mathbf{C}_S, w_H(\mathbf{E}) = 0) & \approx A_P(d_P) \left[ \frac{(1 - \mu_{SD})}{2} \right]^{d_P} \\ & \times \sum_{l=0}^{d_P-1} \binom{d_P-1+l}{l} \left[ \frac{(1 + \mu_{SD})}{2} \right]^l. \end{aligned} \quad (32)$$

An error event in the relay occurs when at least one source codeword estimate differs from the true codeword. The complement of an error event occurs when each source codeword estimate coincides with the true codeword. The probability of no error at the relay soft-decision decoder output is thus given by

$$\begin{aligned} \Pr(w_H(\mathbf{E}) = 0) & = \prod_{k=1}^K \Pr(\hat{\mathbf{c}}_k = \mathbf{c}_k) \\ & = (1 - \Pr(\hat{\mathbf{c}}_k \neq \mathbf{c}_k))^K \\ & = \left( 1 - \sum_{d=d_S}^{n_S} A_S(d) \epsilon(d, \mu_{SR}) \right)^K \end{aligned} \quad (33)$$

with

$$\epsilon(d, \mu_{SR}) = \left[ \frac{1 - \mu_{SR}}{2} \right]^d \sum_{\ell=0}^{d-1} \binom{d-1+\ell}{\ell} \left[ \frac{1 + \mu_{SR}}{2} \right]^\ell. \quad (34)$$

Let us now consider an error event in the relay to calculate the second term in (29). Let  $\mathcal{C}_P$  stand for the set of product codewords. The probability that the relay transmits at least one wrong codeword and that the soft-decision decoder at the destination delivers at least one source codeword estimate different from the true codeword is obtained by enumerating all product codewords different from the reference product codeword  $\mathbf{C}_S$

$$\begin{aligned} \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S, w_H(\mathbf{E}) > 0 | \mathbf{C}_S) & = \sum_{\substack{\mathbf{C} \in \mathcal{C}_P \\ \mathbf{C} \neq \mathbf{C}_S}} \Pr(\hat{\mathbf{C}} = \mathbf{C}, w_H(\mathbf{E}) > 0 | \mathbf{C}_S). \end{aligned} \quad (35)$$

In [18]–[20], the ML decoder at the destination enumerates all error events. The product code-based cooperative scheme cannot use such a receiver at the destination because its complexity is by far too high. We consider instead a suboptimal soft-decision MD-decoder which minimizes the Euclidean distance between the observation and codewords of  $\mathcal{C}_P$ . It does not take into account the error events in the relay. We denote by  $d_E(\mathbf{C})$  the Euclidean distance between the observation and the codeword  $\mathbf{C}$ . Using the well-known union bound (which is asymptotically tight), we can upper bound  $\Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S, w_H(\mathbf{E}) > 0 | \mathbf{C}_S)$  by upper bounding the pairwise error probabilities

$$\begin{aligned} \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S, w_H(\mathbf{E}) > 0 | \mathbf{C}_S) & \leq \sum_{\substack{\mathbf{C} \in \mathcal{C}_P \\ \mathbf{C} \neq \mathbf{C}_S}} \Pr(d_E(\mathbf{C}) < d_E(\mathbf{C}_S), w_H(\mathbf{E}) > 0 | \mathbf{C}_S). \end{aligned} \quad (36)$$

Let us now consider an error event at the relay output. After soft-decision decoding in the relay, the error matrix consists of  $K$  codewords of length  $n_S$ . Assuming a high-enough SNR on the source-relay channel, the dominant error event at the relay soft decoder output corresponds to the configuration with only one erroneous estimated source codeword ( $\mathbf{M}_R$  and  $\mathbf{M}_S$  have all rows identical except one). Due to the product code properties, a single error columnwise in the information part yields an odd number of errors in the parity check part, lower-bounded by

$d_R - 1$  and upper-bounded by  $n_R - k_R$  (we consider extended codes, Hamming weights are thus even). Denoting by  $\mathcal{E}$  the subset of matrices  $\mathbf{E}$  having only one row different from the zero vector, we can go further into the approximation

$$\Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S, w_H(\mathbf{E}) > 0 | \mathbf{C}_S) \approx \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S, \mathbf{E} \in \mathcal{E} | \mathbf{C}_S). \quad (37)$$

Let  $\mathcal{E}_w$  stand for the subset of error patterns  $\mathbf{E}$  in  $\mathcal{E}$  satisfying  $w_H(\mathbf{E}) = w$ . Considering all subsets  $\mathcal{E}_w$  yields

$$\Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S, \mathbf{E} \in \mathcal{E} | \mathbf{C}_S) = \sum_{w=d_S}^{n_S} \Pr(\mathbf{E} \in \mathcal{E}_w) \sum_{\mathbf{E} \in \mathcal{E}_w} \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S | \mathbf{C}_S, \mathbf{E}). \quad (38)$$

For high SNR values, when the decoder fails, the estimated codeword is either  $\mathbf{C}_R$  the product codeword generated in the relay or a product codeword located at minimum Hamming distance from both  $\mathbf{C}_S$  and  $\mathbf{C}_R$ . Given  $\mathbf{E}$  and  $\mathbf{C}_S$ ,  $\mathbf{C}_R$  is perfectly known and we define  $\mathcal{C}_{\mathbf{C}_S, \mathbf{C}_R}^D$  the subset of codewords  $\mathbf{C}$  of  $\mathcal{C}_P$  satisfying  $d_H(\mathbf{C}, \mathbf{C}_S) = d_H(\mathbf{C}, \mathbf{C}_R) = d_P$ . From this analysis, we approximate the probability that the soft-decision decoder at the destination decides in favor of a codeword  $\hat{\mathbf{C}} \neq \mathbf{C}_S$  given an error pattern  $\mathbf{E}$  in the relay by

$$\Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S | \mathbf{C}_S, \mathbf{E}) \approx \left[ \Pr(\hat{\mathbf{C}} = \mathbf{C}_R | \mathbf{C}_S, \mathbf{C}_R) + \sum_{\mathbf{C} \in \mathcal{C}_{\mathbf{C}_S, \mathbf{C}_R}^D} \Pr(\hat{\mathbf{C}} = \mathbf{C} | \mathbf{C}_S, \mathbf{C}_R) \right]. \quad (39)$$

