Multistatic Cloud Radar Systems:

Joint Waveform and Backhaul Optimization

Seongah Jeong∗, Shahrouz Khalili†, Osvaldo Simeone†,
Alexander Haimovich†, and Joonhyuk Kang∗

∗Department of Electrical Engineering, KAIST, Daejeon, Republic of Korea
†CWCSPR, NJIT, Newark, NJ, USA

Email: seongah@kaist.ac.kr

Abstract

In a multistatic cloud radar system, receive elements (REs) measure signals sent by a transmit element (TE) and reflected from a target and possibly clutter, in the presence of interference and noise. The REs communicate over non-ideal backhaul links with a fusion center (FC), or cloud processor, where the presence or absence of the target is determined. Two different backhaul architectures are considered, namely orthogonal-access and multiple-access backhaul. For the former case, the REs quantize the received baseband signals prior to forwarding them to the FC in order to satisfy the backhaul capacity constraints; instead, in the latter case, the REs amplify and forward the received signals so as to leverage the superposition properties of the backhaul channel. This paper addresses the joint optimization of the sensing and communication functions of the cloud radar system adopting the information-theoretic criterion of the Bhattacharyya distance as a proxy for the detection performance. Specifically, the transmitted waveform is jointly optimized with the backhaul quantization in the case of orthogonal-access and with the amplifying gains of the REs in the case multiple-access backhaul. Algorithmic solutions based on successive convex approximation are developed for instantaneous or stochastic channel state information (CSI) on the REs-to-FC channels. Numerical results demonstrate that the proposed schemes outperform conventional solutions that perform separate optimizations of the waveform and backhaul operation, as well as the standard
distributed detection approach.

**Index Terms**

Multistatic radar, cloud radar, detection, waveform design, quantization, power allocation.

I. INTRODUCTION

In radar systems, the design of the waveform sent by the transmit element (TE) for the detection of the presence of a target has received significant interest due to its role by controlling the response to the target and to the clutter [1], [2]. For monostatic radar systems, i.e., with a single TE and a single receive element (RE), the waveform design in terms of the Neyman Pearson (NP) criterion is studied in [3]. In a multistatic radar system, multiple REs communicate to a fusion center (FC), where target detection is performed. In this case, the performance of the NP optimal detector is in general too complex to be suitable as a design metric. As a result, various information-theoretic criteria such as the Bhattacharyya distance, the Kullback Leibler divergence, the J-divergence and the mutual information, which can be shown to provide various bounds to the probability of error (missed detection, false alarm and Bayesian risk), have been considered as alternative design metrics [4]–[6].

While [4]–[6] assume ideal backhaul links, in practice, the REs may be connected to the FC via a wired backhaul, e.g., via coaxial cables or fiber optic links, or through a wireless backhaul, e.g., microwave links. Here, borrowing the nomenclature from Cloud-Radio Access Networks (C-RANs) in communications [7], we refer to multistatic radar systems with non-ideal backhaul links as cloud radar systems. In this paper, we investigate two types of backhaul links, namely orthogonal wired backhaul links, as seen in Fig. 1(a), and a (non-orthogonal) multiple-access wireless channel, as illustrated in Fig. 1(b).

For the case of orthogonal-access wired backhaul in Fig. 1(a), we assume that the backhaul links have finite capacity. Inspired by the C-RAN architecture [7], in order to satisfy the backhaul capacity constraints, the REs quantize the received baseband signals prior to transmission to the FC. With a multiple-access wireless backhaul channel, each RE amplifies and forwards the received signal to the FC.
so that the signals transmitted by the REs are superimposed at the FC (see, e.g., [8], [9]). In both cases, we formulate the problem of jointly designing the waveform, or code vector, and the transmission of the REs over the backhaul. In particular, for orthogonal-access wired backhaul links, we jointly optimize the code vector and the quantization strategy, while joint optimization is carried out over the code vector and the amplifying gains of the REs in the case of a multiple-access wireless backhaul channel. In both cases, we adopt the information-theoretic criterion of the Bhattacharyya distance in order to account for the detection performance [4]–[6]. Throughout, only stochastic information is assumed on the channel gains between target or clutter and the REs. Furthermore, for the multiple-access case, we consider both instantaneous and stochastic channel state information (CSI) on the REs-to-FC wireless channels.

Prior related work has considered the problem of waveform optimization and backhaul design separately. In particular, for the orthogonal-access case, there is a vast literature on the optimization of quantization strategies (see [10]–[12] and references therein). For the multiple-access case, the problem of optimizing the power allocation at the REs was studied in [8], [9] using the minimum mean square error (MMSE) as the performance criterion. Overall, ours seems to be the first work to consider the joint optimization of waveform and backhaul operation. The only exception is our previous work [13], where we presented part of the material on orthogonal-access backhaul (see Remark 1 for further discussion).

The rest of the paper is organized as follows. In Section II we present the detection problems as well as review the optimal detectors for both orthogonal-access and multiple-access backhaul channels. Section III investigates the multistatic cloud radar system with orthogonal backhaul links. For reference, we first consider the standard distributed detection approach combining local hard decisions and majority rule at the FC (see, e.g., [14], [15]) in Section III-A. Then, we formulate the problem of interest in Section III-C and propose a solution based on successive convex approximations [16], [17] in Section III-D. In Section IV we consider multiple-access wireless backhaul channels. In particular, Section IV-A formulates the problems for short-term and long-term adaptive designs that use instantaneous and stochastic CSI about the REs-to-FC channels, respectively, and Section IV-B and Section IV-C propose algorithmic solutions
II. SYSTEM MODEL

Consider a multistatic cloud radar system consisting of a TE, $N$ REs, or sensors, and a FC, or cloud processor, as illustrated in Fig. I. The REs communicate with the FC over orthogonal-access wired backhaul (Fig. I(a)) or a multiple-access wireless backhaul (Fig. I(b)). All the nodes are equipped with single antennas, and the set of REs is denoted as $\mathcal{N} = \{1, \ldots, N\}$.

The system aims to detect the presence of a single stationary target in a homogeneous clutter field. To this end, each RE receives a noisy version of the signal transmitted by the TE and reflected from the surveillance area, which is conveyed to the FC on the backhaul channels as discussed below. It is assumed that perfect timing information is available at the FC, such that samples of the received signal may be associated with specific locations in some coordinate system. For such a location, and based on all the forwarded signals from the different REs, the FC makes a decision about the presence of the target (see, e.g., [6]).

A. Signal Model

Let the transmitted waveform be a train of $K$ standard pulses with complex amplitudes $\mathbf{x} = [x_1, \ldots, x_K]^T$. The set of amplitudes $\mathbf{x}$ is referred to as a code vector [6]. The pulses may form a continuous waveform or serve as individual pulses in a coherent pulse interval. The design of the code vector $\mathbf{x}$ determines the range resolution and clutter response, and thus has a key role in the performance of the radar system.

The Swerling I model is assumed for the amplitude of the target echo, and hence the return has a Rayleigh envelope, which is fixed during the observation interval. The clutter is homogeneous and fixed over the observation interval with a complex-valued Gaussian distribution across the sensors. The returns are assumed to be independent between REs and between target and clutter. Finally, each RE observes
time-correlated complex Gaussian noise that captures the possible presence of various types of interference and jamming signals.

Following the discussion above, the $K \times 1$ discrete received signal by RE $n$, for $n \in \mathcal{N}$, after matched filtering and symbol-rate sampling, is given by

$$H_0: r_n = c_n + w_n, \quad (1a)$$

$$H_1: r_n = s_n + c_n + w_n, \quad n \in \mathcal{N}, \quad (1b)$$

where $H_0$ and $H_1$ represent the hypotheses under which the target is absent or present, respectively, and $s_n$ denotes the signal received from the target at RE $n$, which is given by

$$s_n = h_n x, \quad (2)$$

with $h_n \sim \mathcal{CN}(0, \sigma_{t,n}^2)$ being the random complex amplitude of the target return. The vector $c_n$ represents the clutter contribution, modeled as

$$c_n = g_n x, \quad (3)$$

where $g_n \sim \mathcal{CN}(0, \sigma_{c,n}^2)$ is the random complex amplitude of the clutter. The term $w_n \sim \mathcal{CN}(0, \Omega_{w,n})$ represents signal-independent interference, which aggregates the contributions of thermal noise, interference and jamming, and whose temporal correlation is described by the covariance matrix $\Omega_{w,n}$. The variables $h_n$, $g_n$ and $w_n$ for all $n \in \mathcal{N}$, are assumed to be independent for different values of $n$ and their second-order statistics $\sigma_{t,n}^2$, $\sigma_{c,n}^2$ and $\Omega_{w,n}$ are assumed to be known to the FC, for all $n \in \mathcal{N}$, e.g., from prior measurements or prior information.

In the following, we describe the received signals at the FC, and review the optimal detectors for both the orthogonal-access wired backhaul and the multiple-access wireless backhaul.

1) Orthogonal-Access Wired Backhaul: For the orthogonal-access backhaul case, each RE $n$ is connected to the FC via an orthogonal wired link of limited capacity. Specifically, the capacity between each RE $n$ and the FC is $C_n$ (bits/s/Hz), where the normalization is with respect to the bandwidth of the signal
transmitted by the TE. The capacity $C_n$ is known to the FC and is assumed to change sufficiently slowly to enable the adaptation of code vector and quantization strategy to the capacities $C_n$ for all $n \in \mathcal{N}$.

