# Cooperative Relay with Interference Alignment for Video over Cognitive Radio Networks

Donglin Hu and Shiwen Mao
Dept. Electrical and Computer Engineering, Auburn University, Auburn, AL 36849-5201, USA
Email: dzh0003@tigermail.auburn.edu, smao@ieee.org

***Abstract*—Due to the drastic increase in wireless video traffic, the capacity of existing and future wireless networks will be greatly stressed, while interference will become the dominant capacity limiting factor. In this paper, we investigate cooperative relay in CR networks using video as a reference application. We incorporate interference alignment to allow transmitters collaboratively send encoded signals to all CR users, such that undesired signals will be canceled and the desired signal can be decoded at each CR user. We present a stochastic programming formulation, as well as a reformulation that greatly reduces computational complexity. In the cases of a single licensed channel and multiple licensed channels with channel bonding, we develop an optimal distributed algorithm with proven convergence and convergence speed. In the case of multiple channels without channel bonding, we develop a greedy algorithm with a proven performance bound. The algorithms are evaluated with simulations and are shown to achieve considerable gains over two heuristic schemes that do not consider interference alignment.**

## I. Introduction

Due to significant advances in wireless access technologies and the proliferation of wireless devices and applications, there is a fundamental change in wireless network traffic. As predicted by a Cisco study, wireless data is expected to grow to 6.3 Exabytes per month by 2015, a 26-fold increase over 2010, and 66% of the increase in future wireless data traffic will be video related [1]. The capacity of existing and future wireless networks will be greatly stressed. Coupled with the quickly depleting spectrum, interference will become the major capacity limiting factor.

*Cognitive radios* (CR) provide an effective solution to meeting this critical demand, by exploiting co-deployed networks and sharing underutilized spectrum for future wireless networks [2]. On the other hand, *cooperative communications* represents another effective solution to the capacity problem, where wireless nodes help each other in information delivery to achieve the so-called *cooperative diversity* [3], [4]. Recently, researchers have been exploring the idea of combining these two advanced wireless communication techniques [5]–[7], and the potential of cooperative CR networks has been demonstrated with a testbed implementation [5].

In this paper, we investigate cooperative relay in CR networks, using video as a reference application to make the best use of the enhanced network capacity. We consider a base station (BS) and multiple relay nodes (RN) that collaboratively stream multiple videos to CR users within the network. It has been shown that the performance of a cooperative relay link is mainly limited by two factors: (i) the *half-duplex operation*, since the BS–RN and the RN–user transmissions cannot be scheduled simultaneously on the same channel [3]; and (ii) the *bottleneck channel*, which is usually the BS–user and/or the RN–user channel, which usually has poor quality due to obstacles, attenuation, multipath propagation and mobility [4]. To support high quality video service in such a challenging environment, we assume a well planned relay network where the RNs are connected to the BS with high-speed wireline links. Therefore the video packets will be available at both the BS and the RNs before their scheduled transmission time, thus allowing advanced cooperative transmission techniques to be adopted for streaming videos.

We focus on the bottleneck channel problem. In particular, we consider interference alignment, where the BS and RNs simultaneously transmit encoded signals to all CR users, such that undesired signals will be canceled and the desired signal can be decoded at each CR user [8], [9]. In [10], such cooperative sender-side techniques are termed *interference alignment*, while receiver-side techniques that use overheard (or exchanged via a wireline link) packets to cancel interference is termed *interference cancelllation*. Interference alignment is underpinned by recent advances in information theory [8], [9], and the practical implications have been addressed and proof-of-concept implementations have been reported [8]–[12].

In this paper, we present a stochastic programming formulation of the problem of interference alignment for video streaming in cooperative CR networks. The cross-layer optimization formulation takes into account important design factors including spectrum sensing, opportunistic spectrum access, cooperative relay, interference alignment, and video QoS requirements. The objective is to maximize the received video quality at CR users. We then present a reformulation of the problem based on Linear Algebra theory [13], such that the number of variables and computational complexity can be greatly reduced.

We develop effective solution algorithms to the formulated problem. In particular, we consider three scenarios. In the case of a single licensed channel, we develop a distributed algorithm based on dual decomposition [14], which is guaranteed to converge to the global optimal with a bounded convergence speed. In the case of multiple licensed channels with channel bonding (where a transmitter can aggregate all the available channels to transmit the encoded signal [15], [16]), we show that the distributed algorithm can still be used to achieve optimal solutions. Finally, in the case of multiple licensed channels without channel bonding, we develop a greedy algorithm that leverages the single-channel algorithm for near-optimal solutions, which has a bounded performance. The proposed algorithms are evaluated with simulations, and are shown to outperform two heuristic schemes that do not incorporate interference alignment with considerable gains.

