Author(s): Abdi Mahmoudaliloo, Younes; Ristaniemi, Tapani

Title: Joint local quantization and linear cooperation in spectrum sensing for cognitive radio networks

Year: 2014

Version:

Please cite the original version:

Abdi Mahmoudaliloo, Y., & Ristaniemi, T. (2014). Joint local quantization and linear cooperation in spectrum sensing for cognitive radio networks. IEEE transactions on signal processing, 62(17), 4349 - 4362. https://doi.org/10.1109/TSP.2014.2330803

Joint Local Quantization and Linear Cooperation in Spectrum Sensing for Cognitive Radio Networks

Younes Abdi, *Student Member, IEEE,* and Tapani Ristaniemi, *Senior Member, IEEE*

***Abstract*—In designing cognitive radio networks (CRNs), protecting the license holders from harmful interference while maintaining acceptable quality-of-service (QoS) levels for the secondary users is a challenge effectively mitigated by cooperative spectrum sensing schemes. In this paper, cooperative spectrum sensing in CRNs is studied as a three-phase process composed of local sensing, reporting, and decision/data fusion. Then, a significant tradeoff in designing the reporting phase, i.e., the effect of the number of bits used in local sensing quantization on the overall sensing performance is identified and formulated. In addition, a novel approach is proposed to jointly optimize the linear soft-combining scheme at the fusion phase with the number of quantization bits used by each sensing node at the reporting phase. The proposed optimization is represented using the conventional false alarm and missed detection probabilities, in the form of a mixed-integer nonlinear programming (MINLP) problem. The solution is developed as a branch-and-bound procedure based on convex hull relaxation, and a low-complexity suboptimal approach is also provided. Finally, the performance improvement associated with the proposed joint optimization scheme, which is due to better exploitation of spatial/user diversities in CRNs, is demonstrated by a set of illustrative simulation results.**

***Index Terms*—Cognitive radio (CR), cooperative spectrum sensing, decision fusion, quantization, non-ideal reporting channel.**

I. Introduction

PROMOTING more efficient use of the radio spectrum as a valuable resource has been of the first priority in many scientific debates and research activities worldwide, see e.g. [1], [2]. Consequently, Cognitive Radio (CR) as the best implementation candidate for the emerging Dynamic Spectrum Management procedures has been the core of most related technical discussions and interactions among various academic, industrial, and regulatory groups specialized in wireless communications all over the world, see [3], [4], and the references therein.

Conceptually, CR is an adaptive communication system which offers the promise of intelligent radios that can learn from and adapt to their environment [5]. As a matter of fact, spectrum sensing is the key element in each CR system and enables its user, commonly referred to as Secondary User (SU), to find transmission opportunities in spectrum resources allocated exclusively to license holders. In this context, the license holders are called Primary Users (PUs) and have the exclusive right of using the spectrum. So the aim of a CR Network (CRN) generally is to achieve radio resources for communication within the spectrum band of the PUs without causing any harmful interference and the spectrum sensing capability enables CRs to detect active PUs and avoid causing interference for them. However, due to impairments like shadowing and multipath fading associated with typical wireless environments, there might be some cases in which not all the CRs are able to detect the PU signal, even though it is present and active. This significant issue is known as the Hidden Node Problem and is a major concern in designing CRNs.

Conventionally, the hidden node problem is mitigated by cooperation among spatially diverse sensing nodes, leading to the concept of Cooperative Spectrum Sensing. As a common design strategy, the cooperative sensing is coordinated by and the overall sensing outcome is generated in a special node called the Fusion Center (FC) which might be considered as a more powerful node, like a base station or an access point. This cooperation of the CR nodes with the FC is generally performed as a three-phase process. In the first phase which we call Local Sensing, each node performs spectrum sensing individually, by using its own built-in sensing scheme. In other words, the nodes listen to their environment to detect the PU signal. Accordingly, the wireless channels between the PU and the sensing nodes are referred to as Listening Channels. In the second phase, called Reporting, the sensing nodes send their local sensing outcomes to the FC through dedicated [6]–[8] or non-dedicated [9], [10] Reporting Channels. Finally, in the third phase, i.e., Fusion, the FC combines the received local sensing outcomes by using a soft-decision (SD) [8], [11]–[16] or hard-decision (HD) [17]–[19] method to decide the presence or absence of the PU.

A. Assumed Architecture: Linear Fusion of Quantized Reports

The architecture assumed in this paper has two main parts, namely, the fusion rule at the FC and the reporting links. We have assumed that the FC performs linear combining on the local test summaries which are reported through non-ideal (i.e., erroneous) digital channels. The main considerations motivating us to adopt this structure are as follows.

It has been shown in [20] that for a distributed detection problem with nonideal analogue communication channels between the distributed nodes and FC, the globally optimal structure is to perform the Likelihood Ratio Test (LRT) both at individual nodes and at the fusion center. However, how to efficiently find the optimal LRT thresholds for individual nodes and for the fusion center is still unknown, see [8],

Y. Abdi and T. Ristaniemi are with the Faculty of Information Technology, University of Jyväskylä, P. O. Box 35, FIN-40014, Jyväskylä, Finland, Tel. +358 40 7214 218 (e-mail:younes.abdi@jyu.fi, tapani.ristaniemi@jyu.fi).

This work has been financially supported by Graduate school in Electronics, Telecommunications, and Automation (GETA) coordinated by Aalto University, Espoo, Finland.

[20]. For the quantized SD case, i.e., when the reporting is performed through nonideal digital links, a solution for optimizing the local quantization levels jointly with the LRT threshold at the FC may or may not exist [21]. Even if an optimal solution exists, the threshold calculations are not trivial, and complex optimization schemes are needed to solve them. The complexity cost further rises as these optimizations have to be done each time the listening or reporting channels change. Moreover, the optimal rules are derived under strict assumptions that may not hold in a practical scenario, resulting in lack of robustness [22]. These difficulties are commonly avoided by assuming a linear fusion scheme [8], [12]–[14] which is the base for our considered architecture. In particular, linear combining is shown in [8] to perform very closely to the optimal LRT method but with much less computational complexity.

Two approaches are commonly used in the literature to model the reporting phase. In the first approach, the reporting links are modeled as analogue, i.e., additive white Gaussian noise (AWGN) channels [8], [12]–[14], which is the simplest assumption leading to analytically-tractable formulations. The alternative approach is based on assuming nonideal digital communication links, i.e., Binary Symmetric Channels (BSC) through which the quantized test summaries are sent to the FC [15], [19], [22]. Digital reporting is more practically appealing since, first, the reporting channel bandwidth is limited in practice, and, second, in many cases, the FC is the access point, base station (BS), or network coordinator which performs a set of resource allocation activities besides decision/data fusion, see e.g., the BS in the IEEE 802.22 standard. These tasks basically require establishment of a set of digital communication links between the FC and CR nodes, which can also be used for the reporting purposes. When modeling the reporting facility as a digital link, the effect of reporting channel impairments on the overall sensing performance is captured into the system model by reporting Bit Error Probability (BEP), as it is a convenient and widely applicable method to model the end-to-end performance of the system including the transmitter, the channel, and the receiver.

B. Related Work

Optimal linear combining is a non-convex problem studied in [8] and [23], where it is broken into several subproblems and the optimal solution is derived through a tedious iterative process which fails to cover all possible cases. As an alternative approach, the authors in [8] have proposed a suboptimal solution based on the so-called Modified Deflection Coefficient (MDC) and showed that this method provides very close results to the ones obtained by the optimal LRT method, but with lower complexity. The MDC approach is also used in other works like [11] and [24] to optimize the detection performance where direct formulation of the false alarm and missed detection probabilities leads to a non-convex problem. In [12], a semidefinite programming approach with a divide-and-conquer process is proposed for the linear combining problem, but the most straightforward and complete solution is developed in [13], which covers all the possible cases. The method in [13] is simple to implement in the sense that it only requires solving a polynomial equation in a single scalar variable over a given interval depending on the system parameters. The effect of reporting channel impairments on the overall sensing performance has been investigated in [19], [22], and [25], where the reporting channels are modeled as binary symmetric channels that cause errors with a certain BEP. Existence of a BEP wall has been demonstrated in these works for both soft- and hard-decision combining schemes, and it is shown that if the BEP of the reporting channel is above the BEP wall value, the constraints on the cooperative detection performance cannot be met at the FC, regardless of the signal quality at the listening channels. Moreover, the authors in [22] compare the performance of the HD- and SD-based fusion methods and illustrate that, in general, the SD significantly outperforms the HD scheme when nonideal reporting channels are considered. In [26], a seesaw analogy for distributed detection based on Rayleigh-faded quantized reports is introduced and evaluated. The fusion rule in [26] is based on assigning weights to the quantization levels rather than to reporting nodes.

C. Contribution

In this paper, we jointly optimize the reporting and fusion phases of the cooperative spectrum sensing in CRNs. In particular and different from the previous works, we optimize the linear soft-combining at the FC considering the effect of reporting channel impairments in designing the quantization scheme used at the sensing nodes. In our novel design, we identify, formulate, and take into account a tradeoff which is described in the following paragraph.

For a given set of radio resources dedicated to the reporting phase in terms of transmission power and bandwidth, increasing the number of quantization bits influences the overall sensing performance in two opposite directions. Specifically, on one side, increasing the number of quantization levels leads to a better quantization process and consequently, lowers the quantization errors affecting the reported local sensing outcomes, improving the cooperation performance. On the other side however, increasing the quantization bits raises the BEP induced by the reporting channels and reduces the received sensing outcomes quality at the FC, degrading the overall sensing performance.

We derive exact as well as practical approximate relationships describing the joint impact of the quantization, reporting, and linear fusion processes. Therefore, the spatial/user diversities regarding the listening and reporting channels are considered in a more comprehensive optimization approach, promising a better overall sensing performance.

The rest of the paper is organized as follows. In Section II, general modeling assumptions and details of the CRN considered are introduced. In Section III, the overall structure of the proposed joint optimization problem is presented, and its detailed mathematical formulations are derived. In Section IV, the MINLP problem is formally constructed, and its solution procedures are developed. The effectiveness of the proposed joint optimization is demonstrated through simulation results

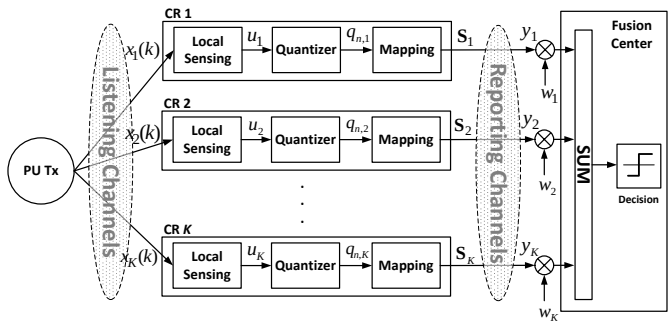

Fig. 1. Basic configuration of cooperative spectrum sensing including the listening and reporting channels, PU, SUs, and FC.

in Section V, followed by concluding remarks provided in Section VI.

II. System Model

A CRN with K sensing nodes is considered. These nodes cooperatively sense the radio spectrum to find temporal and/or spatial vacant bands for their data communication. Fig. 1 shows the basic configuration and major elements in a CRN exploiting cooperative spectrum sensing.