Following the C-RAN paradigm [7], in order to satisfy the backhaul capacity constraints, each RE quantizes the received vector $\mathbf{r}_n$, and sends the quantized vector $\tilde{\mathbf{r}}_n$ to the FC. Note that, since the RE does not know whether the target is present or not, the quantizer cannot depend on the correct hypothesis $\mathcal{H}_0$ or $\mathcal{H}_1$. In order to facilitate analysis and design, we follow the standard approach of modeling the effect of quantization by means of an additive quantization noise (see, e.g., [18], [19]). The signal received by the FC from RE $n$ is hence given by

$$
\mathcal{H}_0 : \quad \tilde{\mathbf{r}}_n = \mathbf{r}_n + \mathbf{q}_n = \mathbf{c}_n + \mathbf{w}_n + \mathbf{q}_n,
$$

$$
\mathcal{H}_1 : \quad \tilde{\mathbf{r}}_n = \mathbf{r}_n + \mathbf{q}_n = \mathbf{s}_n + \mathbf{c}_n + \mathbf{w}_n + \mathbf{q}_n,
$$

where $\mathbf{q}_n \sim \mathcal{CN}(\mathbf{0}, \Omega_{q,n})$ is the quantization error vector. As discussed in [20], a Gaussian quantization noise $\mathbf{q}_n$ with any covariance $\Omega_{q,n}$ can in practice be realized via a linear transform, obtained from the eigenvectors of $\Omega_{q,n}$, followed by a multi-dimensional dithered lattice quantizer such as Trellis Coded Quantization (TCQ) [21]. We recall that a lattice quantizer is characterized by a regular grid of quantization levels, such as the standard scalar uniform quantizer and two-dimensional hexagonal quantizer [20]. Further discussion on the validity of the additive model (4) in the context of the design problem of interest can be found in Section III. As further elaborated on Section III, the covariance matrix $\Omega_{q,n}$ determines the bit rate required for backhaul communication between RE $n$ and the FC [18], [19] and is subject to design.

To set the model (4) in a more convenient form, the signal received at the FC is whitened with respect to the overall additive noise $\mathbf{c}_n + \mathbf{w}_n + \mathbf{q}_n$, and the returns from all sensors are collected, leading to

$$
\mathcal{H}_0 : \quad \mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})
$$

$$
\mathcal{H}_1 : \quad \mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{DSD} + \mathbf{I}),
$$

where $\mathbf{y} = [\mathbf{y}_1^T, \ldots, \mathbf{y}_N^T]^T$, $\mathbf{y}_n = \mathbf{D}_n \tilde{\mathbf{r}}_n$, $\mathbf{D}_n$ is the whitening matrix associated with the RE $n$ and is given by $\mathbf{D}_n = (\sigma_{c,n}^2 \mathbf{xx}^H + \Omega_{w,n} + \Omega_{q,n})^{-1/2}$, $\mathbf{D}$ is the block diagonal matrix $\mathbf{D} = \text{diag}\{\mathbf{D}_1, \ldots, \mathbf{D}_N\}$, and $\mathbf{S}$ is the block diagonal matrix $\mathbf{S} = \text{diag}\{\sigma_{t,1}^2 \mathbf{xx}^H, \ldots, \sigma_{t,N}^2 \mathbf{xx}^H\}$. The detection problem described by (5) has
the standard NP solution given by the test

\[ y^H T y \begin{cases} \mathcal{H}_1 & \text{if } y^H T y \geq \nu, \\ \mathcal{H}_0 & \text{if } y^H T y < \nu, \end{cases} \] (6)

where we have defined \( T = D S D (D S D + I)^{-1} \), and the threshold \( \nu \) is set based on the tolerated false alarm probability [22].

2) Multiple-Access Wireless Backhaul: For the case of a multiple-access wireless backhaul channel, the signal received at the FC is given by the superposition of the signals sent by all REs, as

\[ \tilde{r} = \sum_{n=1}^{N} f_n \tilde{x}_n + z, \] (7)

where \( \tilde{x}_n \) is the signal transmitted by RE \( n \), on which we impose the transmit power constraint \( x^H x \leq P_T \); \( f_n \sim \mathcal{CN}(0, \sigma^2_{f_n}) \) is the complex-valued channel gain between RE \( n \) and the FC; and the noise vector \( z \sim \mathcal{CN}(0, \Omega_z) \) is temporally correlated with correlation matrix \( \Omega_z \). The variables \( h_n, g_n, w_n, f_n \) and \( z \), for all \( n \in \mathcal{N} \), are assumed to be mutually independent. Each RE \( n \) communicates the received signal \( r_n \) in (1) to the FC after amplification by a complex coefficient \( \alpha_n \). This choice is due to its capability to leverage the superposition property of the channel (7) and follows the previous references [8], [9]. The signal received at the FC is hence given as

\[ \mathcal{H}_0 : \tilde{r} = \sum_{n=1}^{N} f_n \tilde{x}_n + z = \sum_{n=1}^{N} f_n \alpha_n r_n + z \]

\[ = \sum_{n=1}^{N} (f_n \alpha_n c_n + f_n \alpha_n w_n) + z \] (8a)

\[ \mathcal{H}_1 : \tilde{r} = \sum_{n=1}^{N} f_n \tilde{x}_n + z = \sum_{n=1}^{N} f_n \alpha_n r_n + z \]

\[ = \sum_{n=1}^{N} (f_n \alpha_n s_n + f_n \alpha_n c_n + f_n \alpha_n w_n) + z. \] (8b)

Based on prior information or measurements, the second-order statistics of the channel gains between the target and the REs, and of the noise terms, namely \( \sigma^2_{t,n}, \sigma^2_{c,n}, \Omega_{w,n} \) and \( \Omega_z \), are assumed to be known to the FC for all \( n \in \mathcal{N} \). The REs-to-FC channels \( f = [f_1 \cdots f_N]^T \) are assumed to be known at the FC,
via training and channel estimation. Moreover, since only the second-order statistics of the channel gains $h_n, n \in \mathcal{N}$, are known to the REs and the FC, no coherent gains may be achieved by optimizing the amplifying gains, and hence it will be seen that one can focus, without loss of optimality, only on the REs’ power gains $p = [p_1 \cdots p_N]^T$, with $p_n = |\alpha_n|^2, n \in \mathcal{N}$.

As in Section II-A1 we can write the hypotheses (8) in a standard form by whitening the signal received at the FC, and consequently the detection problem can be expressed as (5), where we have redefined $y = D\tilde{r}$; $D = (\sum_{n=1}^{N}(|f_n|^2p_n\sigma_{c,n}^2xx^H + |f_n|^2p_n\Omega_{w,n}) + \Omega_x)^{-1/2}$ is the whitening filter with respect to the overall additive noise $\sum_{n=1}^{N}(f_n\alpha_n c_n + f_n\alpha_n w_n) + z$; and $S = \sum_{n=1}^{N}|f_n|^2p_n\sigma_{t,n}^2xx^H$ is the correlation matrix of the desired signal part. Accordingly, the detection problem has the standard estimator-correlator solution given by the test in (6).

III. ORTHOGONAL-ACCESS WIRED BACKHAUL

In this section, we aim to find the optimum code vector $x$ and quantization error covariance matrices $\Omega_{q,n}$ in (4), for $n \in \mathcal{N}$, for given backhaul capacity constraints $C_n$, for all $n \in \mathcal{N}$. To this end, for the sake of tractability, we resort to the information-theoretic metric of the Bhattacharyya distance between the distributions of the quantized received signal (4) under the two hypotheses as a proxy to the detection performance. Before we proceed, for reference, we first discuss the standard distributed detection approach that combines hard local decisions at the REs and a majority-rule detection at the FC (see, e.g., [14], [15]).

A. Distributed Detection

Here, we describe the standard distributed detection approach applied to multistatic radar system (see, e.g., [14], [15]). With this approach, each RE $n$ makes its own decision based on the likelihood test, which is given by

$$y_n^HT_ny_n \gtrless \gamma_n, \quad \mathcal{H}_1 \quad \mathcal{H}_0$$

(9)
where $\gamma_n$ is the threshold for RE $n$, which is calculated based on the tolerated false alarm probability \cite{22}, and we have defined $T_n = D_n S_n D_n (D_n S_n D_n + I)^{-1}$ with $y_n = D_n r_n$, $D_n = (\sigma_{t,n}^2 x x^H + \Omega_{w,n})^{-1/2}$ and $S_n = \sigma_{t,n}^2 x x^H$, for all $n \in \mathcal{N}$. The REs transmit the obtained one-bit hard decision to the FC. Note that this scheme is feasible as long as the backhaul capacity available for each RE-to-FC channel is larger than or equal to 1 bits/s/Hz, i.e., $C_n \geq 1$, for $n \in \mathcal{N}$. The FC decides on the target’s presence based on the majority rule: if the number of REs $k$ that decide for $H_0$ satisfies $k \geq N/2$, the FC chooses $H_0$, and vice versa if $k \leq N/2$.