The remainder of this paper is organized as follows. We present the system model and preliminaries in Section II and the problem statement in Section III. The proposed solution algorithms are presented in Section IV. We discuss simulation results in Section V and related work in Section VI. Section VII concludes this paper.

## II. System Model and Preliminaries

In this section, we present the system model, assumptions, and preliminaries that provide the basis for the problem formulation in Section III. Some of the preliminaries are derived in our prior work [17], [18]. We do not claim contribution in this part but include it for the sake of completeness.

### A. Spectrum and Network Model

We consider a spectrum consisting of one common control channel (indexed 0) and $M$ licensed channels (indexed from 1 to $M$). The $M$ licensed channels are allocated to a primary network, and the common control channel is exclusively used by a cooperative CR network co-located with the primary network. As in prior work, we assume a synchronized time slot structure for the licensed channels [2]. The states of the $M$ licensed channels evolve over time independently, while the occupancy of each channel follows a two-state discrete time Markov process [2]. Let $\vec{S}(t) = [S_1(t), S_2(t), \cdots, S_M(t)]$ denote the network state in time slot $t$, where element $S_m(t)$ represents the status of channel $m$ as: $S_m(t) = 0$ for an idle channel and $S_m(t) = 1$ for a busy channel. The utilization of channel $m$, i.e., $\eta_m = \Pr\{S_m = 1\}$, can be derived as:

$$\eta_m = \lim_{t_0 \to \infty} \frac{1}{t_0} \sum_{t=1}^{t_0} S_m(t), \quad m = 1, 2, \cdots, M. \tag{1}$$

The cooperative CR network is illustrated in Fig. 1. There is a CR BS (indexed 1) and $(K-1)$ CR RNs (indexed from 2 to $K$) deployed in the area to serve $N$ active CR users. Let $\mathcal{U} = \{1, 2, \cdots, N\}$ denote the set of active CR users. We assume that the BS and all the RNs are equipped with multiple transceivers: one is tuned to the common control channel and the others are used to sense multiple licensed channels at the beginning of each time slot, and to transmit encoded signals to CR users. We consider the case where each CR user has one software defined radio (SDR) based transceiver, which can be tuned to operate on any of the $(M+1)$ channels. If the channel bonding/aggregation techniques are used [15], [16], a transmitter can collectively use all the available channels and a CR user can receive from all the available channels simultaneously. Otherwise, only one licensed channel will be used by a transmitter and a CR user can only receive from a single chosen channel at a time.

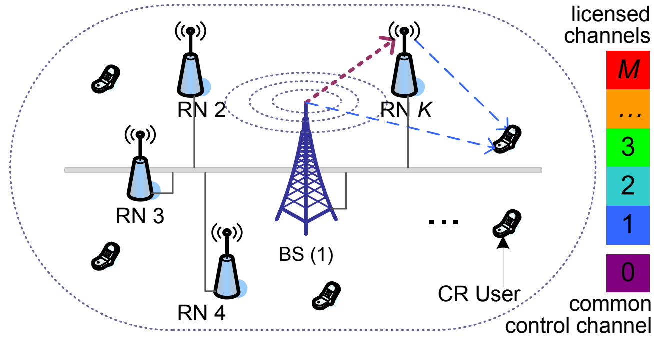


Fig. 1. Illustration of the cooperative CR network.

Consider the three channels in a traditional cooperative relay link. Usually the BS and RNs are mounted on high towers, and the BS–RN channel has good quality due to line-of-sight (LOS) transmissions and absence of mobility. On the other hand, a CR user is typically on the ground level (or indoor). The BS–user and RN–user channels usually have much poorer quality due to obstacles, attenuation, multipath propagation and mobility. To support high quality video service, we assume a well planned relay network, where the RNs are connected to the BS via broadband wireline connections (e.g., as in femtocell networks [18]). Alternatively, free space optical links can also be used to provide multi-gigabit rates between the BS and the RNs [19]. As a result, the video packets will always be available for transmission (with suitable channel coding and retransmission) at the RNs at their scheduled transmission time. To cope with the much poorer BS–user and RN–user channels, the BS and RNs adopt interference alignment to cooperatively transmit video packets to CR users, while exploiting the spectrum opportunities in the licensed channels (see Section III).

### B. Spectrum Sensing

Each time slot consists of a spectrum sensing phase and a data transmission phase. The BS and the RNs sense the licensed channels and exchange their sensing results over the common control channel during the sensing phase. We adopt a hypothesis test to detect channel availability. For channel $m$, the null hypothesis and the alternative hypothesis are:

$$\begin{cases} \mathbb{H}_0^m : \text{channel } m \text{ is idle} \\ \mathbb{H}_1^m : \text{channel } m \text{ is busy.} \end{cases}$$

We model both types of detection errors: (i) false alarm, when an idle channel is considered busy and a spectrum opportunity will be wasted; (ii) miss detection, when a busy channel is considered idle, which may cause collision with primary users.