A. Local Sensing and Quantization

In our adopted model, the kth sample of the received PU signal at the ith CR node is represented as

$$\begin{cases} x_i(k) = \nu_i(k), & \mathcal{H}_0 \\ x_i(k) = h_i s(k) + \nu_i(k), & \mathcal{H}_1 \end{cases} \tag{1}$$

where $\mathcal{H}_1$ and $\mathcal{H}_0$ denote the hypotheses representing the presence or absence of the PU, respectively. $s(k)$ denotes the signal transmitted by the PU and $x_i(k)$ is the received signal by the ith SU. h_i is the listening channel block fading gain. Listening channel gains are assumed to be independent circularly-symmetric Gaussian random variables. $\nu_i(k)$ denotes the circularly-symmetric zero-mean AWGN at the CR sensor receiver, i.e., $\nu_i(k) \sim \mathcal{CN}(0, \sigma^2_{\nu_i})$. Without loss of generality, $s(k)$ and $\{\nu_i(k)\}$ are assumed to be independent of each other.

CR node i, $i = 1, ..., K$ performs spectrum sensing by using its built-in sensor (which can be of any common types like Energy Detection (ED), Cyclostationary Detection (CSD), etc.) to derive a local test statistic u_i, and then uses the following quantization rule to map it on a bit sequence of length d_i

$$\psi_i(u_i) = q_{n,i} \quad \text{if} \quad t_{n,i} \leqslant u_i < t_{n+1,i} \tag{2}$$

where $\psi_i(\cdot)$ denotes the quantization process at the ith SU, $q_{n,i}$, $n = 1, ..., 2^{d_i}$ is its nth quantization level, while $t_{n,i}$ and $t_{n+1,i}$ denote the corresponding boundaries.

Let $\mathbf{d} \triangleq \left[d_1, ..., d_K\right]^T$ denote the number of quantization bits used in all sensing nodes. As described in [22] several quantization methods for signal detectors can be considered here such as Maximum Output Entropy (MOE) quantization and Minimum Average Error (MAE) quantization. Without loss of generality, we have considered uniform and MOE quantization schemes in this paper. We describe the uniform quantization method in the following. For details of the MOE quantization, see [22].

In the uniform quantization incorporated in the ith sensing node, the range covered by the quantization levels is $(\mu_i - m\sigma_i, \mu_i + m\sigma_i)$, where m is determined by the Chebyshev inequality such that $\Pr\{|u_i - \mu_i| \geq m\sigma_i\} \leq \frac{1}{m^2}$. This coverage range is then divided into 2^{d_i} equally-spaced levels, whose boundaries are denoted by

$$t_{n,i} = \mu_i + m\sigma_i \left(\frac{2(n-1)}{2^{d_i} - 1} - 1 \right) \tag{3}$$

The quantization level $q_{n,i}$ lies in the middle of $t_{n-1,i}$ and $t_{n,i}$, i.e.,

$$q_{n,i} = \mu_i + m\sigma_i \left(\frac{2n-1}{2^{d_i} - 1} - 1 \right). \tag{4}$$

And the conditional probability of having level $q_{n,i}$ at the ith quantizer output is

$$\Pr\{\psi_i(u_i) = q_{n,i} | \mathcal{H}_j\} = \int_{t_{n-1,i}}^{t_{n,i}} f_{u_i}(x|\mathcal{H}_j) dx \tag{5}$$

where $f_{u_i}(\cdot|\mathcal{H}_j)$ denotes the probability density function (pdf) of u_i conditioned on $\mathcal{H}_j$, $j = 0, 1$.

The generated reporting bit sequences are then transmitted to the FC through the reporting channel in an orthogonal manner. The effect of reporting channel impairments on the transmitted bit sequences of the ith CR node is modeled as a BEP denoted by $P_{\text{b},i}$. The reporting channel is assumed to affect each node's transmitted reporting bit sequence independently. Moreover, errors introduced on different bits in a transmitted reporting sequence by the reporting channel are assumed to be independent and identically distributed (i.i.d). Therefore, the received quantized test statistics at the FC, y_i, $i = 1, ..., K$ are independent discrete random variables whose probability mass functions (pmf) can be represented (for $j = 0, 1$) as [22]

$$\Pr\{y_i = q_{n,i} | \mathcal{H}_j\} = \sum_{k=1}^{2^{d_i}} P_{\text{b},i}^{D_{n,k}} \left(1 - P_{\text{b},i}\right)^{d_i - D_{n,k}} \Pr\{\psi_i(u_i) = q_{k,i} | \mathcal{H}_j\} \tag{6}$$

where $D_{n,k}$ is the Hamming distance between bit sequences corresponding to levels $q_{n,i}$ and $q_{k,i}$.

B. Mapping and Bit Sequences

In analyzing the system behavior, it is worth considering the BSC effect on the reported bit sequences and relate it to the received test summary statistics. Therefore, focusing on transmitted and received bit strings in the reporting phase, we model the effect of the BSC using the eXclusive OR (XOR) operator as

$$r_i = s_i \oplus e_i \tag{7}$$

where d_i-bit (scalar) random variables s_i, r_i, and e_i denote the sent and received bit sequences and error caused by the

BSC, respectively. If we denote the value of s_i associated with the nth quantization level (i.e., $q_{n,i}$) by $s_{n,i}$, the following invertible mapping describes the correspondence between the quantization levels and the bit sequences

$$\begin{cases} \Gamma : \{1, 2, ..., 2^{d_i}\} \rightarrow \{0, 1, ..., 2^{d_i} - 1\} \\ s_{n,i} = \Gamma(n) \end{cases} \tag{8}$$

In other words, $s_i = \Gamma(n)$ if and only if $\psi_i = q_{n,i}$, or equivalently, $s_i = n$ if and only if $\psi_i = q_{\Gamma^{-1}(n),i}$. Therefore, the pmf of s_i can be expressed as (for $n = 0, 1, ..., 2^{d_i} - 1$, and $j = 0, 1$)

$$P_{s_i|\mathcal{H}_j}(n) \triangleq \Pr\{s_i = n|\mathcal{H}_j\} = \Pr\left\{\psi_i = q_{\Gamma^{-1}(n),i}|\mathcal{H}_j\right\}. \tag{9}$$

Without loss of generality, we have assumed the same mapping process for all CR nodes.

Given the BEP $P_{b,i}$, each bit in the random variable e_i follows the Bernoulli distribution. Consequently, the pmf of e_i is derived as (for $n = 0, 1, ..., 2^{d_i} - 1$)

$$P_{e_i}(n) = P_{b,i}^{w_H(n)} (1 - P_{b,i})^{d_i - w_H(n)} \tag{10}$$

where $w_H(n)$ denotes the Hamming weight of the binary representation of n.

In order to derive the pmf of r_i, we use the fact that the assumed reporting channel contamination does not depend on the reported bit sequence s_i, nor the behavior of the PU. Therefore,

$$\begin{aligned} P_{r_i|\mathcal{H}_j}(n) &= \sum_{k=0}^{2^{d_i}-1} \Pr\{r_i = n|e_i = k|\mathcal{H}_j\} \Pr\{e_i = k\} \\ &= \sum_{k=0}^{2^{d_i}-1} \Pr\{s_i = n \oplus k|e_i = k|\mathcal{H}_j\} \Pr\{e_i = k\} \\ &= \sum_{k=0}^{2^{d_i}-1} P_{s_i|\mathcal{H}_j}(n \oplus k) P_{b,i}^{w_H(k)} (1 - P_{b,i})^{d_i - w_H(k)}. \end{aligned} \tag{11}$$

C. Reporting Channel Error

Assuming a general M-ary modulation for the reporting channel, the reporting BEP can be expressed as

$$P_{\mathrm{b},i} = c_M Q\left(\sqrt{c'_M \gamma_{\mathrm{r},i}}\right) \tag{12}$$

where $Q(x) \triangleq \int_x^\infty \exp(-t^2/2)dt/\sqrt{2\pi}$ is the Q-function, c_M and c'_M are two constants determined by the modulation type and $\gamma_{\mathrm{r},i}$ is the reporting link signal-to-noise ratio (SNR). The reporting SNR depends on the number of bits used in the reporting bit sequence d_i as

$$\gamma_{\mathrm{r},i} = \frac{|h_{\mathrm{r},i}|^2 E_{\mathrm{r}}}{N_0 d_i log_2 M} \tag{13}$$

where $h_{\mathrm{r},i}$, E_{r}, and N_0 denote the reporting channel gain, reporting signal energy, and noise power spectral density, respectively. Therefore, the reporting BEP is a continuous function of d_i

$$P_{\mathrm{b},i} = c_M Q\left(\sqrt{\frac{c''_M}{d_i}}\right) \tag{14}$$

where $c''_M = \frac{c'_M |h_{\mathrm{r},i}|^2 E_r}{N_0 log_2 M}$.

D. Linear Combining

Linear combining is performed at the FC, meaning that the global test statistic y_c is constructed as a weighted sum of the received quantized levels, i.e.,

$$y_{\mathrm{c}} = \sum_{i=1}^{K} w_i y_i = \mathbf{w}^T \mathbf{y} \tag{15}$$

where $\mathbf{w} \triangleq [w_1, ..., w_K]^T$ and $\mathbf{y} \triangleq [y_1, ..., y_K]^T$.

Finally, y_c is compared against a predefined threshold ξ to decide the presence or absence of the PU, i.e.,

$$\begin{cases} \mathcal{H}_1, & y_{\mathrm{c}} \geq \xi \\ \mathcal{H}_0, & y_{\mathrm{c}} < \xi \end{cases} \tag{16}$$

The detector performance is commonly measured using two probabilities, namely the probability of false alarm $P_{\mathrm{fa}} = \Pr\{y_{\mathrm{c}} \geq \xi|\mathcal{H}_0\}$ and the probability of missed detection $P_{\mathrm{md}} = \Pr\{y_{\mathrm{c}} < \xi|\mathcal{H}_1\}$. Both false alarm and missed detection probabilities depend on the probability distribution of the global test statistic y_c, which can be derived as a convolution of the pmfs of K independent random variables $\{y_i\}_{i=1}^K$, i.e.,

$$p(y_{\mathrm{c}}) = p(y_1/w_1) * ... * p(y_K/w_K) \tag{17}$$

where $p(\cdot)$ and $*$ stand for pmf and convolution, respectively.

III. Problem Formulation

Our problem is to jointly optimize the reporting and fusion phases. Specifically, the goal is to jointly optimize $\mathbf{w}$ and $\mathbf{d}$ to achieve the best cooperative sensing performance. We determine the weighting vector $\mathbf{w}$ at the FC and $\mathbf{d}$ used by the sensing nodes through jointly considering the effects of both the listening and reporting channels while taking into account the significant tradeoff explained earlier in specifying the optimal number of quantization bits.

A. Problem Structure

We formulate our proposed optimization based on minimizing the missed detection probability subject to an upper bound on the false alarm probability

$$\begin{aligned} &\min_{\mathbf{w},\mathbf{d}} P_{\mathrm{md}} \quad \text{(P1)} \\ &\text{s.t.} \quad P_{\mathrm{fa}} \leq \alpha \end{aligned}$$

where α is the given upper limit on the false alarm probability.

According to the Central Limit Theorem (CLT), if K is large enough, we can assume a Gaussian distribution for y_c. In Appendix I, we have shown that Lyapunov's CLT condition [27] holds for y_is. Consequently, the false alarm and missed detection probabilities can be expressed in closed form as

$$P_{\mathrm{fa}} = Q\left(\frac{\xi - \boldsymbol{\mu}_{\mathcal{H}_0}^T \mathbf{w}}{\sqrt{\mathbf{w}^T \boldsymbol{\Sigma}_{\mathcal{H}_0} \mathbf{w}}}\right) \tag{18}$$

$$P_{\mathrm{md}} = 1 - Q\left(\frac{\xi - \boldsymbol{\mu}_{\mathcal{H}_1}^T \mathbf{w}}{\sqrt{\mathbf{w}^T \boldsymbol{\Sigma}_{\mathcal{H}_1} \mathbf{w}}}\right) \tag{19}$$

where (for $j = 1, 2$) $\boldsymbol{\mu}_{\mathcal{H}_j} \triangleq \mathrm{E}[\mathbf{y}|\mathcal{H}_j]$ and $\boldsymbol{\Sigma}_{\mathcal{H}_j} \triangleq \mathrm{E}\left[\mathbf{y}\mathbf{y}^T|\mathcal{H}_j\right] = \mathrm{diag}(\sigma^2_{y_1|\mathcal{H}_j}, ..., \sigma^2_{y_K|\mathcal{H}_j})$. We have found through numerical evaluations that Gaussian distribution fits well for $K \geq 5$.