**B. Cloud Processing**

The distributed detection approach does not fully leverage the available backhaul resources, particularly if $C_n \geq 1$ bits/s/Hz. This motivates the cloud radar system introduced in the previous section. To start, we discuss the criterion that is adopted to account for the detection performance, namely the Bhattacharyya distance and the approach used to model the effect of the quantizers at the REs.

1) **Bhattacharyya Distance:** For two zero-mean Gaussian distributions with covariance matrix of $\Sigma_1$ and $\Sigma_2$, the Bhattacharyya distance $B$ is given by \cite{4}

\[
B = \frac{|0.5(\Sigma_1 + \Sigma_2)|}{\sqrt{|\Sigma_1| |\Sigma_2|}}.
\tag{10}
\]

Therefore, for the signal model \cite{4}, the Bhattacharyya distance between the distributions under the two hypotheses can be calculated as

\[
B(x, \Omega_q) = \log \left( \frac{|I + 0.5 DSD|}{\sqrt{|I + DSD|}} \right) = \sum_{n=1}^{N} B_n(x, \Omega_{q,n}) = \sum_{n=1}^{N} \log \left( \frac{1 + 0.5 \lambda_n}{\sqrt{1 + \lambda_n}} \right),
\tag{11}
\]

where we have made explicit the dependence on $x$ and $\Omega_{q,n}$; $\Omega_q$ collects all the covariance matrices of quantization noise and is given as $\Omega_q = \{\Omega_{q,n}\}_{n \in \mathcal{N}}$; and we have defined

\[
\lambda_n = \sigma_{t,n}^2 x x^H \left( \sigma_{t,n}^2 x x^H + \Omega_{w,n} + \Omega_{q,n} \right)^{-1} x.
\tag{12}\]
We observe that (11) is valid under the assumption that the effect of the quantizers can be well approximated by additive Gaussian noise as per (4). This is discussed next.

2) *Quantization:* As discussed above, a dithered lattice quantizer is able to realize the additive quantization noise model (4) when operating over a sufficiently large number of measurement vectors (1). Moreover, the required backhaul rate for the transmission of quantization signal is asymptotically equal to the mutual information $I(r_n; \tilde{r}_n)$ [20]. An alternative justification for the model (4) based on rate-distortion theory can be found in Appendix A. These considerations motivate us to choose the mutual information $I(r_n; \tilde{r}_n)$ as a measure of the backhaul rate required for the transmission to the FC.

While the mutual information $I(r_n; \tilde{r}_n)$ depends on the actual hypothesis $H_0$ or $H_1$, it is easy to see that $I(r_n; \tilde{r}_n)$ is larger under hypothesis $H_1$. Based on this, the mutual information $I(r_n; \tilde{r}_n)$ evaluated under $H_1$ is adopted here as the measure of the bit rate required between RE $n$ and the FC. This can be easily calculated as $I(r_n; \tilde{r}_n) = I_n(x, \Omega_{q,n})$ by using the expression of the mutual information for multivariate Gaussian distribution (see, e.g., [19]) with

$$I_n(x, \Omega_{q,n}) = \log \left| I + (\Omega_{q,n})^{-1}\Omega_{w,n} \right| + \log \left( 1 + (\sigma_{t,n}^2 + \sigma_{c,n}^2)x^H(\Omega_{w,n} + \Omega_{q,n})^{-1}x \right),$$

where again we have made explicit the dependence of mutual information on $x$ and $\Omega_{q,n}$.

In the following, we formulate and solve the problem of jointly optimizing the Bhattacharyya distance criterion over the code vector $x$ at the TE and over the covariance matrices $\Omega_q$ of the quantizers at the REs in Section III-C and in Section III-D respectively.
C. Problem Formulation

The problem of maximizing the Bhattacharyya distance in (11) over the code vector \( \mathbf{x} \) and the covariance matrices \( \mathbf{\Omega}_q \) under the backhaul capacity constraints is stated as

\[
\begin{align*}
\text{minimize} & \quad \bar{\mathcal{B}}(\mathbf{x}, \mathbf{\Omega}_q) = \sum_{n=1}^{N} \bar{\mathcal{B}}_n(\mathbf{x}, \mathbf{\Omega}_{q,n}) \\
\text{s.t.} & \quad \mathcal{I}_n(\mathbf{x}, \mathbf{\Omega}_{q,n}) \leq C_n, \quad n \in \mathcal{N}, \\
& \quad \mathbf{x}^H \mathbf{x} \leq P_T, \\
& \quad \mathbf{\Omega}_{q,n} \succeq 0, \quad n \in \mathcal{N},
\end{align*}
\]

(14a)

(14b)

(14c)

(14d)

where we have formulated the problem as the minimization of the negative distance \( \bar{\mathcal{B}}(\mathbf{x}, \mathbf{\Omega}_q) = \sum_{n=1}^{N} \bar{\mathcal{B}}_n(\mathbf{x}, \mathbf{\Omega}_{q,n}) \), with \( \bar{\mathcal{B}}(\mathbf{x}, \mathbf{\Omega}_q) = -\mathcal{B}(\mathbf{x}, \mathbf{\Omega}_q) \) and \( \bar{\mathcal{B}}_n(\mathbf{x}, \mathbf{\Omega}_{q,n}) = -\mathcal{B}_n(\mathbf{x}, \mathbf{\Omega}_{q,n}) \), following the standard convention in [23]. The power of the code \( \mathbf{x} \) is constrained not to exceed a prescribed value of transmit power \( P_T \). We observe that the constraint (14b) ensures that the transmission rate between each RE and the FC is smaller than \( C_n \), according to the adopted information-theoretic metrics. Note also that the problem (14) is not a convex program, since the objective function (14a) and the constraint (14b) are not convex.

Remark 1: In [13], a total backhaul capacity constraint, rather than individual capacity constraints, was imposed. Moreover, as compared to [13], here we provide additional discussion on the adopted additive quantization noise model, and compare the performance to the standard distributed detection approach.

D. Proposed Algorithm

Since both functions \( \bar{\mathcal{B}}_n(\mathbf{x}, \mathbf{\Omega}_{q,n}) \) and \( \mathcal{I}_n(\mathbf{x}, \mathbf{\Omega}_{q,n}) \) in (14) are non-convex in \( \mathbf{x} \) and \( \mathbf{\Omega}_{q,n} \), the optimization problem (14) is not convex, and hence it is difficult to solve. To obtain a locally optimal solution, we approach the joint optimization of \( \mathbf{x} \) and \( \mathbf{\Omega}_q \) in (14) via successive convex approximations. Specifically, in an outer loop, Block Coordinate Descent (BCD) is applied to update \( \mathbf{x} \) and \( \mathbf{\Omega}_q \) one at a time, while an inner loop implemented via Majorization-Minimization (MM) solves the optimization of \( \mathbf{x} \) and \( \mathbf{\Omega}_q \) separately. This approach was first introduced in [13] for a sum-capacity backhaul constraint. By the properties of
MM (see, e.g., [16], [24]), the algorithm provides a sequence of feasible solution with non-increasing cost function, which guarantees convergence of the cost function. Note that, due to the non-convexity of the problem, no claim of convergence to a local or global optimum is made here.

At the $i$th iteration of outer loop, the optimum code vector $\mathbf{x}^{(i)}$ is obtained by solving (14) for matrices $\Omega_q = \Omega_q^{(i-1)}$ obtained at the previous iteration; subsequently, the matrices $\Omega_q^{(i)}$ are calculated by solving (14) with $\mathbf{x} = \mathbf{x}^{(i)}$. These two separate optimizations are carried out by the MM method, which, as described below, requires the solution of an quadratically constrained quadratic programs (QCQP). The proposed algorithm coupling BCD and MM to solve problem (14), is summarized in Table Algorithm 1. Below we present the MM steps and the overall proposed algorithm in detail.

1) MM Method: We start by reviewing the MM method. For a non-convex function $f(t)$ of a generic variable $t$, which may appear either in the cost function or among the constraints, the MM method substitutes at the $l$th iteration, a convex approximation $f(t|t^{(l-1)})$ of $f(t)$, such that the global upper bound property

$$f(t|t^{(l-1)}) \geq f(t),$$

is satisfied for all $t$ in the domain, along with the local tightness condition

$$f(t^{(l-1)}|t^{(l-1)}) = f(t^{(l-1)}).$$

These properties guarantee the feasibility of all iterates and the descent property that the object function does not increase along the iterations.