Let $\Theta_l^m$ be the $l$th sensing result obtained on channel $m$, with binary (0 or 1) values. The false alarm and miss detection probabilities associated with $\Theta_l^m$, denoted by $\epsilon_l^m$ and $\delta_l^m$, are:

$$\epsilon_l^m = \Pr(\Theta_l^m = 1 | \mathbb{H}_0^m) \tag{2}$$
$$\delta_l^m = \Pr(\Theta_l^m = 0 | \mathbb{H}_1^m). \tag{3}$$

Given $L$ sensing results obtained for channel $m$, the corresponding sensing result vector is $\vec{\Theta}_L^m = [\Theta_1^m, \Theta_2^m, \cdots, \Theta_L^m]$.

Let $P_m^A(\vec{\Theta}_L^m) := P_m^A(\Theta_1^m, \Theta_2^m, \cdots, \Theta_L^m)$ be the conditional probability that channel $m$ is available, which can be computed iteratively as shown in our prior work [18]:

$$
\begin{aligned}
P_m^A(\Theta_1^m) &= \left[1 + \frac{\eta_m}{1-\eta_m} \times \frac{(\delta_1^m)^{1-\Theta_1^m}(1-\delta_1^m)^{\Theta_1^m}}{(\epsilon_1^m)^{\Theta_1^m}(1-\epsilon_1^m)^{1-\Theta_1^m}}\right]^{-1} \\
P_m^A(\vec{\Theta}_l^m) &:= P_m^A(\Theta_1^m, \Theta_2^m, \cdots, \Theta_l^m) \\
&= \left\{1 + \left[\frac{1}{P_m^A(\Theta_1^m, \Theta_2^m, \cdots, \Theta_{l-1}^m)} - 1\right] \times \right. \\
&\quad \left. \frac{(\delta_l^m)^{1-\Theta_l^m}(1-\delta_l^m)^{\Theta_l^m}}{(\epsilon_l^m)^{\Theta_l^m}(1-\epsilon_l^m)^{1-\Theta_l^m}}\right\}^{-1}, l = 2, \cdots, L.
\end{aligned}
$$

### C. Opportunistic Spectrum Access

For each channel $m$, define an index variable $D_m(t)$ for the BS or RNs to access the channel in time slot $t$. That is,

$$
D_m(t) = \begin{cases} 0, & \text{access channel } m \text{ in time slot } t \\ 1, & \text{otherwise,} \end{cases} \qquad m = 1, 2, \cdots, M. \tag{4}
$$

With sensing result $P_m^A(\vec{\Theta}_L^m)$, each channel $m$ will be opportunistically accessed. Let the probability be $P_m^D(\vec{\Theta}_L^m)$ that channel $m$ will be accessed in time slot $t$ (i.e., when $D_m(t) = 0$). When channel $m$ is busy (i.e, with probability $1 - P_m^A(\vec{\Theta}_L^m)$), accessing the channel will cause collision to primary users. For primary user protection, such collision probability should be bounded by a prescribed threshold, denoted as $\gamma_m$. Such a *primary user protection constraint* can be written as:

$$
\left[1 - P_m^A(\vec{\Theta}_L^m)\right] P_m^D(\vec{\Theta}_L^m) \leq \gamma_m. \tag{5}
$$

We can solve (5) to obtain the optimal channel access probability, as:

$$
P_m^D(\vec{\Theta}_L^m) = \min\left\{\frac{\gamma_m}{1 - P_m^A(\vec{\Theta}_L^m)}, 1\right\}. \tag{6}
$$

Let $\mathcal{A}(t)$ be the set of available channels in time slot $t$. It follows that

$$
\mathcal{A}(t) := \{m \mid D_m(t) = 0\}. \tag{7}
$$

### D. Interference Alignment

We next briefly describe the main idea of interference alignment considered in this paper. Interested readers are referred to [10], [12] for insightful examples, a classification of various interference alignment scenarios, and practical considerations.

Consider two transmitters (denoted as $s_1$ and $s_2$ ) and two receivers (denoted as $d_1$ and $d_2$). Let $X_1$ and $X_2$ be the signals corresponding to the packets to be sent to $d_1$ and $d_2$, respectively, which are known at both $s_1$ and $s_2$. With interference alignment, the transmitters $s_1$ and $s_2$ send compound signals $a_{1,1}X_1 + a_{1,2}X_2$ and $a_{2,1}X_1 + a_{2,2}X_2$, respectively, to the two receivers $d_1$ and $d_2$ simultaneously. If channel noise is ignored, the received signals $Y_1$ and $Y_2$ can be written as:

$$
\begin{aligned}
\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} &= \begin{bmatrix} G_{1,1} & G_{1,2} \\ G_{2,1} & G_{2,2} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \\
&:= \mathbf{G}^{\mathsf{T}} \times \mathbf{A} \times \vec{X},
\end{aligned} \tag{8}
$$

where $G_{i,j}$ is the channel gain from transmitter $s_i$ to receiver $d_j$, for all $i$ and $j$.