The error event probability is given by

$$\Pr(\mathbf{E} \in \mathcal{E}_w) = P_e^S(w) \left( 1 - \sum_{w'=d_S}^{n_S} P_e^S(w') \right)^{K-1} \quad (40)$$

where  $P_e^S(w)$  is defined as the probability that a single row of the error matrix  $\mathbf{E}$  in the relay differs from the zero vector and has a weight  $w \neq 0$ . Using the weight enumerating function of the source code  $\mathcal{C}_S$ , we determine  $P_e^S(w)$

$$P_e^S(w) = A_S(w) \left[ \frac{1 - \mu_{SR}}{2} \right]^w \sum_{\ell=0}^{w-1} \binom{w-1+\ell}{\ell} \left[ \frac{1 + \mu_{SR}}{2} \right]^\ell. \quad (41)$$

Let us denote by  $\mathbf{V}_S$  (respectively  $\mathbf{V}_R$ ) the relay-generated parity check matrix associated to  $\mathbf{C}_S$  (respectively to  $\mathbf{C}_R$ ). We now consider the set defined as  $\{\mathbf{C}_R \in \mathcal{C}_P : d_H^r(\mathbf{C}_R) = wd, \mathbf{E} \in \mathcal{E}_w\}$ . It assumes that only one row in the matrix  $\mathbf{M}_R$  be erroneous, which produces a parity check matrix  $\mathbf{V}_R$  with at least  $d - 1$  erroneous rows with error patterns identical to the one in  $\mathbf{M}_R$ . Thanks to the product code properties, whatever the error pattern of the erroneous row in  $\mathbf{M}_R$ , the error pattern (rowwise) will be the same in  $\mathbf{V}_R$ . So the cardinal of  $\{\mathbf{C}_R \in \mathcal{C}_P : d_H^r(\mathbf{C}_R) = wd, \mathbf{E} \in \mathcal{E}_w\}$  is the number of codewords of  $\mathcal{C}_R$  of weight  $d$  generated by the encoding of messages of weight 1. We can thus write

$$|\{\mathbf{C}_R \in \mathcal{C}_P : d_H^r(\mathbf{C}_R) = wd, \mathbf{E} \in \mathcal{E}_w\}| = B_R(1, d). \quad (42)$$

To complete the derivation of (39),  $\Pr(\hat{\mathbf{C}} = \mathbf{C} | \mathbf{C}_S, \mathbf{C}_R)$ , with  $\mathbf{C} \neq \mathbf{C}_S$ , has to be computed. We assume that  $d_H(\mathbf{C}_S, \mathbf{C}_R) = d$  and  $w_H(\mathbf{E}) = w$ . We deduce that  $d_H(\mathbf{V}_S, \mathbf{V}_R) = d - w$ . Let  $\mathbf{C}$  be a codeword of  $\mathcal{C}_P$  different from  $\mathbf{C}_S$ . We denote  $\mathbf{M}$  (respectively  $\mathbf{V}$ ) the matrix obtained by stacking the  $K$  first (respectively  $n_R - K$  last) rows of  $\mathbf{C}$ . We can upper bound  $\Pr(\hat{\mathbf{C}} = \mathbf{C} | \mathbf{C}_S, \mathbf{C}_R)$  by

$$\Pr(\hat{\mathbf{C}} = \mathbf{C} | \mathbf{C}_S, \mathbf{C}_R) \leq \Pr(d_E(\mathbf{C}) < d_E(\mathbf{C}_S) | \mathbf{C}_S, \mathbf{C}_R). \quad (43)$$

Let us define  $\mathcal{I}_s = \{1, \dots, K\}$ ,  $\mathcal{I}_r = \{K + 1, \dots, n_R\}$  and  $\mathcal{J} = \{1, \dots, n_S\}$ . Given  $\mathbf{C}$ ,  $\mathbf{C}_S$  and  $\mathbf{C}_R$ , we define three index subsets:

- $\mathcal{A}_M = \{(i, j) \in \mathcal{I}_s \times \mathcal{J} : M(i, j) \neq M_S(i, j)\}$ , the subset of positions where  $\mathbf{M}$  and  $\mathbf{M}_S$  differ,
- $\mathcal{A}_1 = \{(i, j) \in \mathcal{I}_r \times \mathcal{J} : V(i, j) \neq V_S(i, j), V_R(i, j) \neq V_S(i, j)\}$ , the subset of positions where  $\mathbf{V}$  differs from  $\mathbf{V}_S$  but coincides with  $\mathbf{V}_R$ ,
- $\mathcal{A}_2 = \{(i, j) \in \mathcal{I}_r \times \mathcal{J} : V(i, j) \neq V_S(i, j), V_R(i, j) = V_S(i, j)\}$ , the subset of positions where  $\mathbf{V}$  differs both from  $\mathbf{V}_S$  and  $\mathbf{V}_R$ .