2) Proposed Algorithm: In the following, we use the superscript $i$ to identify the iterations of the outer loop of Algorithm 1 and the superscript $j$ as the index of the inner iteration of the MM method (e.g., $\mathbf{x}^{(i,j)}$ indicates the code vector optimized at the $j$th iteration of the inner loop of the MM method and the $i$th iteration of the outer loop). In Section III-D2a and Section III-D2b, we discuss the application of the MM method to perform optimizations over $\mathbf{x}$ and $\Omega_q$ in Algorithm 1, respectively.

a) Optimization over $\mathbf{x}$: Here, the goal is to obtain the optimal value of $\mathbf{x}^{(i)}$ for problem (14) given $\Omega_q = \Omega_q^{(i-1)}$. To this end, we apply the MM method. Specifically, at the $j$th iteration of the MM method
and the $i$th iteration of the outer loop, the MM method solves a QCQP and obtains a solution $x^{(i,j)}$ by substituting the non-convex objective function $\tilde{B}(x, \Omega_q)$ with a tight upper bound $U^B(x, \Omega_q|x^{(i,j-1)})$ around the current iterate $x^{(i,j-1)}$. This bound is obtained by linearizing the difference-of-convex functions in $\tilde{B}(x, \Omega_q)$ via the first-order Taylor approximation \cite{16}, which follows the same steps as in \cite{6} eq. (34) and (50) in Section IV, and is given by

$$U^B(x, \Omega_q|x^{(i,j-1)}) = \sum_{n=1}^{N} U^B_n(x, \Omega_{q,n}|x^{(i,j-1)}) = \sum_{n=1}^{N} \phi^{(i,j-1)}_n x^H(\Omega_{w,n} + \Omega_{q,n})^{-1} x - \Re e\left(\left(d^{(i,j-1)}_n\right)^H x\right),$$  \hspace{1cm} (17)

where

$$\phi^{(i,j-1)}_n = \frac{\beta_n}{1 + \beta_n y^{(i,j-1)}_n} + \beta_n (1 + 0.5 \gamma_n) + \frac{0.5 \gamma_n}{1 + \lambda^{(i,j-1)}_n}\left(\lambda^{(i,j-1)}_n\right)^2,$$

$$d^{(i,j-1)}_n = \left(\frac{2 \beta_n (1 + 0.5 \gamma_n)}{1 + \beta_n y^{(i,j-1)}_n (1 + 0.5 \gamma_n)} + 2 \beta_n (1 + 0.5 \gamma_n)\right)(\Omega_{w,n} + \Omega_{q,n})^{-1} x^{(i,j-1)},$$

$$y^{(i,j-1)}_n = (x^{(i,j-1)})^H(\Omega_{w,n} + \Omega_{q,n})^{-1} x^{(i,j-1)}; \text{ and}$$

$$\lambda^{(i,j-1)}_n = \gamma_n - \frac{\gamma_n}{1 + \beta_n y^{(i,j-1)}_n},$$

with $\beta_n = \sigma^2_{c,n}$ and $\gamma_n = \sigma^2_{t,n}/\beta_n$. A bound with the desired property can also be easily derived for $I_n(x, \Omega_{q,n})$ by using the inequality $\log(1 + t) \leq \log(1 + t^{(l)}) + 1/(1 + t^{(l)})(t - t^{(l)})$, for $t = (\sigma^2_{t,n} + \sigma^2_{c,n})x^H(\Omega_{w,n} + \Omega_{q,n})^{-1} x$, leading to

$$U^I_n(x, \Omega_{q,n}|x^{(i,j-1)}) = \log|I + (\Omega_{q,n})^{-1}\Omega_{w,n}| + \log(1 + t^{(i,j-1)})$$

$$+ \frac{1}{1 + t^{(i,j-1)}}((\sigma^2_{c,n} + \sigma^2_{t,n})x^H(\Omega_{w,n} + \Omega_{q,n})^{-1} x - t^{(i,j-1)})^{13}.$$  \hspace{1cm} (18)

At the $j$th iteration of the MM method and the $i$th outer loop, we evaluate the new iterate $x^{(i,j)}$ by solving the following QCQP problem

$$x^{(i,j)} \leftarrow \arg \min_{x} U^B(x, \Omega_q|x^{(i,j-1)})$$  \hspace{1cm} (19a)

s.t.  \hspace{1cm} $U^I_n(x, \Omega_{q,n}|x^{(i,j-1)}) \leq C_n, \ n \in N,$  \hspace{1cm} (19b)

$$x^H x \leq P_T.$$  \hspace{1cm} (19c)

The MM method obtains the solution $x^{(i)}$ for the $i$th iteration of the outer loop by solving the problem \cite{19} iteratively over $j$ until a convergence criterion is satisfied.
b) Optimization over $\Omega_q$: In this part, we consider the optimization of matrices $\Omega_q^{(i)}$ for a given $\mathbf{x} = \mathbf{x}^{(i)}$. Similar to the optimization over $\mathbf{x}^{(i)}$ in Section III-D2a, we use upper bounds of $\bar{B}(\mathbf{x}, \Omega_q)$ and $\bar{T}(\mathbf{x}, \Omega_{q,n})$ for optimization. First, by rewriting $\bar{T}(\mathbf{x}, \Omega_{q,n}) = \log |\Omega_{q,n} + (\sigma_{t,n}^2 + \sigma_{c,n}^2) \mathbf{x} \mathbf{x}^H + \Omega_{w,n}| - \log |\Omega_{q,n}|$, we obtain difference-of-convex functions with respect to $\Omega_{q,n}$. Then, by linearizing negative convex component via its first-order Taylor approximation, upper bounds $U^{T}(\mathbf{x}, \Omega_{q,n}|\Omega_{q,n}^{(i,j-1)})$ and $U^{B}(\mathbf{x}, \Omega_q|\Omega_q^{(i,j-1)})$ with the desired properties (15) and (16) are derived for functions $\bar{T}(\mathbf{x}, \Omega_{q,n})$ and $\bar{B}(\mathbf{x}, \Omega_q)$, respectively, as follows

$$U^{T}(\mathbf{x}, \Omega_{q,n}|\Omega_{q,n}^{(i,j-1)}) = \log |\Omega_{q,n}^{(i,j-1)} + (\sigma_{t,n}^2 + \sigma_{c,n}^2) \mathbf{x} \mathbf{x}^H + \Omega_{w,n}|$$
$$- \log |\Omega_{q,n}| + \text{Tr} \left\{ (\Omega_{q,n}^{(i,j-1)} + (\sigma_{t,n}^2 + \sigma_{c,n}^2) \mathbf{x} \mathbf{x}^H + \Omega_{w,n})^{-1} (\Omega_{q,n} - \Omega_{q,n}^{(i,j-1)}) \right\}$$

and

$$U^{B}(\mathbf{x}, \Omega_q|\Omega_q^{(i,j-1)}) = \sum_{n=1}^{N} U^{B}_{n}(\mathbf{x}, \Omega_{q,n}|\Omega_{q,n}^{(i,j-1)})$$
$$= \sum_{n=1}^{N} - \log |(0.5 \sigma_{t,n}^2 + \sigma_{c,n}^2) \mathbf{x} \mathbf{x}^H + \Omega_{w,n} + \Omega_{q,n}|$$
$$+ 0.5 \text{Tr} \left\{ (0.5 \sigma_{t,n}^2 + \sigma_{c,n}^2) \mathbf{x} \mathbf{x}^H + \Omega_{w,n} + \Omega_{q,n}^{(i,j-1)})^{-1} \Omega_{q,n} \right\}$$
$$+ 0.5 \text{Tr} \left\{ (0.5 \sigma_{c,n}^2 \mathbf{x} \mathbf{x}^H + \Omega_{w,n} + \Omega_{q,n}^{(i,j-1)})^{-1} \Omega_{q,n} \right\}.$$  

The $j$th iteration of the MM method then evaluates the matrices $\Omega_q^{(i,j)} = \{\Omega_q^{(i,j)}\}_{n \in \mathcal{N}}$ by solving the following convex optimization problem

$$\Omega_q^{(i,j)} \leftarrow \text{argmin}_{\Omega_q^{(i,j)}} U^{B}(\mathbf{x}, \Omega_q|\Omega_q^{(i,j-1)})$$

s.t. $U^{T}(\mathbf{x}, \Omega_{q,n}|\Omega_{q,n}^{(i,j-1)}) \leq C_n, \ n \in \mathcal{N},$  

$$\Omega_{q,n} \succeq 0, \ n \in \mathcal{N}.$$ 

By repeating the procedure (22) over $j$ until the convergence is attained, the solution $\Omega_q^{(i)}$ is obtained for the $i$th outer loop.
IV. MULTIPLE-ACCESS WIRELESS BACKHAUL

In this section, we consider the case of a multiple-access wireless backhaul, for which we seek to optimize the detection performance with respect to the code vector $x$ and the power gains $p$, under power constraints on the TE and REs. As in the orthogonal-access backhaul links, we adopt the Bhattacharyya distance as the performance metric. As per [10], the Bhattacharyya distance between the distributions (8) under the two hypotheses $H_0$ and $H_1$ can be calculated as

$$B(x, p; f) = \log \left( \frac{|I + 0.5DDSD|}{\sqrt{|I + DSD|}} \right) = \log \frac{1 + 0.5\lambda}{\sqrt{1 + \lambda}}, \quad (23)$$

where $\lambda = f^H P \Sigma_t f x^H (f^H P \Sigma_c f x x^H + (f \otimes I_K)(P \otimes I_K)\Omega_w(f \otimes I_K) + \Omega_z)^{-1} x; \quad \Sigma_t = \text{diag}\{\sigma^2_{t,1}, \ldots, \sigma^2_{t,N}\}$ and $\Sigma_c = \text{diag}\{\sigma^2_{c,1}, \ldots, \sigma^2_{c,N}\}$ are the diagonal matrices whose components are the second-order statistics of channel amplitudes of target return and clutter, respectively; $\Omega_w = \text{diag}\{\Omega_{w,1}, \ldots, \Omega_{w,N}\} \in R^{NK \times NK}$ is a block diagonal matrix containing all the noise covariance matrices at the REs; and $P = \text{diag}\{p\} \in R^{N \times N}$ is the diagonal matrix that contains the RE’s power gains. Note that we have made explicit the dependence of the Bhattacharyya distance $B(x, p; f)$ on the channels $f$ at the FC, as well as on the code vector $x$ and the REs’ power gains $p$.