Now, if we eliminate ξ in Eqs. (18) and (19) by considering a target false alarm probability $P_{\text{fa}} = \alpha$, (P1) is converted to

$$\max_{\mathbf{w},\mathbf{d}} Q\left(\frac{Q^{-1}(\alpha)\sqrt{\mathbf{w}^T\boldsymbol{\Sigma}_{\mathcal{H}_0}\mathbf{w}} - \mathbf{a}^T\mathbf{w}}{\sqrt{\mathbf{w}^T\boldsymbol{\Sigma}_{\mathcal{H}_1}\mathbf{w}}}\right) \quad \text{(P2)}$$

where $Q^{-1}(\cdot)$ is the functional inverse of the Q-function, $\mathbf{a} \triangleq [a_1, ..., a_K]^T \triangleq \boldsymbol{\mu}_{\mathcal{H}_1} - \boldsymbol{\mu}_{\mathcal{H}_0}$ and, for $i = 1, ..., K$, we have $a_i \triangleq \mathrm{E}[y_i|\mathcal{H}_1] - \mathrm{E}[y_i|\mathcal{H}_0]$. a_i and $\sigma^2_{y_i|\mathcal{H}_j}$ are related to the number of quantization bits d_i through the total probability theorem as

$$\begin{aligned} a_i = &\sum_{n=1}^{2^{d_i}} q_{n,i} \sum_{k=1}^{2^{d_i}} P_{\text{b},i}^{D_{n,k}} (1 - P_{\text{b},i})^{d_i - D_{n,k}} \\ &\times \left[\Pr(\psi_i = q_{k,i}|\mathcal{H}_1) - \Pr(\psi_i = q_{k,i}|\mathcal{H}_0)\right] \end{aligned} \tag{20}$$

$$\begin{aligned} &\sigma^2_{y_i|\mathcal{H}_j} = \\ &\sum_{n=1}^{2^{d_i}} q^2_{n,i} \sum_{k=1}^{2^{d_i}} P_{\text{b},i}^{D_{n,k}} (1 - P_{\text{b},i})^{d_i - D_{n,k}} \Pr(\psi_i = q_{k,i}|\mathcal{H}_j) - \\ &\left[\sum_{n=1}^{2^{d_i}} q_{n,i} \sum_{k=1}^{2^{d_i}} P_{\text{b},i}^{D_{n,k}} (1 - P_{\text{b},i})^{d_i - D_{n,k}} \Pr(\psi_i = q_{k,i}|\mathcal{H}_j)\right]^2. \end{aligned} \tag{21}$$

It is worth noting that the Hamming distance $D_{n,k}$ in (20) and (21) is a complicated term which depends on the number of quantization bits d_i, as well as on the mapping process Γ between the quantization levels and the reported bit sequences. We use two approaches to derive explicit formulas describing the effect of d_i on the desired test statistics at the FC. In the first approach, two simplifying but practical assumptions considering the mapping process and reporting BSC are used to derive approximate relationships. In the second approach, the quantization and mapping processes are considered separately, and exact formulas for the desired statistics are derived in general case. These approaches follow as the next two subsections.

B. Gray Coding and Reliable Reporting

We first assume that Gray coding is used to map the reporting bit sequences to their corresponding quantization levels. In other words, we assume that the bit sequences representing the adjacent quantization levels, differ only in one bit. Secondly, we assume a reliable reporting channel, i.e., small $P_{b,i}$s. Consequently, we neglect the cases in which more than one bit in a sequence arrive erroneously at the FC. Then, we use the following lemma which relates the bit sequences with one-bit distance in a Gray-encoded mapping mechanism.

Lemma 1) Let $s_{n,i}$, $n = 1, ..., 2^{d_i}$ denote the d_i-bit Gray-coded bit string corresponding to $q_{n,i}$. By flipping the l_eth bit in $s_{n,i}$, it turns into $s_{k,i}$ where k can be derived as a function of n and l_e as

$$k(n, l_e) = n + 2^{l_e} - 2\text{mod}(n - 1, 2^{l_e}) - 1 \tag{22}$$

Proof: Please refer to Appendix II. ■

Hence, we know that if $q_{n,i}$ (whose bit string is $s_{n,i}$) is sent by the ith CR node and the reporting channel changes only the l_eth element of the reported bit sequence, then $q_{k(n,l_e),i}$ will be received at the FC. In addition, by setting $l_e = 0$, (22) yields $k = n$, i.e., no change in the code index. So we can consider $l_e = 0$ for the error-free bit sequences which arrive at the FC. *Lemma 1* and the aforementioned assumptions lead to the following simplified forms of a_i and $\sigma^2_{y_i|\mathcal{H}_j}$

$$\begin{aligned} a_i \approx &\sum_{n=1}^{2^{d_i}} q_{n,i} \sum_{l_e=0}^{d_i} P_{\text{b},i}^{1-\delta_{l_e,0}} (1 - P_{\text{b},i})^{d_i + \delta_{l_e,0} - 1} \\ &\times \left[\Pr(\psi_i = q_{k(n,l_e),i}|\mathcal{H}_1) - \Pr(\psi_i = q_{k(n,l_e),i}|\mathcal{H}_0)\right] \end{aligned} \tag{23}$$

where $\delta_{l_e,0}$ equals 1 when $l_e = 0$ and 0 otherwise. $\sigma^2_{y_i|\mathcal{H}_j}$ has been expressed on top of the following page.

C. Bit-by-Bit Considerations

Now we proceed with the second approach to derive exact relations for the desired statistics in general case. We relate the bit-sequence interpretations to our detector optimization by considering a general mapping process. For simplicity, first assume a linear mapping scheme, i.e.,

$$s_{n,i} = n - 1, \quad n = 1, ..., 2^{d_i}. \tag{25}$$

Extension to the general case will be considered later. Using uniform quantization (4) at the CR nodes, the quantization levels correspond to the bit sequences according to

$$q_{n,i} = \mu_i + m\sigma_i \left(\frac{2s_{n,i} + 1}{2^{d_i} - 1} - 1\right). \tag{26}$$

This correspondence can be expressed in terms of random variables ψ_i and s_i as

$$\psi_i = \mu_i + m\sigma_i \left(\frac{2s_i + 1}{2^{d_i} - 1} - 1\right). \tag{27}$$

Thus, at the FC, the relation between the received bit sequences and quantization levels is

$$y_i = \mu_i + m\sigma_i \left(\frac{2r_i + 1}{2^{d_i} - 1} - 1\right) \tag{28}$$

which indicates that we can derive the desired statistics, i.e., a_i and $\sigma^2_{y_i|\mathcal{H}_j}$, in terms of the received code statistics as

$$a_i = \frac{2m\sigma_i}{2^{d_i} - 1} \left(E[r_i|\mathcal{H}_1] - E[r_i|\mathcal{H}_0]\right) \tag{29}$$

$$\sigma^2_{y_i|\mathcal{H}_j} = \left(\frac{2m\sigma_i}{2^{d_i} - 1}\right)^2 \sigma^2_{r_i|\mathcal{H}_j}. \tag{30}$$

These moments are calculated using (11) as (for $j = 0, 1$)

$$E[r_i|\mathcal{H}_j] = \sum_{n=0}^{2^{d_i}-1} n P_{r_i|\mathcal{H}_j}(n) \tag{31}$$

$$\sigma^2_{r_i|\mathcal{H}_j} = \sum_{n=0}^{2^{d_i}-1} n^2 P_{r_i|\mathcal{H}_j}(n) - \left[\sum_{n=0}^{2^{d_i}-1} n P_{r_i|\mathcal{H}_j}(n)\right]^2. \tag{32}$$

$$\sigma^2_{y_i|\mathcal{H}_j} \approx \sum_{n=1}^{2^{d_i}} q^2_{n,i} \sum_{l_e=0}^{d_i} P_{\mathrm{b},i}^{1-\delta_{l_e,0}} (1-P_{\mathrm{b},i})^{d_i+\delta_{l_e,0}-1} \mathrm{Pr}(\psi_i = q_{k,i}|\mathcal{H}_j) - \left[\sum_{n=1}^{2^{d_i}} q_{n,i} \sum_{l_e=0}^{d_i} P_{\mathrm{b},i}^{1-\delta_{l_e,0}} (1-P_{\mathrm{b},i})^{d_i+\delta_{l_e,0}-1} \mathrm{Pr}(\psi_i = q_{k,i}|\mathcal{H}_j) \right]^2 \tag{24}$$

To clearly recognize the role of d_i in the derived statistics, we now focus on the pmf of the received bit sequences, i.e., $P_{r_i|\mathcal{H}_j}(n)$. More specifically, we reconsider (11) from a bit-by-bit perspective to deal with the XOR operator and the Hamming weight $w_H(\cdot)$ by using the following lemma.

Lemma 2) For two d_i-bit integers n, and k, if $w_H(k) = n_e$, and $n_e \neq 0$ then

$$n \oplus k = g(n, k_1, k_2, ..., k_{n_e}) \tag{33}$$

where

$$\begin{aligned} g(n, k_1, k_2, ..., k_{n_e}) \triangleq\ & n + [2u(0.5 - b_{k_1}(n)) - 1]\, 2^{k_1-1} \\ & + [2u(0.5 - b_{k_2}(n)) - 1]\, 2^{k_2-1} + ... \\ & + \left[2u(0.5 - b_{k_{n_e}}(n)) - 1\right] 2^{k_{n_e}-1} \end{aligned} \tag{34}$$

where k_i, $i = 1, ..., n_e$ denotes the location of ith 1 in k, $b_j(n)$, $j = 1, ..., d_i$ denotes the value of jth bit in n and $u(\cdot)$ is the step function, which equals to 1 when its argument is positive and 0 otherwise.

Proof: The proof is given in Appendix III. ■

Now we can eliminate the XOR operator and Hamming weights in (11) and rewrite it as (35) (see the top of the next page), where n_e acts as the number of errors introduced by the BSC in the reported bit sequences and g represents the index of bit sequences with n_e-bit distance from the nth sequence.

In addition, *Lemma 2* enables us to approximate our desired statistics by limiting the maximum number of errors considered. Specifically, by adopting N_e, $1 \leq N_e \leq d_i$ as the upper limit of the first summation in (35), we derive an approximation whose accuracy can be controlled by N_e. For instance, if we have a reliable reporting channel, we can neglect the cases with more than one-bit error by using the following approximation, which is derived by restricting the first summation in (35) to $n_e = 1$

$$\begin{aligned} P_{r_i|\mathcal{H}_j}(n) \approx\ & (1-P_{b,i})^{d_i}\, P_{s_i|\mathcal{H}_j}(n) \\ & + P_{b,i}\,(1-P_{b,i})^{d_i-1} \sum_{k_1=1}^{d_i} P_{s_i|\mathcal{H}_j}(g(n,k_1)). \end{aligned} \tag{36}$$

By relaxing the linear mapping and uniform quantization assumptions, the developed analysis structure remains the same. However, there will no longer be linear relationships (29), and (30) between the statistics of the received bit sequences r_i, and their corresponding quantization levels y_i. Nevertheless, the moments of y_i are obtained in terms of the received bit-sequence pmfs as (for $j = 0, 1$)

$$E[y_i^\lambda|\mathcal{H}_j] = \sum_{n=1}^{2^{d_i}} q^\lambda_{n,i} P_{r_i|\mathcal{H}_j}\left(\Gamma(n)\right), \quad \lambda = 1, 2, ... \tag{37}$$

and the desired statistics are obtained accordingly. Hence, through analyzing the received bit sequences, we can derive the desired test statistics a_i and $\sigma^2_{y_i|\mathcal{H}_j}$ in terms of the number of quantization bits d_i in general case.