Introducing  $a_M$ ,  $a_1$  and  $a_2$  the respective cardinalities of  $\mathcal{A}_M$ ,  $\mathcal{A}_1$  and  $\mathcal{A}_2$ , we can observe that  $a_M + a_1 + a_2 = d_H(\mathbf{C}, \mathbf{C}_S)$ . We introduce the random variable  $\Phi_{q,\ell}$  defined as

$$\Phi_{q,\ell} = \frac{\left( \sum_{i=q+1}^{\ell} \alpha_i^2 - \sum_{i=1}^q \alpha_i^2 \right)}{\sqrt{\sum_{i=1}^{\ell} \alpha_i^2}} \quad (44)$$

where  $\{\alpha_i\}_{1 \leq i \leq \ell}$  are i.i.d Rayleigh-distributed random variables satisfying  $\mathbb{E}[\alpha_i^2] = 1$ . The second term in (43) is then given by (see Appendix for demonstration)

$$\Pr(d_E(\mathbf{C}) < d_E(\mathbf{C}_S) | \mathbf{C}_S, \mathbf{C}_R) = \mathbb{E} \left[ \frac{1}{2} \operatorname{erfc} \left( \Phi_{a_1, d_H(\mathbf{C}, \mathbf{C}_S)} \sqrt{\frac{\gamma_{SD}}{2}} \right) \right]. \quad (45)$$

This expression applies for any cooperative scheme that uses network coding in the relay. It shows that the error probability under MD-decoding depends both on the weight of the error on the source information and the weight of the error on the redundancy added in the relay. Given  $\ell$ , the expectation of  $\Phi_{q,\ell}$  decreases as  $q$  increases from 1 to  $\ell - 1$  ( $1 \leq a_1 < d_H(\mathbf{C}, \mathbf{C}_S)$  and  $a_M \geq 1$  as  $\mathbf{C} \neq \mathbf{C}_S$ ). As the function  $\operatorname{erfc}$  is monotonically decreasing, the probability given in (45) is maximum for  $a_1$  maximum. The maximum value for  $a_1$  is obtained for a minimum value of  $a_2$ . The minimum value of  $a_2$  is zero, corresponding to an empty set  $\mathcal{A}_2$ , that is to say when the decoded codeword coincides with  $\mathbf{C}_R$ . In this case, (45) equals

$$\Pr(d_E(\mathbf{C}_R) < d_E(\mathbf{C}_S) | \mathbf{C}_S, \mathbf{C}_R) = \mathbb{E} \left[ \frac{1}{2} \operatorname{erfc} \left( \Phi_{d-w, d} \sqrt{\frac{\gamma_{SD}}{2}} \right) \right]. \quad (46)$$

For high-SNR source-relay channel, when the relay soft decoder fails, we can consider that the erroneous codeword is located at Hamming distance  $d_S$  from the source codeword, yielding  $w = d_S$ . In conclusion, the word error probability in

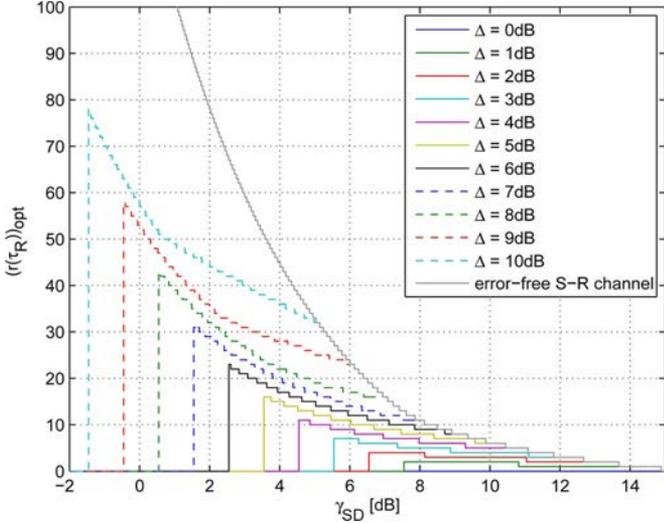


Fig. 3. Optimal redundancy size in the case of  $K = 4$  sources transmitting codewords of length  $n_S = 8$  on the AWGN-MARC.  $\gamma_{SD} = \gamma_{RD}$ ,  $\Delta = 10 \log_{10} \left( \frac{\gamma_{SR}}{\gamma_{SD}} \right)$ . Target bit error probability is  $10^{-6}$ .

the presence of soft-decision decoding errors in the relay is well approximated at high SNR by

$$\begin{aligned} & \Pr(\hat{\mathbf{C}} \neq \mathbf{C}_S, \mathbf{E} \in \mathcal{E} | \mathbf{C}_S) \\ & \approx P_e^S(d_S) \left( 1 - \sum_{w'=d_S}^{n_S} P_e^S(w') \right)^{K-1} \sum_{\substack{d_R \leq d \leq n_R - k_R \\ d \text{ even}}} B_R(1, d) \\ & \times \left[ \mathbb{E} \left[ \frac{1}{2} \operatorname{erfc} \left( \Phi_{d_S(d-1), d_S d} \sqrt{\frac{\gamma_{SD}}{2}} \right) \right] \right. \\ & \left. + \sum_{\mathbf{C} \in \mathcal{C}_{C_S, C_R}^D} \mathbb{E} \left[ \frac{1}{2} \operatorname{erfc} \left( \Phi_{a_1, d_P} \sqrt{\frac{\gamma_{SD}}{2}} \right) \right] \right]. \quad (47) \end{aligned}$$

In summary, a closed-form high-SNR approximation of the error probability under MD-decoding at the destination and assuming soft-decision decoding in the relay is obtained by combining (29), (32), (33), (37), and (47).