A. Problem Formulation

Here, the goal is to jointly optimize the code vector $x$ used at the TE and power gains $p$ used at the REs with the aim of maximizing the Bhattacharyya distance under TE’s and REs’ power constraints. We consider an adaptive design in which the code vector and gains may depend on the instantaneous gains of the CSI or on the stochastic CSI of the REs-to-FC channels. The former is referred to as short-term adaptive design while the latter is called as long-term adaptive design. In the following discussion, we formulate the problems for short-term adaptive design in Section IV-A1 and long-term adaptive design in Section IV-A2.

1) Short-Term Adaptive Design: Here, we consider the case in which design of the code vector and of the REs’ gains depends on the instantaneous gain of the CSI of the REs-to-FC channels $f$. Note that
this design requires to modify the solution vector \((x, p)\) at the time scale at which the channel vector \(f\) varies, hence entailing a potentially large feedback overhead from the FC to the REs and the TE. The problem of maximizing the Bhattacharyya distance \((23)\) over the code vector \(x\) and the power gains \(p\) under the power constraints for TE and REs, is stated as

\[
\begin{align*}
\text{minimize} & \quad \bar{B}(x, p; f) \\
\text{s.t.} & \quad x^H x \leq P_T, \quad (24b) \\
& \quad 1^T p \leq P_R, \quad (24c) \\
& \quad p_n \geq 0, \quad n \in \mathcal{N}, \quad (24d)
\end{align*}
\]

where we have defined \(\bar{B}(x, p; f) = -B(x, p; f)\) to formulate the problem as the minimization of the negative Bhattacharyya distance \(\bar{B}(x, p; f)\). We observe that the problem \((24)\) can be easily modified to include individual power constraints at the REs, but this is not further explored here. Note also that the problem \((24)\) is not a convex program, since the objective function \((24a)\) is not convex.

2) Long-Term Adaptive Design: Here, in order to avoid the possibly excessive feedback overhead between FC and the TE and REs of the short-term adaptive solution, we adopt the average Bhattacharyya distance, as the performance criterion, where the average is taken with respect to the distribution of the REs-to-FC channels \(f\). In this way, the code vector \(x\) and REs’ gains \(p\) have to be updated only at the time scale at which the statistics of channels and noise terms vary. Then, the problem for the long-term adaptive design is formulated from problem \((24)\) by substituting the objective function \(\bar{B}(x, p; f)\) with \(E_f[\bar{B}(x, p; f)]\), yielding

\[
\begin{align*}
\text{minimize} & \quad E_f \left[\bar{B}(x, p; f)\right] \\
\text{s.t.} & \quad (24b) - (24d).
\end{align*}
\]

Note that the problem \((25)\) is a stochastic program with a non-convex objective function \((25a)\).
B. Short-Term Adaptive Design

Here, we propose an algorithm to solve the optimization problem (24). As in Section III-D, due to the difficulty of obtaining a global optimal solution, we aim to develop a descent algorithm and adopts the BCD method coupled with MM. The proposed algorithm is summarized in Table Algorithm 2 and further discussed next. Note that we use the superscript $i$ to identify the iterations of the outer loop of Algorithm 2 and the superscript $j$ as the index of the inner iteration of the MM method.

1) Optimization over $x$: Here, the goal is to optimize the objective function (24) over the code vector $x^{(i)}$ given the gains $p = p^{(i-1)}$. For this purpose, we apply the MM method. Specifically, at the $j$th iteration of the MM method and the $i$th iteration of the outer loop, the MM method solves a convex QCQP and obtains a solution $x^{(i,j)}$ by substituting the non-convex objective function $\tilde{B}(x, p; f)$ with a tight upper bound $U(x, p; f|x^{(i,j-1)})$ around the current iterate $x^{(i,j-1)}$. This bound is obtained by following the same steps as in Section III-D2a and is given by

$$U(x, p; f|x^{(i,j-1)}) = \phi^{(i,j-1)}x^H \left( (f \otimes I_K)^H (P \otimes I_K)\Omega_w (f \otimes I_K) + \Omega_z \right)^{-1} x - \Re \left\{ (d^{(i,j-1)})^H x \right\},$$

where

$$\phi^{(i,j-1)} = \frac{\beta}{1 + \beta y^{(i,j-1)}} + \frac{0.5 \gamma}{1 + \chi^{(i,j-1)}} \frac{\beta}{(1 + \beta y^{(i,j-1)})^2};$$

$$d^{(i,j-1)} = \left( \frac{2\beta (1 + 0.5 \gamma)}{1 + \beta y^{(i,j-1)} (1 + 0.5 \gamma)} + 2\beta (1 + 0.5 \gamma) \right) \left( (f \otimes I_K)^H (P \otimes I_K)\Omega_w (f \otimes I_K) + \Omega_z \right)^{-1} x^{(i,j-1)};$$

$$\beta = f^H P\Sigma_e f;$$

$$\gamma = \frac{f^H P\Sigma_e f}{\beta};$$

$$y^{(i,j-1)} = (x^{(i,j-1)})^H \left( (f \otimes I_K)^H (P \otimes I_K)\Omega_w (f \otimes I_K) + \Omega_z \right)^{-1} x^{(i,j-1)};$$

and

$$\chi^{(i,j-1)} = \gamma - \frac{\gamma}{1 + \beta y^{(i,j-1)}},$$
At the $j$th iteration of the MM method and the $i$th outer loop, we evaluate the new iterate $x^{(i,j)}$ by solving the following QCQP problem

\[
x^{(i,j)} \leftarrow \arg\min_x U(x,p,f|x^{(i,j-1)}) \quad (27a)
\]
\[
\text{s.t.} \quad x^H x \leq P_T. \quad (27b)
\]

The MM method obtains the solution $x^{(i)}$ for the $i$th iteration of the outer loop by solving the problem \((27)\) iteratively over $j$ until a convergence criterion is satisfied.

2) **Optimization over $p$**: We consider now the optimization of the gains $p^{(i)}$, when the code vector $x = x^{(i)}$ is given. Similar to the optimization over $x^{(i)}$ in the previous section, we also use the MM method for the optimization over $p$. Towards this goal, we obtain the upper bound $U(x,p,f|p^{(i,j-1)})$ of the objective function $\bar{B}(x,p,f)$ around the current iterate $p^{(i,j-1)}$. This bound is derived by linearizing the difference-of-convex functions via the first-order Taylor approximation \([16]\). The following bound can then be obtained

\[
U(x,p,f|p^{(i,j-1)}) = -\ln \left| P (0.5\Sigma_t + \Sigma_c) f xx^H + (f \otimes I_K)^H (P \otimes I_K) \Omega_w (f \otimes I_K) + \Omega_z \right|
+ 0.5 \text{tr} \left\{ \left( P (0.5\Sigma_t + \Sigma_c) f xx^H + (f \otimes I_K)^H (P^{(i,j-1)} \otimes I_K) \Omega_w (f \otimes I_K) + \Omega_z \right)^{-1}
\right.
\]
\[
\left. \left( f^H P \Sigma_t + \Sigma_c f xx^H + (f \otimes I_K)^H (P^{(i,j-1)} \otimes I_K) \Omega_w (f \otimes I_K) + \Omega_z \right) \right\}
+ 0.5 \text{tr} \left\{ \left( P^{(i,j-1)} \Sigma_c f xx^H + (f \otimes I_K)^H (P^{(i,j-1)} \otimes I_K) \Omega_w (f \otimes I_K) + \Omega_z \right)^{-1}
\right.
\]
\[
\left. \left( f^H P\Sigma_c f xx^H + (f \otimes I_K)^H (P \otimes I_K) \Omega_w (f \otimes I_K) \right) \right\} . \quad (28)
\]

Then, the new iterate $p^{(i,j)}$ at the $j$th iteration of the MM method and the $i$ iteration of the outer loop can be obtained by solving the following optimization problem:

\[
p^{(i,j)} \leftarrow \arg\min_p U(x,p,f|p^{(i,j-1)}) \quad (29a)
\]
\[
\text{s.t.} \quad 1^T p \leq P_R. \quad (29b)
\]
\[
p_n \geq 0, \quad n \in \mathcal{N}. \quad (29c)
\]
By repeating the procedure (29) over $j$ until a convergence criterion is satisfied, the solution $p^{(i)}$ is determined for the $i$th outer loop.