IV. Joint Reporting-Fusion Optimization

So far, we have thoroughly investigated the statistics governing our joint reporting-fusion optimization and derived analytical formulations describing the effect of local test summary quantization on the overall cooperative sensing performance. Now, we are ready to consider the joint optimization problem. Recall that we are dealing with (P2), which aims at joint optimization of $\mathbf{w}$ and $\mathbf{d}$.

For a given $\mathbf{d}$, (P2) can be solved for optimal weighting vector $\tilde{\mathbf{w}}$ through considering the Lagrange dual problem and Karush-Kuhn-Tucker (KKT) conditions [13] which yield

$$\tilde{\mathbf{w}} = \mathbf{\Sigma}_{\mathcal{H}_0}^{-1/2} \left[Q^{-1}(\alpha)\mathbf{I}_K + \zeta\mathbf{A}\right]^{-1} \mathbf{c} \tag{38}$$

where $\mathbf{A} \triangleq \mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{\Sigma}_{\mathcal{H}_0}^{-1}$ and $\mathbf{c} \triangleq \mathbf{\Sigma}_{\mathcal{H}_0}^{-1/2}\mathbf{a}$. ζ is the single root of the polynomial equation

$$\left\| \left[Q^{-1}(\alpha)\mathbf{I}_K + \zeta\mathbf{A}\right]^{-1} \mathbf{c} \right\| = 1 \tag{39}$$

and satisfies

$$Q^{-1}(\alpha)\mathbf{I}_K + \zeta\mathbf{A} \succ \mathbf{0} \tag{40}$$

where $\mathbf{0}$ stands for the null matrix and $\succ$ represents the element-wise inequality. Note that (39) and (40) specify a unique ζ as a function of $\mathbf{d}$.

Now, using (38), (39), and (40), we remove $\mathbf{w}$ from (P2) and convert it to an optimization in $\mathbf{d}$ and ζ. In addition, since the number of quantization levels cannot be infinite in practice, we limit $\mathbf{d}$ to lie between a minimum $\mathbf{d}_{\text{min}}$ and a maximum value $\mathbf{d}_{\text{max}}$. Moreover, as the q-function is strictly decreasing with respect to its argument, we remove it from (P2) and turn the problem into a minimization, i.e.,

$$\min_{\mathbf{d},\zeta} \varphi(\mathbf{d}, \zeta) \qquad \text{(P3)}$$

$$\text{s.t.} \begin{cases} \left\| \left[Q^{-1}(\alpha)\mathbf{I}_K + \zeta\mathbf{A}\right]^{-1} \mathbf{c} \right\| = 1 \\ Q^{-1}(\alpha)\mathbf{I}_K + \zeta\mathbf{A} \succ \mathbf{0} \\ \mathbf{d}_{\text{min}} \preceq \mathbf{d} \preceq \mathbf{d}_{\text{max}} \end{cases}$$

where

$$\varphi(\mathbf{d}, \zeta) \triangleq \frac{Q^{-1}(\alpha)\sqrt{\tilde{\mathbf{w}}^T\mathbf{\Sigma}_{\mathcal{H}_0}\tilde{\mathbf{w}}} - \mathbf{a}^T\tilde{\mathbf{w}}}{\sqrt{\tilde{\mathbf{w}}^T\mathbf{\Sigma}_{\mathcal{H}_1}\tilde{\mathbf{w}}}}. \tag{41}$$

The cost function in (P3) is a nonlinear function of a_i, $\sigma^2_{y_i|\mathcal{H}_0}$, $\sigma^2_{y_i|\mathcal{H}_1}$ (for $i = 1, ..., K$). As we have already studied,

$$P_{r_i|\mathcal{H}_j}(n) = (1-P_{b,i})^{d_i} P_{s_i|\mathcal{H}_j}(n) + \sum_{n_e=1}^{d_i} P_{b,i}^{n_e} (1-P_{b,i})^{d_i-n_e} \sum_{k_1=1}^{d_i} \sum_{\substack{k_2=1 \\ k_2\neq k_1}}^{d_i} \cdots \sum_{\substack{k_{n_e}=1 \\ k_{n_e}\neq k_1,\ldots,k_{n_e-1}}}^{d_i} P_{s_i|\mathcal{H}_j}(g(n,k_1,k_2,\ldots,k_{n_e})) \quad (35)$$

TABLE I
JOINT REPORTING-FUSION OPTIMIZATION AS A BRANCH-AND-BOUND PROCEDURE

1. Define set L of subproblems;
2. $L \leftarrow \{(P3)\}$; $B_U \leftarrow +\infty$; $\varphi_\varepsilon \leftarrow \varnothing$;
3. Solve (LP) for $\tilde{\mathbf{d}}_1$ and denote its minimum cost function by $B_L^{(1)}$;
4. **while** $L \neq \varnothing$
5. Choose $\mathcal{P}_K \in L$ with the minimum $B_L^{(K)}$;
6. $L \leftarrow L \backslash \{\mathcal{P}_K\}$;
7. $B_L \leftarrow B_L^{(K)}$;
8. Find a feasible solution $\tilde{\varphi}$ for (P3) via local search around $\tilde{\mathbf{d}}_K$;
9. $B_U^{(K)} \leftarrow \tilde{\varphi}$;
10. **if** $B_U^{(K)} < B_U$
11. $\varphi_\varepsilon \leftarrow B_U^{(K)}$;
12. $B_U \leftarrow B_U^{(K)}$;
13. **if** $B_L \geq (1-\varepsilon)B_U$
14. output φ_ε;
15. **else**
16. Remove from L all $\mathcal{P}_{K'}$ with $B_L^{(K')} \geq (1-\varepsilon)B_U$;
17. **end if**
18. **end if**
19. Choose a branching variable d_i and a branching point d_{branch};
20. Create subproblems $\mathcal{P}_{K+}$ and $\mathcal{P}_{K-}$;
21. Solve linear relaxations of $\mathcal{P}_{K+}$ and $\mathcal{P}_{K-}$ for $\tilde{\mathbf{d}}_{K+}$ and $\tilde{\mathbf{d}}_{K-}$ and denote their optimal cost functions by $B_L^{(K+)}$ and $B_L^{(K-)}$;
22. **if** $B_L^{(K+)} \leq (1-\varepsilon)B_U$
23. $L \leftarrow L \cup \{\mathcal{P}_{K+}\}$;
24. **end if**
25. **if** $B_L^{(K-)} \leq (1-\varepsilon)B_U$
26. $L \leftarrow L \cup \{\mathcal{P}_{K-}\}$;
27. **end if**
28. **end while**
29. output φ_ε;

and derived these parameters in terms of the number of quantization bits (i.e., d_i), we can clearly see that the cost function in (P3) is highly sophisticated and nonlinear. Since $\mathbf{d}$ represents the number of quantization levels used in the CR nodes, (P3) is a MINLP problem, which is NP hard in general [28]. We develop a Branch-and-Bound (BnB) algorithm to solve this MINLP.

A. The Branch-and-Bound Algorithm

Our optimization problem is a nonlinear program further constrained by integrality restrictions. Clearly, the optimal value of cost function in a continuous linear relaxation of (P3) will always be a lower bound on the optimal value of our cost function. Moreover, in any minimization, any feasible point always specifies an upper bound on the optimal cost function value. The idea of the BnB is to utilize these observations to subdivide MINLP's feasible region into more-manageable subdivisions and then, if required, to further partition the subdivisions. These subdivisions make a so-called *enumeration tree* whose branches can be pruned in a systematic search for the global optimum.

Table I shows the pseudocode of our BnB algorithm. Assuming φ^* as the global minimum of the cost function in (P3), this algorithm provides a $(1-\varepsilon)$ optimal solution φ_ε, which means φ_ε is close enough to φ^* such that $\varphi^* \geq (1-\varepsilon)\varphi_\varepsilon$.

In this algorithm, a lower bound for the cost function is first derived through solving a linear relaxation of (P3) denoted by (LP) (see line 3 in Table I). Construction of the linear relaxation is described later. Then, a local search is performed around the solution of the linear program to obtain a feasible point for (P3) and an upper bound for the global minimum (line 8). Note that any feasible solution of (P3) gives an upper bound on the minimum of φ. This process, i.e., finding the lower and upper bounds for the cost function, is called *bounding*. The algorithm terminates if the derived upper and lower bounds are within the ε-vicinity of each other (lines 13, 14). Otherwise, it continues with the so-called *branching* step, which refers to dividing the feasible region of the problem into two narrower subsets (lines 19, 20).

In this algorithm, maximum relaxation error is considered as metric for choosing the branching variable. That is, the variable d_i with maximum relaxation error is selected for the branching process. The relaxation error for the variable d_i is defined as $\left|d_i^{\text{ls}} - d_i^{\text{lp}}\right|$, where d_i^{ls} and d_i^{lp} denote the value of d_i obtained by the local search and by solving the linear program, respectively. The branching point is $d_{\text{branch}} = \left\lfloor d_i^{\text{lp}} \right\rfloor$, i.e., the problems $\mathcal{P}_{K-}$ and $\mathcal{P}_{K+}$ are constructed through imposing the constraints $d_i \leq \left\lfloor d_i^{\text{lp}} \right\rfloor$ and $d_i \geq \left\lceil d_i^{\text{lp}} \right\rceil$ on $\mathcal{P}_K$, respectively.

Through an iterative branching procedure, subsets are further divided into smaller ones, and the enumeration tree is built. This tree structure allows the algorithm to remove some branches and search for the solution in a very effective way. Moreover, narrowing down the subsets of the optimization variables leads to tighter linear relaxations (i.e., increases B_L) and provides the next local search processes with a closer starting point to the optimal solution (i.e., reduces B_U). Hence, the gap between B_L and B_U is reduced as the process continues. More specifically, at each iteration, the global lower bound B_L is updated to contain the minimum of the lower bounds of all subsets (lines 5, 7). The global upper bound B_U is also updated at each iteration (lines 10, 12), and the branches with a lower bound greater than $(1-\varepsilon)B_U$ are pruned (line 16). This procedure is continued until the difference between the global lower and upper bounds satisfy the accuracy ε (lines 13, 14). Clearly, we may lose the global optimum by pruning the branches. However, if the global optimum is in a pruned branch with the lower bound $B_L^{(K)}$, then $\varphi^* \geq B_L^{(K)}$, and consequently, $\varphi^* \geq (1-\varepsilon)B_U$. Therefore, the current best feasible solution with objective value B_U is already an $(1-\varepsilon)$ optimal solution, and we can still guarantee $(1-\varepsilon)$ optimality. Indeed, this guarantee is the key feature of the BnB algorithm which makes it very effective in solving the MINLPs.