## V. SIMULATION RESULTS

### A. AWGN-MARC Achievable Rate Upper-Bound in the Finite-Blocklength Regime

In Figs. 3 and 4 (resp. Figs. 5 and 6), we have plotted the optimal redundancy length defined by  $r(\tau_R) = (\tau_R/\tau_S)n_S$  as well as the upper bound on the achievable rate per source  $R_s$  for different values of  $\Delta = 10 \log_{10} \left( \frac{\gamma_{SR}}{\gamma_{SD}} \right)$ , assuming  $\gamma_{SD} = \gamma_{RD}$  and  $K = 8$  (resp.  $K = 26$ ) sources transmitting codewords of length  $n_s = 8$  (resp.  $n_s = 64$ ). Thus, for a given value of  $(\Delta, \gamma_{SD})$ , there is no point in increasing the redundancy length beyond the value given by the curve. In both cases, when  $\Delta$  is equal to 0 dB, that is to say when the source-relay channel is as noisy as the source-destination channel, cooperation is not recommended. For a given strictly positive value of  $\Delta$ , we observe an SNR threshold depending on  $\Delta$  above which the redundancy length becomes different from

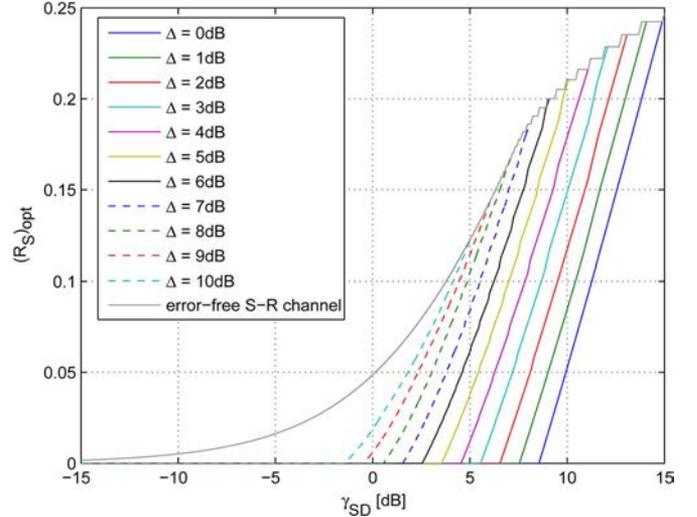


Fig. 4. Optimal rate upper-bound in the case of  $K = 4$  sources transmitting codewords of length  $n_S = 8$  on the AWGN-MARC.  $\gamma_{SD} = \gamma_{RD}$ ,  $\Delta = 10 \log_{10} \left( \frac{\gamma_{SR}}{\gamma_{SD}} \right)$ . Target bit error probability is  $10^{-6}$ .

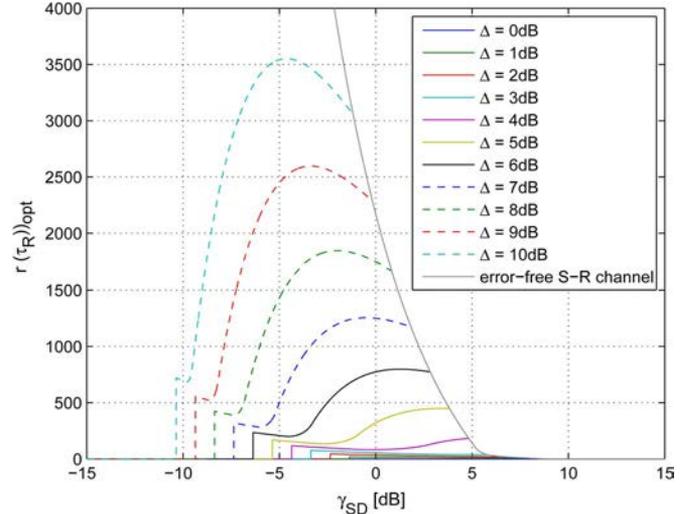


Fig. 5. Optimal redundancy size in the case of  $K = 26$  sources transmitting codewords of length  $n_S = 64$  on the AWGN-MARC.  $\gamma_{SD} = \gamma_{RD}$ ,  $\Delta = 10 \log_{10} \left( \frac{\gamma_{SR}}{\gamma_{SD}} \right)$ . Target bit error probability is  $10^{-6}$ .

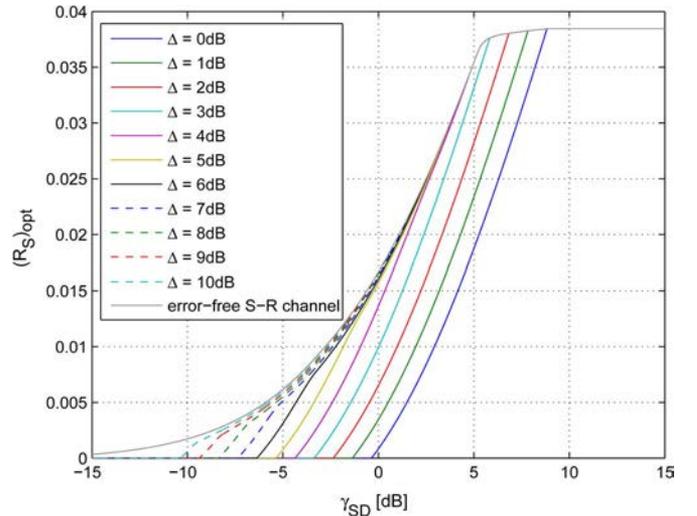


Fig. 6. Optimal rate upper-bound in the case of  $K = 26$  sources transmitting codewords of length  $n_S = 64$  on the AWGN-MARC.  $\gamma_{SD} = \gamma_{RD}$ ,  $\Delta = 10 \log_{10} \left( \frac{\gamma_{SR}}{\gamma_{SD}} \right)$ . Target bit error probability is  $10^{-6}$ .