From (8), it can be seen that both receivers can perfectly decode their signals if the precoding matrix $\mathbf{A}$ is chosen to be $\left\{\mathbf{G}^{\mathsf{T}}\right\}^{-1}$, i.e., the inverse of the channel gain matrix. With this technique, the transmitters are able to send packets simultaneously and the interference between the two concurrent transmissions can be effectively canceled at both the receivers [10].

### E. Video Performance Measure

We assume that the BS streams multiple real-time videos, one to each CR user, with help from the RNs. The MGS option of H.264/SVC is adopted in our model because (i) the scalability is very useful to achieve a graceful quality degradation under the highly dynamic CR network environment [17], [20], and (ii) MGS has better rate distortion performance over MPEG-4 FGS [21]. Due to the real-time constraint, we assume that each Group of Pictures (GOP) must be delivered in the next $T$ time slots. Video packets are transmitted in the decreasing order of their significance to the quality of reconstructed video. Retransmissions are made for lost video packets and overdue packets will be discarded.

The quality of reconstructed MGS video can be modeled with a linear equation [21]:

$$
W(R) = \alpha + \beta \times R, \tag{9}
$$

where $W(R)$ is the average peak signal-to-noise ratio (PSNR) of the reconstructed MGS video, $R$ is the average data rate, and $\alpha$ and $\beta$ are constants depending on the specific video sequence and codec. This relationship is validated in our simulations with several video test sequences and the H.264/SVC codec. We observe that the simulated points match very well with the model-predicted curves.

## III. Problem Formulation

We formulate the problem of interference alignment for scalable video streaming over cooperative CR networks in this section. As discussed in Section II, the video packets are available at the BS and RNs before their scheduled transmission time; the BS and RNs adopt interference alignment to overcome the poor BS–user and RN–user channels.

Let $X_j$ denote the signal to be transmitted to user $j$, which has unit power. As shown in Section II-D, transmitter $k$ sends a compound signal $\sum_{j \in \mathcal{U}} a_{k,j} X_j$ to all the active CR users, where $a_{k,j}$'s are the weights (or, transmit power) to be determined. Ignoring channel noise, we can compute the received signal $Y_n$ at a user $n$ as:

$$\begin{aligned} Y_n &= \sum_{k=1}^{K} G_{k,n} \sum_{j=1}^{N} a_{k,j} X_j = \sum_{k=1}^{K} \sum_{j=1}^{N} a_{k,j} G_{k,n} X_j \\ &= \sum_{j=1}^{N} X_j \sum_{k=1}^{K} a_{k,j} G_{k,n}, \ n = 1, 2, \cdots, N, \end{aligned} \tag{10}$$

where $G_{k,n}$ is the channel gain from the BS (i.e., $k = 1$) or an RN $k$ to user $n$. For user $n$, only signal $X_n$ should be decoded and the coefficients of all other signals should be forced to zero. The *zero-forcing constraints* can be written as:

$$\sum_{k=1}^{K} a_{k,j} G_{k,n} = 0, \quad \text{for all } j \neq n. \tag{11}$$

Usually the total transmit power of the BS and every RN is limited by a peak power $P_{max}$. Since $X_j$ has unit power, for all $j$, the power of each transmitted signal is the square sum of all the related coefficients $a_{k,j}^2$. The *peak power constraint* can be written as

$$\sum_{j=1}^{N} |a_{k,j}|^2 \leq P_{max}, \quad k = 1, \cdots, K. \tag{12}$$

Recall that each CR user has one SDR transceiver that can be tuned to receive from any of the $(M+1)$ channels (without channel bonding). Let $b_j^m$ be a binary variable indicating that user $j$ selects licensed channel $m$. It is defined as

$$b_j^m = \begin{cases} 1, & \text{if user } n \text{ receives from channel } m \\ 0, & \text{otherwise,} \end{cases} \qquad j = 1, \cdots, N, \ m = 1, \cdots, M. \tag{13}$$

Then, we have the following *transceiver constraint*:

$$\sum_{m \in \mathcal{A}(t)} b_j^m \leq 1, \quad j = 1, \cdots, N. \tag{14}$$

After introducing the channel selection variables $b_j^m$'s, the overall channel gain becomes

$$G_{k,j} = \sum_{m \in \mathcal{A}(t)} b_j^m H_{k,j}^m, \tag{15}$$

where $H_{k,j}^m$ is the channel gain from the BS (i.e., $k = 1$) or an RN $k$ to user $j$ on channel $m$.