TABLE II
DECOMPOSITION OF (P3) INTO SMALL BLOCKS AND SIMPLE CONSTRAINTS USED IN CONVEX HULL RELAXATION

Elements of $\mathbf{X}_{\mathcal{H}_k}^{(j)}$, $\mathbf{Z}_{\mathcal{H}_k}$, and $\mathbf{z}$ for $j=1,...,7$, and $k=0,1$.
$X_{i,n}^{(1)} = m\sigma_i\left(\frac{2(n-1)}{Z_{i,1}}-1\right)$, $X_{i,n\vert\mathcal{H}_k}^{(2)} = Q\left(X_{i,n}^{(1)}/\sigma_{i\vert\mathcal{H}_k}\right)$, $X_{i,n\vert\mathcal{H}_k}^{(3)} = X_{i,n\vert\mathcal{H}_k}^{(2)} - X_{i,n+1\vert\mathcal{H}_k}^{(2)}$, $X_{i,n\vert\mathcal{H}_k}^{(4)} = \sum_{k_1=1}^{d_i} X_{i,g(n,k_1)\vert\mathcal{H}_k}^{(3)}$, $X_{i,n\vert\mathcal{H}_k}^{(5)} = Z_{i,5}X_{i,n\vert\mathcal{H}_k}^{(4)}$, $X_{i,n\vert\mathcal{H}_k}^{(6)} = Z_{i,2\vert\mathcal{H}_k}X_{i,n\vert\mathcal{H}_k}^{(3)}$, $X_{i,n\vert\mathcal{H}_k}^{(7)} = X_{i,n\vert\mathcal{H}_k}^{(5)} + X_{i,n\vert\mathcal{H}_k}^{(6)}$.
$Z_{i,1} = 2^{d_i}-1$, $Z_{i,2} = c_M Q\left(\sqrt{c_M''/d_i}\right)$, $Z_{i,3} = 1-d_iZ_{i,2}$, $Z_{i,4} = Z_{i,2}+Z_{i,3}$, $Z_{i,5} = Z_{i,2}Z_{i,4}$, $Z_{i,6\vert\mathcal{H}_k} = \sum_{n=0}^{2^{d_i}-1} nX_{i,n\vert\mathcal{H}_k}^{(7)}$, $Z_{i,7} = Z_{i,6\vert\mathcal{H}_1} - Z_{i,6\vert\mathcal{H}_0}$, $Z_{i,8} = \frac{2m\sigma_i}{Z_{i,1}}Z_{i,7}$, $Z_{i,9\vert\mathcal{H}_k} = \left(Z_{i,6\vert\mathcal{H}_k}\right)^2$, $Z_{i,10\vert\mathcal{H}_k} = \sum_{n=0}^{2^{d_i}-1} n^2X_{i,n\vert\mathcal{H}_k}^{(7)}$, $Z_{i,11\vert\mathcal{H}_k} = Z_{i,10\vert\mathcal{H}_k} - Z_{i,9\vert\mathcal{H}_k}$, $Z_{i,12\vert\mathcal{H}_k} = \zeta Z_{i,11\vert\mathcal{H}_k}$, $Z_{i,13} = Q^{-1}(\alpha)Z_{i,11\vert\mathcal{H}_0} + Z_{i,12\vert\mathcal{H}_1}$, $Z_{i,14} = Z_{i,8}^2$, $Z_{i,15} = \frac{Z_{i,14}}{Z_{i,13}}$, $Z_{i,16\vert\mathcal{H}_k} = Z_{i,14}Z_{i,11\vert\mathcal{H}_k}$, $Z_{i,17} = Z_{i,13}^2$, $Z_{i,18\vert\mathcal{H}_k} = \frac{Z_{i,16\vert\mathcal{H}_k}}{Z_{i,17}}$, $Z_{i,19\vert\mathcal{H}_k} = \frac{Z_{i,16\vert\mathcal{H}_k}}{Z_{i,17}}$.
$z_1 = Q^{-1}(\alpha) - \sum_{i=1}^K Z_{i,15}$, $z_2 = \sum_{i=1}^K Z_{i,18}$, $z_3 = \sqrt{z_2}$, $z_4 = \sum_{i=1}^K Z_{i,19\vert\mathcal{H}_0} - 1$, $z_5 = z_1/z_3$.

It has been shown that under very general conditions, a BnB solution procedure always converges [29], [30]. Moreover, although the worst-case complexity of such a procedure is exponential, the actual running time could be fast when all partition variables are integers (e.g., the problem considered in this paper).

B. Convex Hull Relaxation

To derive a linear relaxation of our joint optimization problem, we reconfigure (P3) by introducing a number of auxiliary variables along with some additional constraints. In this process, the cost and constraints in (P3) are decomposed into a set of small easy-to-handle functions. We refer to these small functions as *blocks*. These blocks build a set of simple constraints which, as a whole, represent (P3). Finally, the derived constraint functions are replaced by appropriate linear inequalities.

We have decomposed our cost and constraint functions into a set of blocks represented in Table II. Through combination of these blocks, it can be easily verified that the following optimization problem is equivalent to (P3)

$$\min_{\mathbf{v}} z_5 \qquad \text{(P4)}$$
$$\text{s.t.} \begin{cases} \text{Constraints in Table II} \\ z_4 = 0 \\ Z_{i,13} > 0, \quad \text{for} \quad i = 1,...,K \\ \mathbf{v}_{\text{min}} \preceq \mathbf{v} \preceq \mathbf{v}_{\text{max}} \end{cases}$$

where $\mathbf{v}$ contains the old optimization variables, namely, $\mathbf{d}$ and ζ, as well as the new ones defined in Table II, i.e.,

$$\mathbf{v} \triangleq [\mathfrak{X}_{\mathcal{H}_0}^{(1)},...,\mathfrak{X}_{\mathcal{H}_0}^{(7)},\mathfrak{X}_{\mathcal{H}_1}^{(1)},...,\mathfrak{X}_{\mathcal{H}_1}^{(7)},\mathfrak{Z}_{\mathcal{H}_0},\mathfrak{Z}_{\mathcal{H}_1},\mathbf{z},\mathbf{d}^T,\zeta]^T \quad (42)$$

in which row vectors $\mathfrak{X}_{\mathcal{H}_k}^{(j)}$ and $\mathfrak{Z}_{\mathcal{H}_k}$, $j=1,...,7$, $k=0,1$, denote the vectorized forms of $\mathbf{X}_{\mathcal{H}_k}^{(j)}$ and $\mathbf{Z}_{\mathcal{H}_k}$ respectively, i.e., $\mathfrak{X}_{\mathcal{H}_k}^{(j)} \triangleq \text{vec}\left(\mathbf{X}_{\mathcal{H}_k}^{(j)}\right)$ and $\mathfrak{Z}_{\mathcal{H}_k} \triangleq \text{vec}\left(\mathbf{Z}_{\mathcal{H}_k}\right)$. The elements of $\mathbf{X}_{\mathcal{H}_k}^{(j)}$, $\mathbf{Z}_{\mathcal{H}_k}$, and $\mathbf{z}$ represent the blocks defined in Table II. These elements are denoted as[1]

$$\left[\mathbf{X}_{\mathcal{H}_k}^{(j)}\right]_{i,n} \triangleq X_{i,n\vert\mathcal{H}_k}^{(j)} \quad (43)$$
$$\left[\mathbf{Z}_{\mathcal{H}_k}\right]_{i,j} \triangleq Z_{i,j\vert\mathcal{H}_k} \quad (44)$$
$$\mathbf{z} \triangleq [z_1,...,z_5]. \quad (45)$$

The vectors $\mathbf{v}_{\text{min}}$ and $\mathbf{v}_{\text{max}}$ denote the lower and upper bounds on the elements of $\mathbf{v}$, respectively. These bounds are directly obtained by applying $\mathbf{d}_{\text{min}} \preceq \mathbf{d} \preceq \mathbf{d}_{\text{max}}$ on the blocks.

It is worth noting that any element v_i in $\mathbf{v}$ can be represented as a function of other elements, i.e., for $v_i \in [v_{i,\text{min}}, v_{i,\text{max}}]$, we have $v_i = \vartheta_i(\mathbf{v})$. ϑ_i is either a convex (or concave) function or it represents a product (or ratio) of two elements in $\mathbf{v}$, i.e., $\vartheta_i(\mathbf{v}) = v_k v_j$.

If ϑ_i is convex, we linearize it by partitioning each interval $[v_{l,\text{min}}, v_{l,\text{max}}]$ into $N_{\text{lin}}-1$ subintevals, $1 \le l \le |\mathbf{v}|$. This partitioning is realized by considering N_{lin} points as $v_{l,\text{min}} = \hat{v}_{l,1} \le \hat{v}_{l,2} \le ... \le \hat{v}_{l,N_{\text{lin}}} = v_{l,\text{max}}$. In this way, we make a grid over the space $\mathbf{v}_{\text{min}} \preceq \mathbf{v} \preceq \mathbf{v}_{\text{max}}$. Now for any subspace within this grid $\hat{\mathbf{v}}_k \preceq \mathbf{v} \preceq \hat{\mathbf{v}}_{k+1}$, $k \in \{1,...,N_{\text{lin}}\}$, we have the following linear lower bound for v_i

$$v_i \ge \vartheta_i(\tilde{\mathbf{v}}) + \nabla\vartheta_i(\tilde{\mathbf{v}})^T(\mathbf{v}-\tilde{\mathbf{v}}) \quad (46)$$

where $\tilde{\mathbf{v}}$ can be any point such that $\hat{\mathbf{v}}_k \preceq \tilde{\mathbf{v}} \preceq \hat{\mathbf{v}}_{k+1}$, and $\hat{\mathbf{v}}_k \triangleq \left[\hat{v}_{1,k},...,\hat{v}_{|\mathbf{v}|,k}\right]$. The number of points N_{lin} can be used to control the precision of the partitioning process.

For $v_i \in [v_{i,\text{min}}, v_{i,\text{max}}]$, the upper bound in our linearization is denoted by the following relation,

$$v_i \le \vartheta_i(\mathbf{v}_{\text{min}}) + \frac{\vartheta_i(\mathbf{v}_{\text{max}}) - \vartheta_i(\mathbf{v}_{\text{min}})}{v_{k,\text{max}} - v_{k,\text{min}}}(v_k - v_{k,\text{min}}) \quad (47)$$

where v_k can be any element in $\mathbf{v}$ such that $\partial\vartheta_i/\partial v_k \neq 0$, and $v_{k,\text{min}}$, and $v_{k,\text{max}}$ denote the minimum and maximum values of v_k, respectively. Similar bounds can be obtained for the concave functions.

For the blocks which represent the product of two elements, i.e., $\vartheta_i(\mathbf{v}) = v_k v_j$, the tightest linear constraints are denoted

[1] We have dropped $\mathcal{H}_k$ from the element representations whenever there is no difference between the element values for $\mathcal{H}_0$ and $\mathcal{H}_1$.

by [31], [32]

$$v_i \geq v_{j,\min} v_k + v_{k,\min} v_j - v_{j,\min} v_{k,\min} \quad (48)$$
$$v_i \geq v_{j,\max} v_k + v_{k,\max} v_j - v_{j,\max} v_{k,\max} \quad (49)$$
$$v_i \leq v_{j,\min} v_k + v_{k,\max} v_j - v_{j,\min} v_{k,\max} \quad (50)$$
$$v_i \leq v_{j,\max} v_k + v_{k,\min} v_j - v_{j,\max} v_{k,\min}. \quad (51)$$

Applying (46)-(51) on the functions defined in Table II, we derive a linear relaxation of (P4) whose solution provides a tight lower bound on the minimum value of our cost function in (P3). This linear program can be expressed as

$$\min_{\mathbf{v}} z_5 \quad \text{(LP)}$$
$$\text{s.t.} \begin{cases} \text{Linear constraints (46)-(51)} \\ z_4 = 0 \\ Z_{i,13} > 0, \quad \text{for} \quad i = 1, ..., K \\ \mathbf{v}_{\min} \preceq \mathbf{v} \preceq \mathbf{v}_{\max} \end{cases}$$

which is solved in polynomial time.