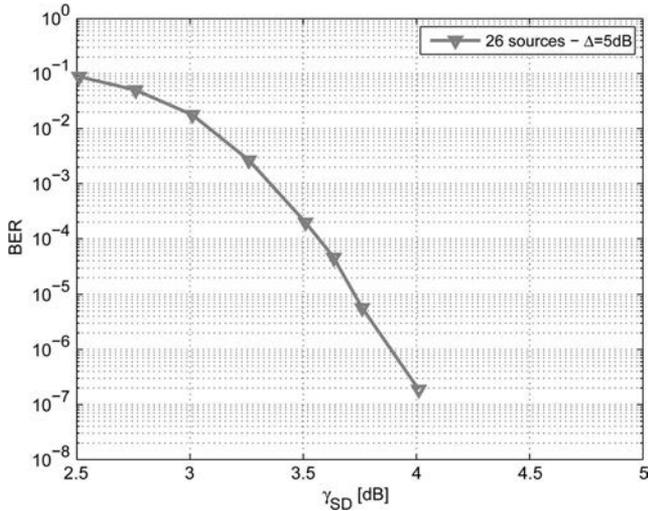


Fig. 7. Simulated bit error rate for  $K = 26$  sources transmitting extended BCH(64,51) codewords on the AWGN-MARC. Relay code: extended BCH code (32,26), BPSK.  $\Delta = 5$  dB.

zero. Indeed (14) is an approximation and may be negative, whereas a capacity value is always positive. In that case the approximation is arbitrarily forced to zero. We can observe that this SNR threshold decreases when  $\Delta$  increases. In the case of  $K = 4$  (Fig. 3), the redundancy length is a monotone decreasing function of  $\gamma_{SD}$ . As  $\gamma_{SD}$  increases, the source-destination link becomes more reliable, and less redundancy is required to ensure the transmission quality, which contributes to a higher data rate (see Fig. 4). The maximum data rate achieved on the  $\gamma_{SD}$  range increases with  $\Delta$ . When  $\gamma_{SD}$  increases, all curves converge to and achieve the same upper bound (which corresponds to an error-free source-relay channel). In the case of  $K = 26$  (Figs. 5 and 6), the behavior is slightly different. We observe three zones defined beyond the SNR threshold defined above. In the first zone (lower values of  $\gamma_{SD}$ ), the optimal value of  $r$  decreases with  $\gamma_{SD}$ . In the second zone (medium values of  $\gamma_{SD}$ ),  $r^*$  increases with  $\gamma_{SD}$ . In the third zone (higher values of  $\gamma_{SD}$ ),  $r^*$  decreases going to zero with  $\gamma_{SD}$ . The first zone gets wider as  $\Delta$  decreases (we observe the same shape as for  $K = 4$ ). In that zone the maximum rate is determined by the source-relay channel and is equal to  $\tau_S C_{n_S}(\gamma_{SR})$  [cf. (19)]. As  $r$  is inversely proportional to  $\tau_S$  and as the rate increases with  $\tau_S$ ,  $r$  is a decreasing function of  $\gamma_{SD}$ . The second zone corresponds to the diversity zone: the destination exploits the coding diversity achieved thanks to the cooperation through joint decoding of the noisy source codewords. The third zone is the data rate increase zone. As the SNR becomes high enough to ensure correct source-destination transmission, the required relay-generated redundancy can be reduced. As for  $K = 4$ , when  $\gamma_{SD}$  increases, all curves converge to and achieve the same upper bound (error-free source-relay channel).

To evaluate the tightness of the proposed upper bound on the achievable rate for the AWGN-MARC, we have plotted in Fig. 7 the bit error rate obtained with 26 sources transmitting extended BCH(64,51) codewords on the AWGN-MARC. The relay uses the extended BCH code (32,26) and  $\Delta = 5$  dB. We have reported the performance in Fig. 8 to compare with the

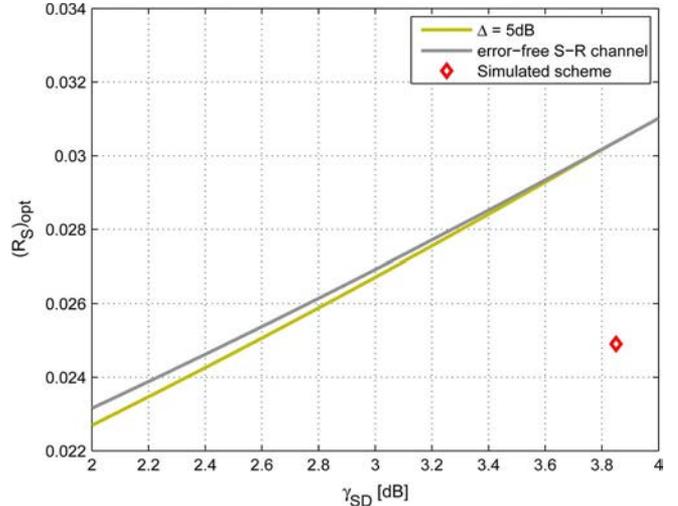


Fig. 8. Comparison of the upper bound on the achievable rate with the practical redundancy size for  $K = 26$  sources transmitting extended BCH(64,51) codewords on the AWGN-MARC. Relay code: extended BCH code (32,26), BPSK.  $\Delta = 5$  dB.

rate predicted with the upper bound (19) as the achievable rate. The practical redundancy size is 12.7% (56 bits) less than that given in Fig. 5 and the average SNR required to achieve a BER of  $10^{-6}$  is around 1.3 dB away from the SNR given by the rate upper bound in Fig. 6. For  $\Delta = 5$  dB and  $\gamma_{SD} = 3.85$  dB, the bit error rate at the relay encoder output is around  $10^{-7}$ , that is to say the source-to-relay channel can be considered error-free and soft-decision decoding at the destination achieves optimal performance. We can thus conclude that the proposed upper bound is suitable for the selection of the redundancy in the relay as a function of  $\gamma_{SD}$  and  $\gamma_{SR}$ .