Let $w_j^t$ be the PSNR of user $j$'s reconstructed video at the beginning of time slot $t$ and $W_j^t$ the PSNR of user $j$'s reconstructed video at the end of time slot $t$. In time slot $t$, $w_j^t$ is already known, while $W_j^t$ is a random variable depending on the resource allocation and primary user activity during the time slot. That is, $w_j^{t+1}$ is a realization of $W_j^t$.

We formulate a multistage stochastic programming problem to maximize the expected logarithm-sum of the PSNR's at the end of the GOP, i.e., $\sum_{j=1}^{N} \mathbb{E}\left[\log(W_j^T)\right]$, for proportional fairness among the video sessions. The multistage stochastic programming problem can be decomposed into $T$ serial sub-problems, one for each time slot $t$, as [17]:

$$\begin{aligned} \text{maximize:} \quad & \sum_{j=1}^{N} \mathbb{E}\left[\log(W_j^t)|w_j^t\right] & (16) \\ \text{subject to:} \quad & W_j^t = w_j^t + \Psi_j^t & (17) \\ & b_j^m \in \{0,1\}, \ a_{k,j} \geq 0, \ \text{for all } m, j, k & (18) \\ & \text{Constraints (11), (12) and (14),} \end{aligned}$$

where $\Psi_j^t$ is a random variable that depends on spectrum sensing, power allocation, and channel selection. This is a mixed integer nonlinear programming problem (MINLP), with binary variables $b_j^m$'s and continuous real variables $a_{k,j}$'s.

In particular, $\Psi_j^t$ can have two possible values: (i) zero, if the packet is not successfully received due to collision with primary users; (ii) the objective value increase achieved in time slot $t$ if the packet is successfully received, denoted as $\lambda_j^t$. The increase in objective value can be computed as:

$$\lambda_j^t = \frac{\beta_j B}{T} \log_2 \left( 1 + \frac{1}{N_0} \left( \sum_{k=1}^{K} a_{k,j} G_{k,j} \right)^2 \right), \tag{19}$$

where $N_0$ is the noise power and $B$ is the channel bandwidth.

User $j$ can successfully receive a video packet from channel $m$ if it tunes to channel $m$ (i.e., $b_j^m = 1$) and the BS and RNs transmit on channel $m$ (i.e., with probability $P_m^D(\vec{\Theta}_L^m)$). The probability that user $j$ successfully receives a video packet, denoted as $P_j^t$, is

$$P_j^t = \sum_{m \in \mathcal{A}(t)} b_j^m P_m^D(\vec{\Theta}_L^m). \tag{20}$$

Therefore, we can expand the expectation in (16) to obtain a reformulated problem:

$$\begin{aligned} \text{maximize:} \quad & \sum_{j=1}^{N} \mathbb{E}\left[P_j^t \log(w_j^t + \lambda_j^t) + (1 - P_j^t) \log(w_j^t)\right] & (21) \\ \text{subject to:} \quad & \text{constraints (11), (12), (14), and (18).} \end{aligned}$$

## IV. SOLUTION ALGORITHMS

In this section, we develop effective solution algorithms to the formulated problem (16). In Section IV-A, we first consider the case of a single licensed channel, and derive a distributed, optimal algorithm with guaranteed convergence and bounded convergence speed. We then address the case of multiple licensed channels. If channel bonding/aggregation techniques are used [15], [16], the distributed algorithm in Section IV-A can still be applied to achieve optimal solutions. We finally consider the case of multiple licensed channels without channel bonding, and develop a greedy algorithm with a performance lower bound in Section IV-C.

### A. Case of Single Channel

*1) Property :* Consider the case when there is only one licensed channel, i.e., when $M = 1$. The $K$ transmitters, including the BS and $(K-1)$ RNs, send video packets to active users using the licensed channel when it is sensed idle.

*Definition 1:* A set of vectors is *linearly independent* if none of them can be written as a linear combination of the other vectors in the set [13].

For user $j$, the weight and channel gain vectors are: $\vec{a}_j = [a_{1,j}, a_{2,j}, \cdots, a_{K,j}]^{\mathsf{T}}$ and $\vec{G}_j = [G_{1,j}, G_{2,j}, \cdots, G_{K,j}]^{\mathsf{T}}$, where $\mathsf{T}$ denotes *matrix transpose*. Due to spatial diversity, we assume that the $\vec{G}_j$ vectors are linearly independent [8].

*Lemma 1:* To successfully decode each signal $X_j$, $j = 1, 2, \cdots, N$, the number of active users $N$ should be smaller than or equal to the number of transmitters $K$.