3) Proposed Algorithm: In summary, in order to solve problem (24), we propose an algorithm (described in Table Algorithm 2) that alternates between the optimization over $x$, described in Section IV-B1 and the optimization over $p$, discussed in Section IV-B2. In particular, at the $i$th iteration of the outer loop, the iterate $x^{(i)}$ is obtained by solving a sequence of convex problems (Section IV-B1) via the MM method for a fixed $p = p^{(i-1)}$. Then, the iterate $p^{(i)}$ is found by solving a sequence of convex problems (Section IV-B2) via the MM method with $x = x^{(i)}$ attained in the previous step. According to the properties of the MM method [16], [24], the proposed scheme yields a non-increasing objective function along the outer and inner iterations, hence ensuring convergence of the cost function.

C. Long-Term Adaptive Design

In order to prevent the possibly excessive feedback overhead from the FC to the REs and the TE, we consider a long-term adaptive design and solve the problem (25). Since the stochastic program (25) has a non-convex objective function, we apply the stochastic successive upper-bound minimization method (SSUM) [17], which minimizes at each step an approximate ensemble average of a locally tight upper bound of the cost function. Specifically, we develop a BCD scheme similar to the one detailed in Table Algorithm 2 that uses SSUM in lieu of the MM scheme. We first present the optimization over the code vector $x$ given the gains $p$ via SSUM, and then describe the optimization over $p$ with fixed $x$ via SSUM. Here, similar to Algorithm 2, we use the superscript $i$ to identify the iterations of the outer loop, and the superscript $j$ as the index of the inner iterations of SSUM.

1) Optimization over $x$: Following the SSUM scheme, at the $j$th inner iteration and the $i$th outer iteration, we optimize the code vector $x^{(i,j)}$ given $p = p^{(i-1)}$ by solving the following convex problem

$$x^{(i,j)} \leftarrow \arg\min_x \frac{1}{j} \sum_{l=1}^j U^{(l)}(x, p; f^{(l)} | x^{(i,l-1)})$$

s.t. $x^H x \leq P_T$.
where \( f^{(l)} \) denotes a channel vector \( f \) for the FC that is randomly and independently generated at the \( l \)th iteration according to the known distribution of \( f \), and \( U^{(l)}(x, p; f^{(l)}|x^{(i,l-1)}) \) is the locally tight convex upper bound (26) on the negative Bhattacharyya distance around the point \( x^{(i,l-1)} \). Note that the cost function (30a) depends on all the realizations of the channel vectors \( f^{(l)} \) for \( l = 1, \ldots, j \). The solution \( x^{(i)} \) for the \( i \)th iteration of the outer loop is obtained by solving the problem (30) iteratively over \( j \), until a convergence criterion is satisfied.

2) Optimization over \( p \): With the optimized code vector \( x = x^{(i)} \), SSUM calculates the iterates \( p^{(i,j)} \) by solving iteratively the following problems

\[
\begin{align*}
p^{(i,j)} &\leftarrow \underset{p}{\text{argmin}} \quad \frac{1}{j} \sum_{l=1}^{j} U^{(l)}(x, p; f^{(l)}|p^{(i,l-1)}) \\
&\text{s.t.} \quad 1^T p \leq P_R, \\
p_n &\geq 0, \quad n \in \mathcal{N},
\end{align*}
\]

where \( U^{(l)}(x, p; f^{(l)}|p^{(i,l-1)}) \) is the convex upper bound (28) on the negative Bhattacharyya distance around the point \( p^{(i,l-1)} \). The iterate \( p^{(i)} \) is obtained by solving the problem (31) iteratively over \( j \) until convergence of the cost function.

3) Proposed Algorithm: The final algorithm for long-term adaptive design can be summarized as in Table Algorithm 2 by substituting (27) and (29) with (30) and (31), respectively.

V. NUMERICAL RESULTS

In the following, the performance of the proposed algorithms that perform joint optimization of the code vector \( x \) and of the quantization noise covariance matrices \( \Omega_q \) for the orthogonal-access wired backhaul case, and of the code vector \( x \) and of the power gains \( p \) for the multiple-access wireless backhaul case, are investigated via numerical results in Section V-A and in Section V-B respectively. Throughout, we set the length of the code vector to \( K = 6 \) and the variances of the target amplitudes as \( \sigma_{t,n}^2 = 1 \) for \( n \in \mathcal{N} \). Moreover, we model the noise with covariance matrices \( [\Omega_w,n]_{i,j} = (1 - 0.12n)^{|i-j|} \) and
\[ [\Omega_z]_{i,j} = (1 - 0.45)^{|i-j|} \] as in [6], hence accounting for temporally correlated interference. The channel coefficients \( f_n \) have unit variance, i.e., \( \sigma_{f_n}^2 = 1 \).

A. Orthogonal-Access Wired Backhaul

In this section, the performance of the proposed joint optimization of the code vector \( \mathbf{x} \) and of the quantization noise covariance matrices \( \Omega_q \) in Section III is verified via numerical results. For reference, we consider the performance of the following strategies: (i) No optimization (No opt.): Set \( \mathbf{x} = \sqrt{P_T/K} \mathbf{1}_K \) and \( \Omega_{q,n} = \epsilon \mathbf{I} \), for \( n \in \mathcal{N} \), where \( \epsilon \) is a constant that is found by satisfying the constraint (14b) with equality; (ii) Code vector optimization (Code opt.): Optimize the code vector \( \mathbf{x} \) by using the algorithm in [6], which is given in Algorithm 1 by setting \( \Omega_{q,n} = 0 \) for \( n \in \mathcal{N} \), and set \( \Omega_{q,n} = \epsilon \mathbf{I} \), for \( n \in \mathcal{N} \), as explained above; (iii) Quantization noise optimization (Quantization opt.): Optimize the covariance matrices \( \Omega_q \) as per Algorithm 1 with \( \mathbf{x} = \sqrt{P_T/K} \mathbf{1}_K \). In the following, we set the number of REs, the transmit power and the variance of the clutter amplitudes as \( N = 3 \), \( P_T = 10 \) dB, \( \sigma_{c,1}^2 = 0.125, \sigma_{c,2}^2 = 0.25 \) and \( \sigma_{c,3}^2 = 0.5 \), respectively. Also, the backhaul rate constraints \( C_n \) are assumed to be equal, i.e., \( C_n = C \) for all \( n \in \mathcal{N} \).

In Fig. 2 the Bhattacharyya distance is plotted versus the available backhaul capacity \( C \). For intermediate and large values of \( C \), the proposed joint optimization of code vector and quantization noise is seen to be significantly beneficial over all separate optimization strategies. In order to study the actual detection performance and validate the results in Fig. 2, Fig. 3 shows the detection probability \( P_d \) as a function of the available backhaul capacity \( C \) when the false alarm probability is \( P_{fa} = 0.01 \). The curve was evaluated via Monte Carlo simulations by implementing the optimum test detector [6]. We also implemented the distributed detection scheme described in Section III-A by setting the threshold \( \gamma_n \) in (9) to be equal for \( n \in \mathcal{N} \) for simplicity. It can be noted that the relative gains predicted by the Bhattacharyya distance criterion in Fig. 2 are consistent with the performance shown in Fig. 3. Moreover, for small values of \( C \), distributed detection outperforms cloud detection due to the performance degradation caused by the large quantization noise on the cloud-based schemes. However, as the available backhaul capacity \( C \) increases,
the cloud detection approach considerably outperforms distributed detection.

Fig. 4 plots the Receiving Operating Characteristic (ROC), i.e., the detection probability $P_d$ versus false alarm probability $P_{fa}$, for $C = 5$ bits/s/Hz. It is confirmed that the proposed joint optimization method provides remarkable gains over all separate optimization schemes as well as over the distributed detection approach. For instance, for $P_{fa} = 0.01$, joint optimization yields $P_d = 0.7125$, while code optimization only yields $P_d = 0.599$.

**B. Multiple-Access Wireless Backhaul**

In this section, we evaluate the performance of the proposed algorithms that perform the joint optimization of the code vector $\mathbf{x}$ and of the amplifying power gains $\mathbf{p}$ for the short-term (Section IV-B) and long-term (Section IV-C) adaptive designs. For reference, we consider the following schemes: (i) No opt.: Set $\mathbf{x} = \sqrt{P_T/K} \mathbf{1}_K$ and $\mathbf{p} = P_R/N \mathbf{1}_N$; (ii) Code opt.: Optimize the code vector $\mathbf{x}$ as per Algorithm 2 (with (30) in lieu of (27) for the long-term adaptive design) with $\mathbf{p} = P_R/N \mathbf{1}_N$; and (iii) Gain optimization (Gain opt.): Optimize the gains $\mathbf{p}$ as per Algorithm 2 (with (31) in lieu of (29) for the long-term adaptive design) with $\mathbf{x} = \sqrt{P_T/K} \mathbf{1}_K$. We set the total REs’ power as $P_R = 10$ dB.