### B. Theoretical Upper Bound Approximation in the Case of $K = 4$ Sources Assuming Independent Fast Fading Rayleigh Channels

In this subsection, our aim is to validate the upper bound approximation of the word error probability at the destination, which enables to analyse the decoder behavior: at low SNR values, the BER is determined by the conditional error probability given an error-free source-relay channel case (usual product code performance upper bound), whereas at high SNR values, it is limited by the conditional error probability given an erroneous source-relay channel. This approximation results from the fact that the most probable error event at the soft decoder output in the relay consists of a single erroneous row (i.e., only one wrong source codeword) and that the erroneous decoded codeword is either the product codeword generated by the relay (after erroneous soft decoding) or a product codeword located at minimum Hamming distance from the source-equivalent product codeword and the relay-generated product codeword. The weight enumerators of both the product code and its component codes are required to compute the upper bound and are not available in most cases. That is why we consider the extended Hamming code (8,4) for both the source code and relay code. In this case it is easy to obtain all parameters required for the computation of the error probability approximation.

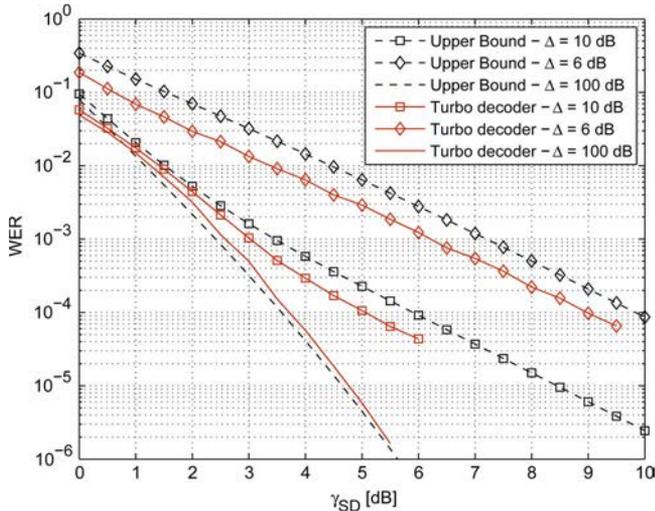


Fig. 9. Theoretical upper bound approximations and simulated frame error rates for  $K = 8$  sources transmitting extended Hamming codewords on a MARC with frequency-flat and fast Rayleigh fading channel. Relay code: extended Hamming code (8,4), BPSK.  $\Delta \in \{6, 10, 100\}$  dB.

We assume that  $K = 4$  sources transmit independent codewords from the extended Hamming (8,4) code. The relay itself applies the same code columnwise:  $n_R = 8$ . We consider independent fast-fading Rayleigh channels and different values of  $\Delta = \frac{\gamma_{SR}}{\gamma_{SD}}$ . The extended Hamming (8,4) code weight enumerator is  $A(d) = 1 + 14d^4 + d^8$  and its input-output weight enumerator is  $B(w, d) = 1 + (4w + 6w^2 + 4w^3)d^4 + w^4d^8$ . In the evaluation of the second term in (39), we restrict ourselves to error events with only one erroneous row that generate codewords  $\mathbf{C}_R$  located at the minimum Hamming distance from the source product codeword  $\mathbf{C}_S$ . As codes are linear, it is tantamount to considering all product codewords of weight  $d_P = 16$  and given one, we have to identify among all minimum Hamming weight product codewords (196 in all), the codewords  $\mathbf{C}$  located at minimum Hamming distance from it to evaluate  $a_1$ . An exhaustive search shows that given such a codeword  $\mathbf{C}_R$ , the histogram of  $a_1$  does not depend on  $\mathbf{C}_R$ . Denoting by  $N_{a_1}$  the number of studied codewords  $\mathbf{C}$ , we obtain that  $N_{a_1} = 6, 12, 6$  for  $a_1 = 4, 6, 8$  respectively. In Fig. 9, we have plotted the simulated frame error rates and the theoretical upper bound approximations for  $\Delta \in \{6, 10, 100\}$  dB. The simulated frame error rate is obtained with the turbo decoding algorithm described in [11]. We can observe that corresponding curves are very close, showing the accuracy of the upper bound approximation. The upper bound approximation coincides with the conditional error probability given an error-free relay decoder output in the low SNR region and with the conditional error probability given an erroneous relay decoding in the high SNR region. This analysis clearly demonstrates that the source-relay channel is the bottleneck of the cooperation scheme based on distributed coding.

## VI. CONCLUSION

In this paper, we have studied a cooperation scheme for sensor networks based on distributed coding for the MARC. The considered scheme involves the transmission of short code-

words from the sources and also from the relay, that ultimately form an overall, powerful concatenated code at the destination. We have obtained an approximate upper-bound on the maximum achievable throughput for the AWGN-MARC under practical assumptions, namely the use of finite-length codes with non-vanishing error probabilities (and thus an erroneous source-relay channel), with discrete, finite signal constellations. The proposed upper bound has been used to optimize the relative amount of redundancy to be sent by the relay for cooperation. Our analysis shows in particular that cooperation is useless when the SNR is too low. At low to medium SNRs, there are two adjacent zones. In the first zone, the source-relay channel is the weakest link and determines the network achievable rate: the redundancy length is a decreasing function of the SNR on the source-destination channel. In the second zone, the source-relay channel is reliable enough and cooperation provides diversity that enhances the overall network quality. In the high SNR region, the source-destination channel conditions get better and the required relay redundancy decreases, thus improving the network throughput. The maximum required relay redundancy also depends on the difference of SNR between the source-relay and source-destination channels: it is all the more important as the difference is large.