*Proof:* From (11), it can be seen that $\vec{a}_j$ is orthogonal to the $(N-1)$ vectors $\vec{G}_n$'s, for $n \neq j$. Since $\vec{a}_j$ is a $K$ by 1 vector, there are at most $(K-1)$ vectors that are orthogonal to $\vec{a}_j$. Since the $\vec{G}_j$ vectors are linearly independent, it follows that $(N-1) \leq (K-1)$ and therefore $N \leq K$. ■

According to Lemma 1, the following additional constraints should be enforced for the channel selection variables.

$$\sum_{j=1}^{N} b_j^m \leq K, \quad \text{for all } m \in \mathcal{A}(t). \tag{22}$$

That is, the number of active users receiving from any channel $m$ cannot be more than the number of transmitters on that channel, which is $K$ in the single channel case and less than or equal to $K$ in the multiple channels case. We first assume that $N$ is not greater than $K$, and will remove this assumption in the following subsection.

*2) Reformulation and Complexity Reduction :* With a single channel, all active users receive from channel 1. Therefore $b_j^1 = 1$, and $b_j^m = 0$, for $m > 1$, $j = 1, 2, \cdots, N$. The formulated problem is now reduced to a nonlinear programming problem with constraints (11), (12), and (18).

If the number of active users is $N = 1$, the solution is straightforward: all the transmitters send the same signal $X_1$ to the single user using their maximum transmit power $P_{max}$.

In general, the reduced problem can be solved with the dual decomposition technique [14] (i.e., a primal dual algorithm). This problem has $K \times N$ primal variables (i.e., the $a_{k,j}$'s), and we need to define $N(N-1)$ dual variables (or, Lagrangian Multipliers) for constraints (11) and $K$ dual variables for constraints (12). These numbers could be large for even moderate-sized systems. Before presenting the solution algorithm, we first derive a reformulation of the original problem (21) that can greatly reduce the number of primal and dual variables, such that the computational complexity can be reduced.

*Lemma 2:* Each vector $\vec{a}_j = [a_{1,j}, a_{2,j}, \cdots, a_{K,j}]^{\mathsf{T}}$ can be represented by the linear combination of $r$ nonzero, linearly independent vectors, where $r = K - N + 1$.

*Proof:* From (11), each vector $\vec{a}_j$ is orthogonal to $\vec{G}_i$ where $j \neq i$. Define a reduced matrix $\mathbf{G}_{-j}$ obtained by deleting $\vec{G}_j$ from $\mathbf{G}$, i.e., $\mathbf{G}_{-j} = [\vec{G}_1, \cdots, \vec{G}_{j-1}, \vec{G}_{j+1}, \cdots, \vec{G}_N]$. Then $\vec{a}_j$ is a solution to the homogeneous linear system $\mathbf{G}_{-j}^{\mathsf{T}} \vec{x} = 0$. Since we assume that the $\vec{G}_i$'s are all linearly independent, the columns of $\mathbf{G}_{-j}$ are also linearly independent [13]. Thus the rank of $\mathbf{G}_{-j}$ is $(N-1)$. The solution

TABLE I
BASIS COMPUTATION ALGORITHM

| | |
|---|---|
| 1: | IF $(K > N)$ |
| 2: | Solve homogeneous linear system $\mathbf{G}^{\mathsf{T}} \vec{x} = 0$ and get basis $[\vec{v}_1, \cdots, \vec{v}_{K-N}]$; |
| 3: | FOR $i = 1$ to $K - N$ |
| 4: | $\vec{e}_{j,i} = \vec{v}_i$, for all $j$; |
| 5: | END FOR |
| 6: | END IF |
| 7: | FOR $j = 1$ to $N$ |
| 8: | Orthogonalize $\mathbf{G}_{-j}$ and get $N-1$ orthogonal vectors $\vec{\omega}_{j,i}$'s; |
| 9: | Calculate $\vec{e}_{j,r}$ as in (23); |
| 10: | END FOR |

belongs to the null space of $\mathbf{G}_{-j}$. The dimension of the null space is $r = K - (N-1)$ according to the Rank-nullity Theorem [13]. Therefore, each $\vec{a}_j$ can be presented by the linear combination of $r$ linearly independent vectors. ■

Let $\mathbf{e}_j = \{\vec{e}_{j,1}, \vec{e}_{j,2}, \cdots, \vec{e}_{j,r}\}$ be a *basis* for the null space of $\mathbf{G}_{-j}$. There are many methods to obtain the basis, such as Gaussian Elimination. However, we show that it is not necessary to solve the homogeneous linear system $\mathbf{G}_{-j}^{\mathsf{T}} \vec{x} = 0$ to get the basis for every different $j$ value. Therefore the computational complexity can be further reduced.