Fig. 5 shows the Bhattacharyya distance as a function of the TE’s power $P_T$, with $N = 3$, $\sigma_{c,1}^2 = 0.125$, $\sigma_{c,2}^2 = 0.25$ and $\sigma_{c,3}^2 = 0.5$. For smaller values of $P_T$, optimizing the code vector is more advantageous than optimizing the amplifying gains, due to the fact that performance is limited by the TE-to-REs connection. In contrast, for sufficiently large values of $P_T$, the optimization of the REs’ gains is to be preferred, since the performance becomes limited by the channels between the REs and the FC. Joint optimization significantly outperforms all other schemes, except in the very low- and large-power regimes, in which, as discussed, the performance is limited by either the TE-to-REs or the REs-to-FC channels. In addition, we observe that the long-term adaptive scheme loses about 10% in terms of the Bhattacharyya distance with respect to the short-term adaptive design in the high SNR regime. The results in Fig. 5 can be interpreted by noting that the joint optimization seeks to design the transmitted signal $\mathbf{x}$ such that it reduces the power transmitted at the frequencies in which the REs observe the largest interference, while, at the same time,
allocating more power to REs suffering from less interference and, with the short-term adaptive design, having better channels to the FC. Fig. 5 also reveals an interesting property of cloud radar: properties of the communication channel affect the design of the radar waveform. This is evidenced in the different Bhattacharyya distances resulting from code only optimization for instantaneous CSI versus stochastic CSI.

In Fig. 6 the Bhattacharyya distance is plotted versus the number REs $N$ with $P_T = 12$ dB, $\sigma^2_{c,1} = 1$, $\sigma^2_{c,2} = 0.9$, $\sigma^2_{c,3} = 0.75$, $\sigma^2_{c,4} = 0.5$, $\sigma^2_{c,5} = 0.35$, $\sigma^2_{c,6} = 0.25$, $\sigma^2_{c,7} = 0.125$ and $\sigma^2_{c,8} = 0.05$. Optimizing the REs’ power gains is seen to be especially beneficial at large $N$, due to the ability to allocate more power to the REs in better condition in terms of interference and channels to the FC. For instance, even with the long-term adaptive design, optimizing the REs’ power gains outperforms code optimization with short-term adaptive design for sufficiently large $N$.

Fig. 7 plots the ROC curves with $P_T = 20$ dB, $N = 3$, $\sigma^2_{c,1} = 0.125$, $\sigma^2_{c,2} = 0.25$ and $\sigma^2_{c,3} = 0.5$. The curve was evaluated via Monte Carlo simulations by implementing the optimum test detector (6). It can be observed that the gains observed in the previous figures directly translate into a better ROC performance of joint optimization. Note also that power gain optimization is seen to be advantageous due to large value of $P_T$ as predicted based on Fig. 5.

VI. CONCLUDING REMARKS

We have studied a multistatic cloud radar system, where the REs and FC are connected via an orthogonal-access wired backhaul or a multiple-access wireless backhaul channel. In the former case, each RE quantizes and forwards the signal sent by TE to a FC, while amplify-and-forward of the received signal is carried out over the multiple-access wireless backhaul. The FC collects the signals from all the REs and determines the target’s presence or absence. We have investigated the joint optimization of waveform and backhaul transmission so as to maximize the detection performance. As the performance metric, we adopted the Bhattacharyya distance and the proposed algorithmic solutions were based on successive convex approximations. Overall, joint optimization was seen to have remarkable gains over
the standard separate optimization of waveform and backhaul transmission. Moreover, cloud processing is found to outperform the standard distributed detection approach as long as the backhaul capacity is large enough.

APPENDIX A

DISCUSSION ON MODEL (4) USING RATE-DISTORTION THEORY

In this Appendix, we elaborate on the justification of the additive quantization noise model (4) and the related adoption of the mutual information as a performance criterion using rate-distortion theory. We refer to [25] for an introduction on rate-distortion theory and particularly on the notion of joint typicality. Here, we keep the discussion informal in order to emphasize the main ideas.

To start, consider a vector of \( m \) samples of the received signal (1), which we denote as \( r^m = [r_1, \ldots, r_m] \) with the subscript indexing the samples. Similarly, denote as \( \tilde{r}^m = [\tilde{r}_1, \ldots, \tilde{r}_m] \) the corresponding quantized vector of \( m \) samples (4). From the standard covering lemma [25] of rate-distortion theory, it follows that there exist vector quantization schemes with rate close to the mutual information \( I(r; \tilde{r}) \) for which we have, for large \( m \),

\[
\Pr[(r^m, \tilde{r}^m) \text{ are strongly jointly typical}] \simeq 1. 
\]  
(32)

The probability (32) is taken with respect to the distribution of the sequence \( r^m \), and hence the condition (32) says that, for almost every \( r^m \), a jointly typical sequence \( \tilde{r}^m \) can be found by the quantizer. Joint typicality is with respect to the (properly quantized) distribution \( p_j(r^m, \tilde{r}^m) \) under hypothesis \( \mathcal{H}_j \) as dictated by (4). As a reminder, joint typicality implies that the empirical distribution, or histogram, \( \pi(r, \tilde{r}|r^m, \tilde{r}^m) \) of the vector pair \( (r^m, \tilde{r}^m) \), namely

\[
\pi(r, \tilde{r}|r^m, \tilde{r}^m) = \sum_{i=1}^{m} 1(r_i = r, \tilde{r}_i = \tilde{r}), \quad (33)
\]

where \( 1(\cdot) \) is the indicator function, satisfies

\[
\pi(r, \tilde{r}|r^m, \tilde{r}^m) \simeq p_j(r, \tilde{r}) \quad (34)
\]
with high probability. We refer to [25] for more precise definitions and to [26] for related results.

From the discussion above, for most sources $r^m$, the empirical marginal distribution of the quantized signal $\tilde{r}^m$ is approximately Gaussian with distribution given by (4). Therefore, the Bhattacharyya distance (11) can be interpreted as the distance between the empirical distributions of $\tilde{r}^m$ under two hypotheses.

For an alternative related justification, define as $q_{j,i}(\tilde{r}^m)$ the distribution of the $i$th quantized symbol $\tilde{r}_i$ under hypothesis $\mathcal{H}_j$. Consider now the Bhattacharyya distance between $q_{0,i}(\tilde{r}^m)$ and $q_{1,i}(\tilde{r}^m)$ averaged over the block, namely $1/m \sum^{m}_{i=1} B(q_{0,i}, q_{1,i})$. We have the inequality

$$\frac{1}{m} \sum^{m}_{i=1} B(q_{0,i}, q_{1,i}) \geq B\left(\frac{1}{m} \sum^{m}_{i=1} q_{0,i}, \frac{1}{m} \sum^{m}_{i=1} q_{1,i}\right)$$

$$m \to \infty \ B(p_0(\tilde{r}), p_1(\tilde{r})), \quad (35a)$$

where (35a) follows from Jensen’s inequality, due to the convexity of the Bhattacharyya distance, and (35b) holds in the limit of large $m$ with high probability thanks to the covering lemma discussed above. The bound (35) shows that the average Bhattacharyya distance between the distributions of the quantized signals under the two assumptions over the block is lower-bounded by the Bhattacharyya distance (11), which equals $B(p_0(\tilde{r}), p_1(\tilde{r}))$.

**References**

[1] M. Bernfeld, *Radar signals: An introduction to theory and application*. Elsevier, 2012.

[2] A. W. Rihaczek, *Principles of high-resolution radar*. Wiley, 1967.

[3] S. M. Kay, “Optimal signal design for detection of Gaussian point targets in stationary Gaussian clutter/reverberation,” *IEEE Jour. Select. Topics in Sig. Proc.*, vol. 1, no. 1, pp. 31–41, Jun. 2007.

[4] T. Kailath, “The divergence and Bhattacharyya distance measures in signal selection,” *IEEE Trans. Comm.*, vol. 15, no. 1, pp. 52–60, Feb. 1967.

[5] S. M. Kay, “Waveform design for multistatic radar detection,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 3, pp. 1153–1166, Jul. 2009.

[6] M. Naghsh, M. Modarres-Hashemi, S. Shahbazpanahi, M. Soltanalian, and P. Stoica, “Unified optimization framework of multi-static radar code design using information-theoretic criteria,” *IEEE Trans. Sig. Proc.*, vol. 61, no. 21, pp. 5401–5416, Nov. 2013.

[7] China Mobile, “C-RAN: the road towards green RAN,” White Paper, ver. 2.5, China mobile Research Institute, Oct. 2011.
[8] G. Alirezaei, M. Reyer, and R. Mathar, “Optimum power allocation in sensor networks for passive radar applications,” IEEE Trans. Wireless Comm., vol. 13, no. 6, pp. 3222–3231, Jun. 2014.

[9] G. Alirezaei, O. Taghizadeh, and R. Mathar, “Optimum power allocation with sensitivity analysis for passive radar applications,” IEEE Sensors Journal, vol. 14, no. 11, pp. 3800–3809, Jun. 2014.

[10] J. Kang, O. Simeone, J. Kang, and S. S. Shitz, “Joint signal and channel state information compression for the backhaul of uplink network MIMO systems,” IEEE Trans. Wireless Comm., vol. 13, no. 3, pp. 1555–1567, Mar. 2014.

[11] ——, “Fronthaul compression and precoding design for C-RANs over ergodic fading channel,” arXiv preprint [arXiv:1412.7713].

[12] S. Jeong, O. Simeone, A. Haimovich, and J. Kang, “Optimal fronthaul quantization for cloud radio positioning,” arXiv preprint [arXiv:1311.1567].