C. Low-complexity Suboptimal Solution

As mentioned earlier, a commonly used suboptimal approach for the design of linear cooperation is based on maximizing the MDC. The MDC provides a good measure of the detection capability, as it characterizes the system performance as the variance-normalized distance between the centers of two conditional pdfs of the global test summary y_c. Therefore, we propose an alternative approach for the discussed joint reporting-fusion optimization based on the MDC, which leads to nearly optimal performance with much less effort. The MDC is defined as

$$\Delta_m^2 \triangleq \frac{\left(\mathrm{E}\left[y_c|\mathcal{H}_1\right] - \mathrm{E}\left[y_c|\mathcal{H}_0\right]\right)^2}{\mathrm{Var}\left\{y_c|\mathcal{H}_1\right\}}. \quad (52)$$

Replacing y_c with its weighted sum definition, we have

$$\Delta_m^2 = \frac{\left(\mathbf{a}^T\mathbf{w}\right)^2}{\mathbf{w}^T\mathbf{\Sigma}_{\mathcal{H}_1}\mathbf{w}}. \quad (53)$$

Using the MDC approach, we aim at finding $\mathbf{w}$ and $\mathbf{d}$ such that

$$\max_{\mathbf{w},\mathbf{d}} \Delta_m^2 \quad \text{(P5)}$$
$$\text{s.t.} \begin{cases} \|\mathbf{w}\| = 1 \\ \mathbf{d}_{\min} \preceq \mathbf{d} \preceq \mathbf{d}_{\max} \end{cases}$$

The constraint on the weight vector norm is necessary here to derive a unique solution since the MDC does not depend on $\|\mathbf{w}\|$.

In order to derive an analytical solution for (P5), we first eliminate $\mathbf{w}$ as follows. Through the linear transformation [8]

$$\mathbf{w}' = \mathbf{\Sigma}_{\mathcal{H}_1}^{1/2}\mathbf{w} \quad (54)$$

the MDC is converted to

$$\Delta_m^2 = \frac{\mathbf{w}'^T\mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\mathbf{a}\mathbf{a}^T\mathbf{\Sigma}_{\mathcal{H}_1}^{-1/2}\mathbf{w}'}{\mathbf{w}'^T\mathbf{w}'} \leq \left\|\mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\mathbf{a}\right\|^2 \quad (55)$$

and the equality is achieved when $\mathbf{w}' = \mathbf{\Sigma}_{\mathcal{H}_1}^{-T/2}\mathbf{a}$. Therefore, the optimal $\mathbf{w}$ which maximizes the MDC is derived as function of $\mathbf{d}$ as

$$\mathbf{w}_{\text{mdc}} = \frac{\mathbf{\Sigma}_{\mathcal{H}_1}^{-1/2}\mathbf{w}'}{\left\|\mathbf{\Sigma}_{\mathcal{H}_1}^{-1/2}\mathbf{w}'\right\|}. \quad (56)$$

Replacing $\mathbf{w}$ with its MDC-optimal value $\mathbf{w}_{\text{mdc}}$, the MDC can be rewritten as a function of $\mathbf{d}$

$$\Delta_m^2 = \mathbf{a}^T\mathbf{\Sigma}_{\mathcal{H}_1}^{-1}\mathbf{a} = \sum_{i=1}^{K} \frac{a_i^2}{\sigma_{y_i|\mathcal{H}_1}^2}. \quad (57)$$

Therefore, (P5) is converted to the following system of optimization problems

$$\text{for } i = 1, ..., K, \quad \max_{d_i} \frac{a_i^2}{\sigma_{y_i|\mathcal{H}_1}^2} \quad \text{(P6)}$$
$$\text{s.t.} \qquad d_{i,min} \leq d_i \leq d_{i,max}$$

We are now dealing with K one-dimensional problems and it is clear that the computational complexity of solving this set of optimizations linearly increases with the number of sensing nodes K.

V. Numerical Results

Two typical distributed detection scenarios have been considered to illustrate the effectiveness of the proposed optimization scheme. In the first scenario, the ED and uniform quantization have been adopted as local sensing and test summary quantization methods at the CR nodes, respectively. As the second scenario, CSD and MOE quantization have been considered. Although not presented here, we also tested other combinations such as CSD with uniform quantization and ED with MOE quantization and obtained similar results, which are expected since our proposed optimization does not depend on any specific local sensing or quantization scheme. In all simulations, there are $K = 5$ cooperating nodes which transmit their sensing outcomes over the reporting channels using the Binary Phase Shift Keying (BPSK) modulation. The PU signal is modeled as a Direct-Sequence (DS) spread-spectrum BPSK signal by using Walsh-Hadamard code with the length of 16, i.e., the processing gain of 16 is considered in all simulation results. The maximum number of quantization bits in each node is 7.

In ED, the energy of the received PU signal is measured using N signal samples, i.e.,

$$u_i = \sum_{k=1}^{N} |x_i(k)|^2 \quad (58)$$

and in CSD the test summary is formed as an estimation of the PU signal autocorrelation function as

$$u_i = \frac{1}{N-\tau} \sum_{k=1}^{N-\tau} x(k+\tau)x^*(k)e^{-j2\pi fk} \quad (59)$$

where f is the cycle frequency, and τ is the lag used in calculating the autocorrelation.

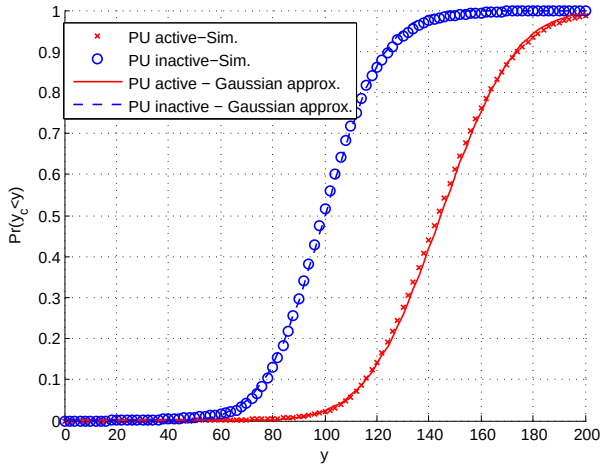

Fig. 2. Cumulative probability distribution of y_c under hypotheses $\mathcal{H}_0$ (the two leftmost curves) and $\mathcal{H}_1$ (the two rightmost curves). Solid curves correspond to the analytic Gaussian approximations. Marked curves are obtained by Monte-Carlo simulation.

The number of PU signal samples N used in ED is 20 and the Chebyshev probability for the uniform quantization coverage interval has been set to 95%. When CSD is considered, the cyclic autocorrelation function is estimated in each node using $N = 100$ samples of the PU signal and the estimated autocorrelation corresponds to $f = 1/T_c$ and $\tau = 0$. T_c is the chip period of the PU signal. For the ED, the local listening channel SNR levels at sensor inputs are {0, -2.7, -3.1, -1.4, -6.9} in dB, and the reporting channel SNRs are {10, 13, 12, 14, 11} in dB. For the CSD, {5, 2.3, 1.9, 3.6, -1.9} in dB are the listening channel SNR levels and the reporting channel SNRs are {11, 14, 13, 12, 11} in dB. Each point on the CROC curves has been derived by averaging over 10,000 realizations. The averaging has been performed on the noise for a fixed set of channel gains and noise variances as in [8], [13].

Fig. 2 shows why Gaussian approximation works well for the global test statistic in (17). It depicts the cumulative distribution function (CDF) of the global test summary y_c when ED is used as the local sensing method at $K = 5$ sensing nodes, assuming both hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$. It can be seen by comparing the simulation and asymptotic results that y_c behaves very similarly to a normal random variable while the PU is either present or absent.

Figs. 3, and 4 depict the results derived as Complementary Receiver Operational Characteristics (CROC) curves for both energy and cyclostationary detectors. Specifically, for each detector three cases have been considered as

Case#1: Depicts the performance of uniform linear combining at the fusion center and maximum number of quantization bits at the sensing nodes,

Case#2: Depicts the performance of optimal linear combining at the fusion center and maximum number of quantization bits at the sensing nodes,

Case#3: Depicts the performance of the proposed joint optimization, i.e., optimal linear combining at the fusion center and optimal number of quantization bits at the sensing nodes.

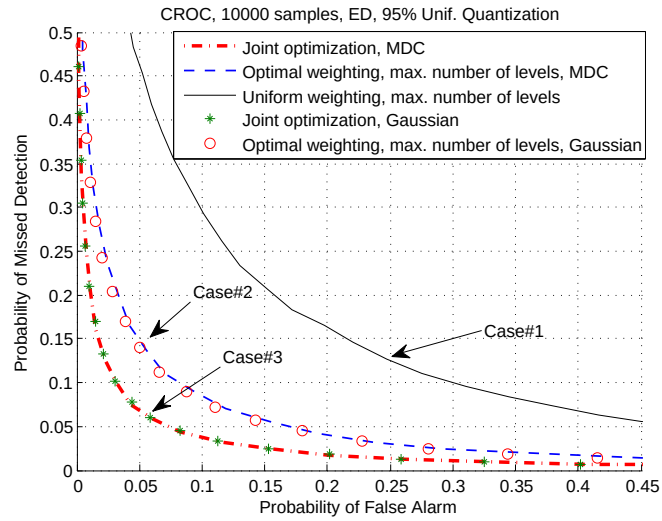

Fig. 3. CROC curves for the energy detection using 20 samples of the PU signal and uniform quantization with Chebyshev probability of 95%. The listening channel SNR levels at sensor inputs are {0, -2.7, -3.1, -1.4, -6.9} in dB. The reporting channel SNR levels are {10, 13, 12, 14, 11} in dB. The results are obtained using 10,000 realizations.

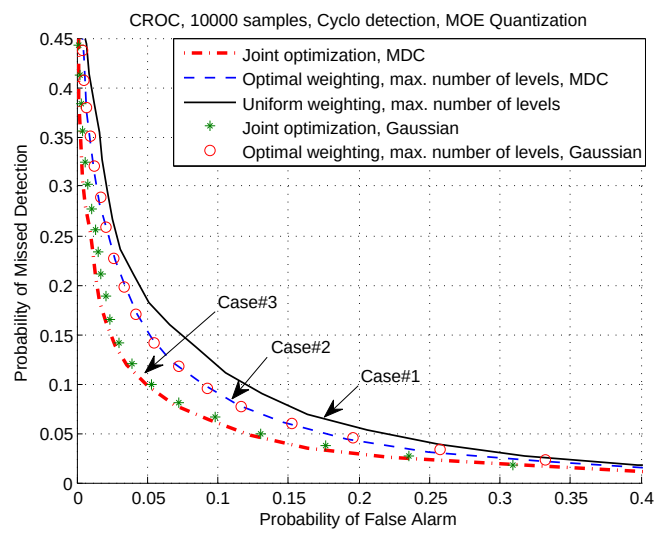

Fig. 4. CROC curves for the cyclostationary detection using 100 samples of the PU signal and MOE quantization. The local SNR levels at sensor inputs are {5, 2.3, 1.9, 3.6, -1.9} in dB. The reporting channel SNR levels are {11, 14, 13, 12, 11} in dB. The results are obtained using 10,000 realizations.