The algorithm for computing a basis is shown in Table I. In Steps 1–6, we first solve the homogeneous linear system $\mathbf{G}^{\mathsf{T}} \vec{x} = 0$ to get a basis $[\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{K-N}]$. Note that if $K$ is equal to $N$, the basis is the empty set $\emptyset$. We then set the $K-N$ basis vectors to be the first $K-N$ vectors in all the basses $\mathbf{e}_j$, $j = 1, 2, \cdots, N$. In Step 8, we orthogonalize each $\mathbf{G}_{-j}$ and obtain $(N-1)$ orthogonal vectors $\vec{\omega}_{j,i}$, $i = 1, 2, \cdots, N-1$. Finally in Step 9, we let the $r$th vector $\vec{e}_{j,r}$ be orthogonal to all the $\vec{\omega}_{j,i}$'s by subtracting all the projections on each $\vec{\omega}_{j,i}$ from $\vec{G}_j$ (recall that $r = K - N + 1$). The operation is:

$$\vec{e}_{j,r} = \vec{e}_{j,N-K+1} = \vec{G}_j - \sum_{i=1}^{N-1} \frac{\vec{G}_j^{\mathsf{T}} \vec{\omega}_{j,i}}{\vec{\omega}_{j,i}^{\mathsf{T}} \vec{\omega}_{j,i}} \vec{\omega}_{j,i}. \tag{23}$$

*Lemma 3:* The solution space constructed by the basis $[\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{K-N}]$ is a sub-space of the solution space of $\mathbf{G}_{-j}^{\mathsf{T}} \vec{x} = 0$ for all $j$.

*Proof:* It can be seen that each vector $\vec{v}_i$ is a solution of $\mathbf{G}_{-j}^{\mathsf{T}} \vec{x} = 0$, for $i = 1, 2, \cdots, K - N$. ■

*Lemma 4:* The vectors $[\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{K-N}, \vec{e}_{j,r}]$ computed in Table I is a basis of the null space of $\mathbf{G}_{-j}$.

*Proof:* Obviously, the $\vec{v}_i$'s are linearly independent. From (23), it is easy to verify that $\vec{e}_{j,r}$ is orthogonal to all the $\vec{\omega}_{j,i}$'s. Therefore, $\vec{e}_{j,r}$ is also a solution to system $\mathbf{G}_{-j}^{\mathsf{T}} \vec{x} = 0$. Since $\vec{G}_j$ and $\vec{\omega}_{j,i}$ are orthogonal to all the $\vec{v}_i$'s, and $\vec{e}_{j,r}$ is a linear combination of $\vec{G}_j$ and $\vec{\omega}_{j,i}$, $\vec{e}_{j,r}$ is also orthogonal and linearly independent to all the $\vec{v}_i$'s. The conclusion follows. ■

Define coefficients $\vec{c}_j = [c_{j,1}, c_{j,2}, \cdots, c_{j,r}]^{\mathsf{T}}$. Then we can represent $\vec{a}_j$ as a linear combination of the basis vectors, i.e., $\vec{a}_j = \sum_{l=1}^{r} c_{j,l} \vec{e}_{j,l} = \mathbf{e}_j \vec{c}_j$. Eq. (19) can be rewritten as

$$\begin{aligned} \lambda_j^t &= \frac{\beta_j B}{T} \log_2 \left( 1 + \frac{1}{N_0} \left( \vec{c}_j^{\mathsf{T}} \mathbf{e}_j^{\mathsf{T}} \vec{G}_j \right)^2 \right) \\ &= \frac{\beta_j B}{T} \log_2 \left( 1 + \frac{1}{N_0} \left( c_{j,r} \vec{e}_{j,r}^{\mathsf{T}} \vec{G}_j \right)^2 \right). \end{aligned} \tag{24}$$

TABLE II
COMPARISON OF COMPUTATIONAL COMPLEXITY

| | Original Problem | Reformulated Problem |
|---|---|---|
| Primal Variables | $KN$ | $(K-N+1)N$ |
| Dual Variables | $N(N-1)+K$ | $K$ |

The second equality is because the first $K-N$ column vectors in $\mathbf{e}_j$ are orthogonal to $G_j$. The random variable $W_j^t$ in the objective function now only depends on $c_{j,r}$. The peak power constraint can be revised as:

$$\sum_{j=1}^{N}[\mathbf{e}_j(k)\vec{c}_j]^2 \leq P_{max}, \quad k=1,\cdots,K, \tag{25}$$

where $\mathbf{e}_j(k)$ is the $k$th row of matrix $\mathbf{e}_j$.

With such a reformulation, the number of primal and dual variables can be greatly reduced. In Table II, we show the numbers of variables in the original problem and in the reformulated problem. The number of primary variables is reduced from $KN$ to $(K-N+1)N$, and the number of dual variables is reduced from $N(N-1)+K$ to $K$. Such reductions result in greatly reduced computational complexity.