[13] S. Khalili, O. Simeone, and A. M. Haimovich, “Cloud Radio-Multistatic Radar: Joint optimization of code vector and backhaul quantization,” IEEE Sig. Proc. Lett., vol. 22, no. 4, pp. 494–498, Oct. 2014.

[14] P. K. Varshney, Distributed Detection and Data Fusion. Springer, 1997.

[15] I. F. Akyildiz, B. F. Lo, and R. Balakrishnan, “Cooperative spectrum sensing in cognitive radio networks: A survey,” Physical Communication, vol. 4, no. 1, pp. 40–62.

[16] D. R. Hunter and K. Lange, “A tutorial on MM algorithms,” The American Statistician, vol. 58, no. 1, pp. 30–37, Feb. 2004.

[17] M. Razaviyayn, M. Sanjabi, and Z.-Q. Luo, “A stochastic successive minimization method for nonsmooth nonconvex optimization with applications to transceiver design in wireless communication networks,” arXiv preprint arXiv:1307.4457, Jul. 2013.

[18] A. Gersho and R. Grey, Vector Quantization and Signal Compression. Boston: MA Kluwer, 1992.

[19] T. M. Cover and J. A. Thomas, Element of Information Theory. John Wiley & Sons, 2006.

[20] R. Zamir and M. Feder, “On lattice quantization noise,” IEEE Trans. Info. Th., vol. 42, no. 4, pp. 1152–1159, Jun. 1996.

[21] M. W. Marcellin and T. R. Fischer, “Trellis coded quantization of memoryless and Gauss-Markov sources,” IEEE Trans. Comm., vol. 38, no. 1, pp. 82–93, Jan. 1996.

[22] S. M. Kay, Fundamentals of Signal Processing-Estimation Theory. Englandwood Cliffs, NJ: Prentice Hall, 1993.

[23] S. P. Boyd and L. Vandenberghe, Convex optimization. Cambridge university press, 2004.

[24] M. Razaviyayn, M. Hong, and Z. Luo, “A unified convergence analysis of block successive minimization methods for nonsmooth optimization,” SIAM Journal on Optimization, vol. 23, no. 2, pp. 1126–1153, Jun. 2013.

[25] A. E. Gamal and Y.-H. Kim, Network Information Theory. Cambridge University Press, 2011.

[26] T. Weissman and E. Ordentlich, “The empirical distribution of rate-constrained source codes,” IEEE Trans. Info. Th., vol. 51, no. 11, pp. 3718–3733, Nov. 2005.
Algorithm 1 Joint optimization of code vector and quantization noise covariances

Initialization (outer loop): Initialize $x^{(0)} \in \mathbb{C}^{K \times 1}$, $\Omega_q^{(0)} \succeq 0$ and set $i = 0$.

Repeat (BCD method)

\[ i \leftarrow i + 1 \]

Initialization (inner loop): Initialize $x^{(i,0)} = x^{(i-1)}$ and set $j = 0$.

Repeat (MM method for $x^{(i)}$)

\[ j \leftarrow j + 1 \]

Find $x^{(i,j)}$ by solving the problem (19) with $\Omega_q = \Omega_q^{(i-1)}$ (see (19)).

Until a convergence criterion is satisfied.

Update $x^{(i)} \leftarrow x^{(i,j)}$

Initialization (inner loop): Initialize $\Omega_q^{(i,0)} = \Omega_q^{(i-1)}$ and set $j = 0$.

Repeat (MM method for $\Omega_q^{(i)}$)

\[ j \leftarrow j + 1 \]

Find $\Omega_q^{(i,j)}$ by solving the problem (22) with $x = x^{(i)}$ (see (22)).

Until a convergence criterion is satisfied.

Update $\Omega_q^{(i)} \leftarrow \Omega_q^{(i,j)}$

Until a convergence criterion is satisfied.

Solution: $x \leftarrow x^{(i)}$ and $\Omega_q \leftarrow \Omega_q^{(i)}$
Algorithm 2 Short-term adaptive design of code vector and amplifier gain (24)

Initialization (outer loop): Initialize \( x^{(0)} \in C^{K \times 1} \), \( p^{(0)} \geq 0 \) and set \( i = 0 \).

Repeat (BCD method)

\[ i \leftarrow i + 1 \]

Initialization (inner loop): Initialize \( x^{(i,0)} = x^{(i-1)} \) and set \( j = 0 \).

Repeat (MM method for \( x^{(i)} \))

\[ j \leftarrow j + 1 \]

Find \( x^{(i,j)} \) by solving the problem (27) with \( p = p^{(i-1)} \) (see (27)).

Until a convergence criterion is satisfied.

Update \( x^{(i)} \leftarrow x^{(i,j)} \)

Initialization (inner loop): Initialize \( p^{(i,0)} = p^{(i-1)} \) and set \( j = 0 \).

Repeat (MM method for \( p^{(i)} \))

\[ j \leftarrow j + 1 \]

Find \( p^{(i,j)} \) by solving the problem (29) with \( x = x^{(i)} \) (see (29)).

Until a convergence criterion is satisfied.

Update \( p^{(i)} \leftarrow p^{(i,j)} \)

Until a convergence criterion is satisfied.

Solution: \( x \leftarrow x^{(i)} \) and \( p \leftarrow p^{(i)} \)
Fig. 1. Illustration of the considered multistatic cloud radar system, which consists of a TE, $N$ REs, and a FC. All the nodes are configured with a single antenna. The REs are connected to the FC via (a) orthogonal-access wired backhaul links or (b) a (non-orthogonal) multiple-access wireless backhaul channel.
Fig. 2. Bhattacharyya distance versus the backhaul capacity $C_n = C$, $n \in \mathcal{N}$, with $P_T = 10$ dB, $K = 6$, $N = 3$, $\sigma_{t,n}^2 = 1$, $\sigma_{c,1}^2 = 0.125$, $\sigma_{c,2}^2 = 0.25$, $\sigma_{c,3}^2 = 0.5$ and $[\Omega_{w,n}]_{i,j} = (1 - 0.12n)^{|i-j|}$ for $n \in \mathcal{N}$.

Fig. 3. Probability of detection $P_d$ versus the backhaul capacity $C_n = C$, $n \in \mathcal{N}$, with $P_T = 10$ dB, $K = 6$, $N = 3$, $\sigma_{t,n}^2 = 1$, $\sigma_{c,1}^2 = 0.125$, $\sigma_{c,2}^2 = 0.25$, $\sigma_{c,3}^2 = 0.5$, $[\Omega_{w,n}]_{i,j} = (1 - 0.12n)^{|i-j|}$ for $n \in \mathcal{N}$ and $P_{fa} = 0.01$. 
Fig. 4. ROC curves with $P_T = 10$ dB, $K = 6$, $N = 3$, $\sigma_{f,n}^2 = 1$, $\sigma_{c,1}^2 = 0.125$, $\sigma_{c,2}^2 = 0.25$, $\sigma_{c,3}^2 = 0.5$, $[\Omega_{w,n}]_{i,j} = (1 - 0.12 n)^{|i-j|}$ and $C_n = C = 5$ bits/s/Hz for $n \in \mathcal{N}$.

Fig. 5. Bhattacharyya distance versus the TE’s power $P_T$ with $P_R = 10$ dB, $K = 6$, $N = 3$, $\sigma_{f,n}^2 = 1$, $\sigma_{t,n}^2 = 1$, $\sigma_{c,1}^2 = 0.125$, $\sigma_{c,2}^2 = 0.25$, $\sigma_{c,3}^2 = 0.5$, $[\Omega_{w,n}]_{i,j} = (1 - 0.12 n)^{|i-j|}$ and $[\Omega_z]_{i,j} = (1 - 0.45)^{|i-j|}$ for $n \in \mathcal{N}$. 
Fig. 6. Bhattacharyya distance versus the number REs $N$ with $P_T = 12$ dB, $P_R = 10$ dB, $K = 6$, $\sigma_{f,n}^2 = 1$, $\sigma_{t,n}^2 = 1$, $\sigma_{c,1}^2 = 1$, $\sigma_{c,2}^2 = 0.9$, $\sigma_{c,3}^2 = 0.75$, $\sigma_{c,4}^2 = 0.5$, $\sigma_{c,5}^2 = 0.35$, $\sigma_{c,6}^2 = 0.25$ and $\sigma_{c,7}^2 = 0.125$, $\sigma_{c,8}^2 = 0.05$, $[\Omega_{w,n}]_{i,j} = (1 - 0.12n)^{|i-j|}$ and $[\Omega_{z}]_{i,j} = (1 - 0.45)^{|i-j|}$ for $n \in \mathcal{N}$.

Fig. 7. ROC curves with $P_T = 20$ dB, $P_R = 10$ dB, $K = 6$, $N = 3$, $\sigma_{f,n}^2 = 1$, $\sigma_{t,n}^2 = 1$, $\sigma_{c,1}^2 = 0.125$, $\sigma_{c,2}^2 = 0.25$, $\sigma_{c,3}^2 = 0.5$, $[\Omega_{w,n}]_{i,j} = (1 - 0.12n)^{|i-j|}$ and $[\Omega_{z}]_{i,j} = (1 - 0.45)^{|i-j|}$ for $n \in \mathcal{N}$. 