It is worth noting that in both figures Case#1, Case#2, and Case#3 represent the detector design without any optimization, only with optimal weighting, and with joint reporting-fusion optimization, respectively. The plots clearly illustrate the effectiveness of our proposed detector in terms of lower false alarm and missed detection probabilities which are shown as CROC curves closer to the origin. Moreover, it can be observed that the achieved optimization results from both the Gaussian approximation and MDC are in close agreement with each other. As another point, we see that performance improvement due to the proposed optimization is higher when ED is used as the local sensing method. This observation is reasonable since ED is known to be more sensitive to the

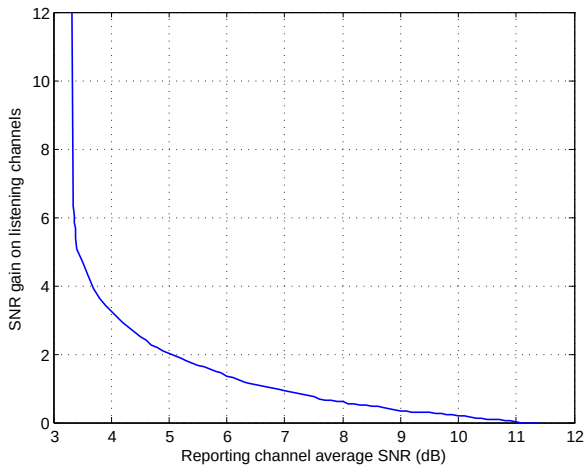

Fig. 5. SNR gains obtained by the proposed joint reporting-fusion optimization scheme vs. the average reporting SNR level experienced by $K = 5$ sensing nodes (SNR_0) and with reporting SNR deviation of $\Delta = 1$ (dB).

TABLE III
OPTIMAL NUMBER OF QUANTIZATION LEVELS FOR DIFFERENT LISTENING, AND REPORTING SNRS

$\text{SNR}_r \downarrow$, $\text{SNR}_l \rightarrow$	-15 dB	-10 dB	0 dB
0 dB	15	15	4
5 dB	30	15	4
15 dB	31	16	4

TABLE IV
OPTIMAL FUSION WEIGHTS, AND OPTIMAL NUMBER OF QUANTIZATION LEVELS FOR NODES OPERATING IN DIFFERENT LISTENING-CHANNEL, AND REPORTING-CHANNEL CONDITIONS

SNR_l (dB)	-15	-10	0
SNR_r (dB)	15	5	0
Opt. Weights	0.0037	0.0171	0.9998
Opt. No. of Levels	31	16	4

PU signal SNR than the cyclostationary detection. In fact, the linear combining scheme at the FC is a technique to exploit spatial diversity among different sensing nodes to increase the effective SNR level experienced by the detector. For a comparison between ED and CSD performances, see [33]. Note also that we have adopted the parameter values to have almost equal performances in both detectors before applying the proposed optimization. More specifically, by comparing the curves labeled Case#2 in Figs. 2 and 3, we see almost-equal performances. In fact, we have set the parameters in a way that both detectors meet $P_{\text{fa}} = 0.1$, and $P_{\text{md}} = 0.1$ (both curves nearly pass the point $(0.1, 0.1)$). In other words, although the SNR levels and sensing times are different, we have evaluated the achieved performance gain by using nearly-equal detectors.

In order to evaluate the performance gain achieved by the proposed joint optimization method, we define the listening channel SNR gain as follows

$$\text{SNR gain (dB)} = \text{SNR}_{\text{maxq}} \text{ (dB)} - \text{SNR}_{\text{optq}} \text{ (dB)} \qquad (60)$$

where SNR_{maxq} is the minimum SNR required at the SUs to meet $P_{\text{fa}} = \alpha$ and $P_{\text{md}} = \beta$ when they use the maximum number of quantization levels with optimal weighting and SNR_{optq} is the minimum SNR required at the SUs when they use optimal number of quantization levels and optimal weighting vector derived through the proposed joint optimization scheme.

Fig. 5 depicts the SNR gain at the listening channels vs. the average SNR at the reporting channels. It represents the case in which $K = 5$ sensing nodes experience different SNRs on their listening and reporting channels. The corresponding SNRs for both the reporting and listening channels of $K = 5$ sensing nodes are $[\text{SNR}_0 + 2\Delta \ \text{SNR}_0 + \Delta \ \text{SNR}_0 \ \text{SNR}_0 - \Delta \ \text{SNR}_0 - 2\Delta]$, where SNR_0 is the average SNR over the 5 channels and $\Delta = 1$ (dB). The target false alarm and missed detection probabilities are both set to 5%. As shown in Fig. 5, the proposed joint optimization leads to a significant performance gain, especially at low SNR regimes experienced at the reporting channels. This superior performance of the proposed design at low SNRs stems from the fact that, the local sensing quantization is also considered when optimizing the linear combining at the FC. Hence, the effect of reporting channel impairments is reduced and, consequently, the proposed detector experiences the reporting channel BEP wall (see [19] and [22]) at lower SNR values, compared to the design which only optimizes the linear fusion. In other words, by decreasing the reporting channel SNR levels, at a certain point, the detector with the maximum number of quantization levels can not meet the target false alarm and missed detection probabilities, no matter how high the SNR levels at the listening channels are. But the proposed detector still reaches the desired performance for a moderate average SNR level at the listening channels. This observation demonstrates the importance of the proposed method, especially when the sensing nodes face stringent energy consumption constraints due to e.g., limited battery life time or green communication considerations, which force the designer to reduce the transmission power used for the reporting phase.

A set of significant observations is obtained by evaluating the derived quantization levels for different listening-channel and reporting-channel SNRs (Table III). In particular, we observe that, when the sensing nodes experience high SNR levels at the listening channel, the proposed optimization scheme reduces the number of quantization levels as much as possible. This clearly means that, when the PU signal is strong, only a small precision in quantizing the sensing outcomes is enough to have the desired detection performance. Therefore, the optimizer takes the advantage of this effect by reducing the number of quantization levels to decrease the reporting channel contaminations on the reported sensing outcomes. Moreover, as long as the listening channel SNR is high, the optimal number of quantization levels remains low for low-, medium-, and high-level reporting SNRs. This means, for the high listening-channel SNR, that the optimal number of levels is not sensitive to the reporting channel SNR levels. In other words, when the local detection processes are reliable, the overall system performance is not affected significantly by the reporting channel degradations. However, we observe a different behavior when the listening-channel SNR is low, i.e.,

when cooperation in sensing is highly necessary. Specifically, since the local sensing has to be performed on weak PU signals in this case, the uncertainty in determining whether the PU is active or not is quite high. The optimizer tries to compensate for this uncertainty (or not to contribute to this uncertainty) by increasing the precision of the local quantization process at the expense of suffering a higher degradation at the reporting phase. Recall that increasing the number of levels improves the quality of local quantization but increases the BEP of the reporting channel as well. As an important conclusion, the proposed joint optimization approach enables the system to trade the reporting channel quality for the local sensing quality in order to achieve the best overall detection performance. These observations are illustrated in Table III, which shows the optimal number of quantization levels for various listening and reporting SNRs.

Table IV shows typical values of the obtained optimal weighting and quantization vectors for sensing nodes operating in different listening- and reporting-channel conditions.

VI. Conclusion

In this paper, after a structured study of major phases in a centralized cooperative sensing scheme, the effect of the number of bits used in local sensing quantization on the overall sensing performance in a CRN with cooperative sensing has been introduced and a joint optimization approach has been proposed to optimize the linear soft-combining scheme at the fusion phase with the number of quantization bits used by each sensing node at the reporting phase. The presented analytical expressions followed by simulation results demonstrate that, through joint consideration of the reporting and fusion phases in a cooperative sensing scheme, considerable performance gains can be obtained. This better performance stems from better exploitation of spatial/user diversities in CRNs. The proposed joint optimization scheme leads to more powerful distributed detection performance, especially when the sensing nodes have to work at low SNR regimes.

Appendix I
Proof of Lyapunov's CLT condition for the received quantized test summaries

Proof: For simplicity, we consider the uniform quantization here. Other quantization schemes can be treated similarly. For any $\delta > 0$, we have (for either $\mathcal{H}_0$ or $\mathcal{H}_1$)

$$E\left[|y_i - \bar{y}_i|^{2+\delta}\right] \leq (2m\sigma_i)^{2+\delta} \leq (2m\sigma_{\max})^{2+\delta} \tag{61}$$

where $\bar{y}_i$ denotes the mean of y_i, and $\sigma_{\max}$ denotes the maximum standard deviation of the sensing outcomes. We also have

$$s_K^2 \triangleq \sum_{i=1}^{K} \sigma_{y_i}^2 \geq K\sigma_{y_i,\min}^2 \tag{62}$$

where $\sigma_{y_i,\min}^2$ denotes the minimum variance of the received test summaries. Using the above inequalities, we can set an upper bound on the ratio in Lyapunov's condition, i.e.,

$$\frac{1}{s_K^{2+\delta}} \sum_{i=1}^{K} E\left[|y_i - \bar{y}_i|^{2+\delta}\right] \leq \frac{K\left(2m\sigma_{i,\max}\right)}{K^{1+\delta/2}\sigma_{y_i,\min}^2}. \tag{63}$$

Finally, since $\sigma_i \neq 0$ requires that $\sigma_{y_i,\min}^2 \neq 0$, the above upper bound approaches zero as $K \to \infty$, hence,

$$\lim_{K\to\infty} \frac{1}{s_K^{2+\delta}} \sum_{i=1}^{K} E\left[|y_i - \bar{y}_i|^{2+\delta}\right] = 0. \tag{64}$$

■

Appendix II
Proof of Lemma 1

Proof: The Gray code structure is depicted in Fig. 6, where a four-bit Gray code is decomposed by four lines L_1, L_2, L_3, and L_4 into four blocks $\mathbf{g}_1$, $\mathbf{g}_2$, $\mathbf{g}_3$, and $\mathbf{g}_4$. The vector $\mathbf{g}_k$, $k = 1, ..., 4$, denotes the k-bit Gray code in this format. Although only four bits are considered here, by repeating this structure we can construct Gray codes with arbitrary number of bits. According to this structure, we can see that $\mathbf{g}_{k+1}$ is generated by vertically concatenating $\mathbf{g}_k$ with $\mathbf{m}_k$, i.e., for $k \geq 1$,

$$\mathbf{g}_{k+1} = \begin{bmatrix} \mathbf{g}_k \\ \mathbf{m}_k \end{bmatrix} \tag{65}$$

where $\mathbf{m}_k$ is generated by mirroring $\mathbf{g}_k$ with respect to L_k and then flipping its $(k+1)$th bit into 1. Therefore, considering only the first k bits, $\mathbf{g}_k$ and $\mathbf{m}_k$ are symmetric with respect to L_k. The leftmost column in Fig. 6 contains the index of codewords in this structure. For $n = 1, ..., 2^k$, we denote the nth elements of $\mathbf{g}_k$ and $\mathbf{m}_k$ by $g_k(n)$ and $m_k(n)$, respectively.

Because of the aforementioned symmetry, for $l_e \geq 1$ and $n = 1, ..., 2^{l_e}$, change in the code index due to flipping the l_eth bit in $g_{l_e}(n)$ and $m_{l_e}(n)$ are the same. In other words, if flipping the l_eth bit in $g_{l_e}(n)$ converts it to $g_{l_e}(n_1)$ and flipping the same bit in $m_{l_e}(n)$ turns it into $m_{l_e}(n_2)$, then $n_1 = n_2$. That is, the change in the element index of $\mathbf{g}_{l_e}$ (i.e., $n_1 - n$) and $\mathbf{m}_{l_e}$ are equal when the l_eth bit is flipped in both $\mathbf{g}_{l_e}$ and $\mathbf{m}_{l_e}$. Moreover, since $\mathbf{g}_{l_e+1}$ is composed of $\mathbf{g}_{l_e}$ and $\mathbf{m}_{l_e}$, the same rule applies for $\mathbf{g}_{l_e+1}$. Consequently, when flipping the l_eth bit, the same jumping rule holds for all $\mathbf{g}_n$, $n \geq l_e$. Thus, in order to find how the code index is changed when flipping the l_eth bit, we only need to consider $\mathbf{g}_{l_e}$.