Considering a distributed coding scheme based on product codes, we have then analysed the frame error rate at the destination, assuming soft-decision decoding in the relay and taking explicitly into account occasional error events at the relay output. We have proposed an approximation for the word error probability upper bound. The reasoning applied in this paper to derive the upper bound and its approximation can be extended to other distributed coding cooperative schemes. Simulations on an independent fast fading Rayleigh channel have shown that the achieved word error rate is close to the proposed upper-bound approximation. For low SNR values the network performance is determined by the product code performance, whereas for high SNR values it is dominated by error events at the relay output.

Both analyses (network throughput, error probability) yield the same fundamental conclusion: the source-relay channel is the bottleneck of such a cooperation scheme. Several approaches may be considered to cope with this issue. Improving the detection in the relay or exploiting diversity through the use of multiple antennas at the relay, or multiple relays, is a possible area of improvement. Another attractive approach is the definition of cooperation-aware protocols, enabling encoding of error-free detected source codewords only, which requires an additional CRC.

## APPENDIX

We reshape  $\mathbf{y}_D$  by cutting it up into length- $n_S$  packets and stacking the resulting packets rowwise in a matrix denoted by  $\mathbf{Y}_D$ . Its  $K$  first rows correspond to the source codewords and its last  $(n_R - K)$  last rows correspond to the relay-generated codewords. Denoting by  $i$  (respectively  $j$ ) the row (respectively column) index, the components can be written as follows:

$$Y_D(i, j) = \alpha_{SD}(i, j) (2M_S(i, j) - 1) + N(i, j), i \in \mathcal{I}_s \quad (48)$$

$$Y_D(i, j) = \alpha_{RD}(i, j) (2V_R(i, j) - 1) + N(i, j), i \in \mathcal{I}_r \quad (49)$$

where  $\{N(i, j)\}$  are i.i.d. AWGN samples with zero mean and unit variance, whereas  $\{\alpha_{SD}(i, j)\}$  as well as  $\{\alpha_{RD}(i, j)\}$  are i.i.d. Rayleigh random variables satisfying  $E[|\alpha_{SD}(i, j)|^2] = \gamma_{SD}$  and  $E[|\alpha_{RD}(i, j)|^2] = \gamma_{RD}$ . The Euclidean distances have the following expressions:

$$d_E(\mathbf{C}_S) = \sum_{\substack{i \in \mathcal{I}_s \cup \mathcal{I}_r \\ j \in \mathcal{J}}} |N(i, j)|^2 + 4 \sum_{\substack{i \in \mathcal{I}_r \\ j \in \mathcal{J}}} \alpha_{RD}^2(i, j) (V_S(i, j) - V_R(i, j))^2 - 4 \sum_{\substack{i \in \mathcal{I}_r \\ j \in \mathcal{J}}} \alpha_{RD}(i, j) (V_S(i, j) - V_R(i, j)) N(i, j) \quad (50)$$

$$d_E(\mathbf{C}) = \sum_{\substack{i \in \mathcal{I}_s \cup \mathcal{I}_r \\ j \in \mathcal{J}}} |N(i, j)|^2 + 4 \sum_{\substack{i \in \mathcal{I}_s \\ j \in \mathcal{J}}} \alpha_{SD}^2(i, j) (M(i, j) - M_S(i, j))^2 - 4 \sum_{\substack{i \in \mathcal{I}_s \\ j \in \mathcal{J}}} \alpha_{SD}(i, j) (M(i, j) - M_S(i, j)) N(i, j) + 4 \sum_{\substack{i \in \mathcal{I}_r \\ j \in \mathcal{J}}} \alpha_{RD}^2(i, j) (V(i, j) - V_R(i, j))^2 - 4 \sum_{\substack{i \in \mathcal{I}_r \\ j \in \mathcal{J}}} \alpha_{RD}(i, j) (V(i, j) - V_R(i, j)) N(i, j). \quad (51)$$

The assumption  $d_E(\mathbf{C}_S) > d_E(\mathbf{C})$  is thus equivalent to

$$z < \sum_{(i, j) \in \mathcal{A}_1} \alpha_{RD}^2(i, j) - \sum_{(i, j) \in \mathcal{A}_2} \alpha_{RD}^2(i, j) - \sum_{(i, j) \in \mathcal{A}_M} \alpha_{SD}^2(i, j) \quad (52)$$

with  $\mathcal{A}_M$ ,  $\mathcal{A}_1$ , and  $\mathcal{A}_2$  defined in Section IV-C and  $z$  a zero mean Gaussian random variable equal to

$$z = \sum_{(i, j) \in \mathcal{A}_M} \alpha_{SD}(i, j) (M_S(i, j) - M(i, j)) N(i, j) + \sum_{(i, j) \in \mathcal{A}_1 \cup \mathcal{A}_2} \alpha_{RD}(i, j) (V_S(i, j) - V(i, j)) N(i, j). \quad (53)$$

One can easily prove that the conditional variance of  $z$  given  $\{\mathbf{C}_S, \mathbf{C}_R, \mathbf{C}\}$  is equal to

$$\sigma_z^2 | \mathbf{C}_S, \mathbf{C}_R, \mathbf{C} = \left( \sum_{(i, j) \in \mathcal{A}_M} \alpha_{SD}^2(i, j) + \sum_{(i, j) \in \mathcal{A}_1 \cup \mathcal{A}_2} \alpha_{RD}^2(i, j) \right). \quad (54)$$