To maintain simplicity, we only focus on $l_e = 3$, but the general case is proved by considering the the Gray code structure's symmetry which stems from the explained mirroring process. Recall that by considering $\mathbf{g}_3$ and $l_e = 3$, we find the jumping rule when the 3rd bit is flipped in a *complete Gray code with arbitrary number of bits*. Focusing on $\mathbf{g}_3$ in Fig. 6, we see that flipping the 3rd bit is equivalent to two successive processes: *i*) increasing the code index by 2^{l_e} (which is 8 in this case), *ii*) mirroring the new codeword with respect to L_3. These two processes are shown by arrows on $g_3(3)$ in Fig. 6 as an example. $g_3(3)$ is first converted to $m_3(3)$ and then mirrored against L_3 to give $g_3(6)$.

According to the presented structure, mirroring a codeword with index $n \geq 2^{l_e}$ with respect to L_{l_e} is equivalent to decrease n by $2\text{mod}(n-1, 2^{l_e}) + 1$. Therefore, the two mentioned processes on the code index can be indicated by the following relationship

$$k(n, l_e) = n + 2^{l_e} - 2\text{mod}(n-1, 2^{l_e}) - 1 \tag{66}$$

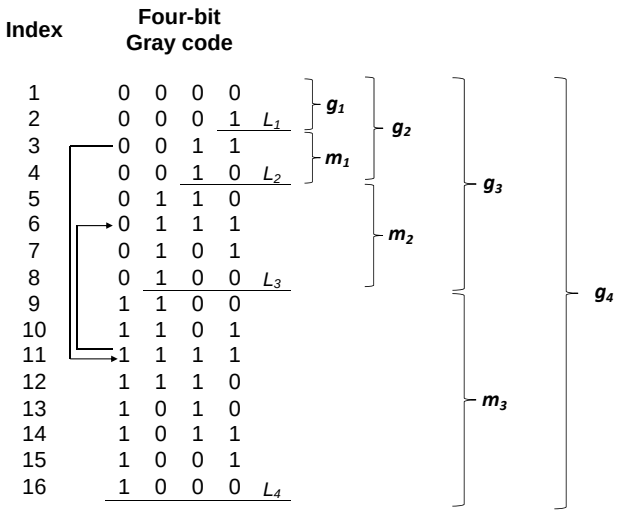

Fig. 6. Four-bit Gray code structure. The least significant bits are the rightmost ones.

where k denotes the new index and n denotes the old index. ■

Appendix III
Proof of Lemma 2

Proof: If k only contains a single 1 in its binary format (i.e., $n_e = 1$), then $k = 2^{k_1-1}$ and the XOR operation leads to

$$n \oplus k = \begin{cases} n + 2^{k_1-1}, & b_{k_1}(n) = 0 \\ n - 2^{k_1-1}, & b_{k_1}(n) = 1 \end{cases} \tag{67}$$

The right-hand side of this equation can be expressed in closed form by using the step function as

$$n \oplus k = n + \left[2u(0.5 - b_{k_1}(n)) - 1\right] 2^{k_1-1} \tag{68}$$

For $n_e \geq 1$, we have

$$k = 2^{k_1-1} \oplus 2^{k_2-1} \oplus ... \oplus 2^{k_{n_e}-1} \tag{69}$$

Hence, (34) is obtained by successively applying (68) on the first, second, ..., and n_eth bit in k. ■

References

[1] N. Jesuale, "Lights and sirens broadband," in *Proc. IEEE DySPAN 2011*, Aachen, Germany, May 3-6 2011, pp. 467–475.

[2] Y. Zeng, Y. Liang, S. W. O. Z. Lei, F. Chin, and S. Sun, "Worldwide regulatory and standardization activities on cognitive radio," in *Proc. IEEE DySPAN 2010*, Washington DC, April 6-9 2010, pp. 1–9.

[3] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Elsevier Physical Commun.*, vol. 4, no. 1, pp. 40–62, March 2011.

[4] FCC, "Unlicensed operation in the TV broadcast bands, additional spectrum for unlicensed devices below 900 MHz and in the 3 GHz band, order," *in ET Docket Nos. 04-186 and 02-380, DA 11-131*, Jan. 2011.

[5] C. Clancy, J. Hecker, E. Stuntebeck, and T. O'Shea, "Applications of machine learning to cognitive radio networks," *IEEE Wireless. Commun.*, vol. 4, no. 1, pp. 47–22, Aug. 2007.

[6] G. Ganesan and Y. G. Li, "Cooperative spectrum sensing in cognitive radio-part I: Two user networks," *IEEE Trans. on Wireless Commun.*, vol. 6, no. 6, pp. 2204–2213, 2007.

[7] J. Ma, G. Zhao, and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. on Wireless Commun.*, vol. 7, no. 11, pp. 28–40, Nov. 2008.

[8] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 28–40, Feb. 2008.

[9] Y. Zou, Y. Yao, and B. Zheng, "A selective-relay based cooperative spectrum sensing scheme without dedicated reporting channels in cognitive radio networks," *IEEE Trans. on Wireless Commun.*, vol. 10, no. 4, pp. 1188–1198, April 2011.

[10] Q. Chen, M. Motani, W. Wong, and A. Nallanathan, "Cooperative spectrum sensing strategies for cognitive radio mesh networks," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 56–67, Feb. 2011.

[11] K. Hossain, B. Champagne, and A. Assra, "Cooperative multiband joint detection with correlated spectral occupancy in cognitive radio networks," *IEEE Trans. Signal.Process.*, vol. 60, no. 5, pp. 2682–2687, May 2012.

[12] Z. Quan, W. Ma, S. Cui, and A. H. Sayed, "Optimal linear fusion for distributed detection via semidefinite programming," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2431–2436, 2010.

[13] G. Taricco, "Optimization of linear cooperative spectrum sensing for cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 77–86, Feb. 2011.

[14] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, "Optimal multiband joint detection for spectrum sensing in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1128–1140, March 2009.

[15] S. Chaudhari, , and V. Koivunen, "Effect of quantization and channel errors on collaborative spectrum sensing," in *Proc. The 43rd Asilomar Conference on Signals, Systems, and Computers*, Nov. 2009, pp. 528–533.

[16] J. Li, Z. Li, J. Si, and Y. Zhang, "Efficient soft decision fusion rule in cooperative spectrum sensing," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 1931–1943, April 2013.

[17] D. Oh, H. Lee, and H. Y. Lee, "Linear hard decision combining for cooperative spectrum sensing in cognitive radio systems," in *Proc. 72nd IEEE Vehic. Technol. Conf.*, Sep. 6-9 2010, pp. 1–5.

[18] T. C. Aysal, S. Kandeepan, and R. Piesiewicz, "Cooperative spectrum sensing with noisy hard decision transmissions," in *Proc. IEEE ICC'09*, Jun. 14-18 2009, pp. 1–5.

[19] S. Chaudhari and V. K. J. Lunden, "BEP walls for collaborative spectrum sensing," in *Proc. IEEE International Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, May 22-27 2011, pp. 2984–2987.

[20] B. Chen and P. K.Willett, "On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 693–699, Feb. 2005.

[21] J. Tsitsiklis, *Decentralized detection, Advances in Statistical Signal Processing*, H. V. Poor and J. B. Thomas, Eds. Greenwich, U.K.: CT: JAI Press, 1993, vol. 2.

[22] S. Chaudhari, J. Lunden, V. Koivunen, and H. V. Poor, "Cooperative sensing with imperfect reporting channels: Hard decisions or soft decisions?" *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 18–28, Jan. 2012.

[23] Z. Quan, S. Cui, and A. H. Sayed, "An optimal strategy for cooperative spectrum sensing in cognitive radio networks," in *Proc. IEEE GLOBECOM*, Washington DC., Nov. 2007, pp. 2947–2951.

[24] Y. C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Wireless. Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[25] S. Chaudhari, J. Lundén, V. Koivunen, and H. V. Poor, "BEP walls for cooperative sensing in cognitive radios using K-out-of-N fusion rules," *Elsevier Signal Process*, vol. 93, no. 7, pp. 1900–1908, July 2013.

[26] H. B. Yilmaz, T. Tugcu, and F. Alagoz, "Novel quantization-based spectrum sensing scheme under imperfect reporting channel and false reports," *Int. J. Commun. Syst.*, July 2012, early view.

[27] P. Billingsley, *Probability and measure*, 3rd ed. New York: John Wiley & Sons, 1995.

[28] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York: W.H. Freeman and Company, 1979.

[29] Y. Shi, Y. T. Hou, and H. Zhou, "Per-node based optimal power control for multi-hop cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5290–5299, Oct. 2009.

[30] H. D. Sherali and W. P. Adams, *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*. Kluwer Academic Publishers, 1999.

[31] G. P. McCormick, *Nonlinear programming: Theory, algorithms, and applications*. New York: John Wiley & Sons, 1983.

[32] P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Wachter, "Branching and bound tightening techniques for non-convex MINLP," *IBM Research Report*, Aug. 2008.

[33] J. Lundén, V. Koivunen, A. Huttunen, and H. V. Poor, "Collaborative cyclostationary spectrum sensing for cognitive radio systems," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4182–4195, Nov. 2009.

Younes Abdi (S'10) received his B.Sc. degree from the University of Tabriz, Tabriz, Iran, in 2008 and his M.Sc. degree from Tarbiat Modares University, Tehran, Iran, in 2011, both in Electrical Engineering.

From Oct. 2010 to Nov. 2011, he was with Radio Communications Group at Iran Telecommunications Research Center (ITRC), Tehran, Iran, working on standardization and regulatory issues of cognitive radio networks. In 2012, he joined the Faculty of Information Technology at the University of Jyväskylä, Jyväskylä, Finland, where he is currently pursuing his studies towards the Ph.D. degree. He is also a member of working group 1900.1 in the IEEE Dynamic Spectrum Access Networks Standards Committee (DySPAN-SC). His research has been financially supported by Finnish National Graduate School in Electronics, Telecommunications, and Automation (GETA) and Jyväskylä Doctoral Program in Computing and Mathematical Sciences (COMAS). His current research interests are in the areas of advanced signal processing and wireless communications.

Tapani Ristaniemi (SM'11) received his M.Sc. in 1995 (mathematics), Ph.Lic. in 1997 (applied mathematics), and Ph.D. in 2000 (wireless communications), all from the University of Jyväskylä, Jyväskylä, Finland. In 2001, he was appointed as a Professor in the Department of Mathematical Information Technology, University of Jyväskylä. In 2004, he moved to the Department of Communications Engineering, Tampere University of Technology, Tampere, Finland, where he was appointed as a Professor of wireless communications. In 2006, he moved back to the University of Jyväskylä to take up his appointment as a Professor of computer science. He is an Adjunct Professor of Tampere University of Technology. In 2013, he was a Visiting Professor in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He has authored or co-authored over 150 publications in journals, conference proceedings, and invited sessions. He served as a Guest Editor of IEEE Wireless Communications in 2011 and currently he is an Editorial Board Member of Wireless Networks and the International Journal of Communication Systems. His research interests are in the areas of brain and communication signal processing and wireless communication systems.

Besides academic activities, Professor Ristaniemi is also active in the industry. In 2005, he co-founded a start-up, Magister Solutions, Ltd., in Finland, specializing in wireless systems (R&D) for telecom and space industries in Europe. Currently, he serves as a consultant and a Member of the Board of Directors.