Considering  $\gamma_{SD} = \gamma_{RD}$ , we obtain that

$$\begin{aligned} & \Pr(d_E(\mathbf{C}) < d_E(\mathbf{C}_S) | \mathbf{C}_S, \mathbf{C}_R) \\ &= \Pr(z < \Psi | \mathbf{C}_S, \mathbf{C}_R, \mathbf{C}) \\ &= \mathbb{E} \left[ \frac{1}{2} \operatorname{erfc} \left( \Phi_{\alpha_1, d_H(\mathbf{C}, \mathbf{C}_S)} \sqrt{\frac{\gamma_{SD}}{2}} \right) \right] \end{aligned} \quad (55)$$

where  $\Psi$  is the random variable defined by

$$\Psi = \sum_{(i, j) \in \mathcal{A}_1} \alpha_{RD}^2(i, j) - \sum_{(i, j) \in \mathcal{A}_M} \alpha_{SD}^2(i, j) - \sum_{(i, j) \in \mathcal{A}_2} \alpha_{RD}^2(i, j) \quad (56)$$

and  $\Phi_{q, \ell}$  is the random variable defined by

$$\Phi_{q, \ell} = \frac{\left( \sum_{i=q+1}^{\ell} \alpha_i^2 - \sum_{i=1}^q \alpha_i^2 \right)}{\sqrt{\sum_{i=1}^{\ell} \alpha_i^2}} \quad (57)$$

with  $\{\alpha_i\}_{1 \leq i \leq \ell}$  i.i.d Rayleigh-distributed random variables satisfying  $E[\alpha_i^2] = 1$ .

#### ACKNOWLEDGMENT

The authors would like to thank PRACOM and Region Bretagne for their financial support.

#### REFERENCES

- [1] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, no. 1, pp. 120–154, 1971.
- [2] T. M. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.
- [3] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, "Hierarchical sensor networks: Capacity bounds and cooperative strategies using the multiple-access relay channel model," in *Proc. IEEE Conf. Sensor Ad Hoc Commun. Netw.*, 2004, pp. 572–584.
- [4] R. Lu, X. Li, X. Liang, and X. Shen, "GRS: The green, reliability, security of emerging machine to machine communications," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 28–35, Apr. 2011.
- [5] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.
- [6] A. Ozgur, O. Leveque, and D. N. C. Tse, "Hierarchical cooperation achieves optimal capacity scaling in Ad Hoc networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3549–3572, Oct. 2007.
- [7] G. Kramer, I. Maric, and R. D. Yates, "Cooperative communications," *Found. Trends Netw.*, vol. 1, no. 3, pp. 271–425, 2007.
- [8] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.
- [9] Y. Yin, R. Pyndiah, and K. Amis, "Multi-source cooperative communications using low-density parity-check product codes," in *Proc. 14th Int. Symp. Wireless Pers. Multimedia Commun.*, 2011, pp. 1–5.
- [10] Z. Xia, Y. Qu, H. Yu, and Y. Xu, "A distributed cooperative product code for multi-source single destination wireless network," in *Proc. 15th Asia-Pacific Conf. on Commun.*, 2009, pp. 736–739.
- [11] R. Pyndiah, F. Guilloud, and K. Amis, "Multiple source cooperative coding using turbo product codes with a noisy relay," in *Proc. IEEE 6th Int. Symp. Turbo Codes Iterative Inf. Process.*, 2010, pp. 98–102.
- [12] T. U. R. Ahsin and S. B. Slimane, "Network coding based on product codes in cooperative relaying," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2010, pp. 1–6.
- [13] E. A. Obiedat, G. Chen, and L. Cao, "Distributed turbo product codes over multiple relays," in *Proc. IEEE 7th Consum. Commun. Netw. Conf.*, 2010, pp. 1–5.
- [14] R. Youssef and A. G. i Amat, "Distributed serially concatenated codes for multi-source cooperative relay networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 253–263, Jan. 2011.
- [15] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Bounds on achievable rates for general multi-terminal networks with practical constraints," in *Proc. IPSN*, 2003, pp. 146–161.
- [16] C. Hausl, "Joint network-channel coding for wireless relay networks," Ph.D. dissertation, Technischen Universitat Munchen, Munchen, Germany, Jun. 2008.
- [17] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel Coding Rate in the Finite Blocklength Regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
- [18] S. L. H. Nguyen, A. Ghayeb, G. Al-Habian, and M. Hasna, "Mitigating error propagation in two-way relay channels with network coding," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3380–3390, Nov. 2010.

- [19] M. Iezzi, M. Di Renzo, and F. Graziosi, "Closed-form error probability of network-coded cooperative wireless networks with channel-aware detectors," in *Proc. IEEE Global Commun. Conf.*, 2011, pp. 1–6.
- [20] M. C. Ju and I.-M. Kim, "ML Performance Analysis of the Decode-and-Forward Protocol in Cooperative Diversity Networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3855–3867, Jul. 2009.
- [21] A. Graell i Amat and I. Land, "Bounds of the probability of error for decode-and-forward relaying with two sources," in *Proc. IEEE 6th Int. Symp. Turbo Codes Iterative Inf. Process.*, 2010, pp. 196–200.
- [22] Y. Yin, R. Pyndiah, and K. Amis, "Performance of turbo product codes on the multiple-access relay channel with relatively poor source-relay links," in *Proc IEEE 13th Int. Workshop Signal Process. Adv. Wireless Commun.*, 2012, pp. 419–423.
- [23] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [24] J. G. Proakis, *Digital communications*. New York, NY, USA: McGraw-Hill, 2000.