*3) Distributed Algorithm :* To solve the reformulated problem, we define non-negative dual variables $\vec{\mu}=[\mu_1,\cdots,\mu_K]^{\mathsf{T}}$ for the inequality constraints. The Lagrangian function is

$$\begin{aligned}\mathcal{L}(\mathbf{c},\vec{\mu}) &= \sum_{j=1}^{N}\mathbb{E}\left[\log(W_j^t(c_{j,r}))|w_j^t\right] + \\ &\quad \sum_{k=1}^{K}\mu_k(P_{max}-\sum_{j=1}^{N}[\mathbf{e}_j(k)\vec{c}_j]^2) \\ &= \sum_{j=1}^{N}\mathcal{L}_j(\vec{c}_j,\vec{\mu})+P_{max}\sum_{k=1}^{K}\mu_k, \end{aligned}\tag{26}$$

where $\mathbf{c}$ is a matrix consisting of all column vector $\vec{c}_j$'s and

$$\mathcal{L}_j(\vec{c}_j,\vec{\mu})=\mathbb{E}\left[\log(W_j^t(c_{j,r}))|w_j^t\right]-\sum_{k=1}^{K}\mu_k[\mathbf{e}_j(k)\vec{c}_j]^2.$$

The corresponding problem can be decomposed into $N$ sub-problems and solved iteratively [14]. In Step $\tau \geq 1$, for given vector $\vec{\mu}(\tau)$, each CR user solves the following sub-problem using local information

$$\vec{c}_j(\tau)=\arg\max \mathcal{L}_j(\vec{c}_j,\vec{\mu}(\tau)). \tag{27}$$

Obviously, the objective function in (27) is concave. Therefore, there is a unique optimal solution. The CR users then exchange their solutions over the common control channel. To solve the primal problem, we adopt the gradient method [14].

$$\vec{c}_j(\tau+1)=\vec{c}_j(\tau)+\phi\nabla\mathcal{L}_j(\vec{c}_j(\tau),\vec{\mu}(\tau)), \tag{28}$$

where $\nabla\mathcal{L}_j(\vec{c}_j(\tau),\vec{\mu}(\tau))$ is the gradient of the primal problem and $\phi$ is a small positive step size.

The master dual problem for a given $\mathbf{c}(\tau)$ is:

$$\min_{\mu_i\geq 0, i=1,\cdots,K} q(\vec{\mu})=\sum_{j=1}^{N}\mathcal{L}_j(\vec{c}_j(\tau),\vec{\mu})+P_{max}\sum_{k=1}^{K}\mu_k. \tag{29}$$

TABLE III
ALGORITHM FOR THE CASE OF A SINGLE CHANNEL

IF ($N=1$)
    Set $a_{k,j}$ to $P_{max}$ for all $k$;
ELSE
    Set $\tau=0$; $\vec{\mu}(0)$ to positive values; $\mathbf{c}(0)$ to random values;
    Compute bases $\mathbf{e}_j$'s as in Table I;
    DO
        $\tau=\tau+1$ ;
        Compute $\vec{c}_j(\tau)$ as in (28);
        Broadcast $\vec{c}_j(\tau)$ on the common control channel;
        Update $\vec{\mu}(\tau)$ as in (30);
    WHILE ($||\vec{\mu}(\tau)-\vec{\mu}(\tau-1)|| > \kappa$);
    Compute $a_{k,j}$'s;
END IF

Since the Lagrangian function is differentiable, the subgradient iteration method can be adopted.

$$\vec{\mu}(\tau+1)=[\vec{\mu}(\tau)-\rho(\tau)\vec{g}(\tau)]^+, \tag{30}$$

where $\rho(\tau)=\frac{q(\vec{\mu}(\tau))-q(\vec{\mu}^*)}{||\vec{g}(\tau)||^2}$ is a positive step size, $\vec{\mu}^*$ is the optimal solution, $\vec{g}(\tau)=\nabla q(\vec{\mu}(\tau))$ is the gradient of the dual problem, and $[\cdot]^+$ denotes the projection onto the nonnegative axis. Since the optimal solution $\vec{\mu}^*$ is unknown a priori, we choose the mean of the objective values of the primal and dual problems as an estimate for $\vec{\mu}^*$ in the algorithm. The updated $\mu_k(\tau+1)$ will again be used to solve the sub-problems (27). Since the problem is convex, we have strong duality; the duality gap between the primal and dual problems will be zero. The distributed algorithm is presented in Table III, where $0 \leq \kappa \ll 1$ is a threshold for convergence.

We analyze the performance of the distributed algorithm. In particular, we prove that it converges to the optimal solution at a speed faster than $1/\sqrt{\tau}$ as $\tau$ goes to infinity. The proofs are omitted due to lack of space.

### B. Case of Multiple Channels with Channel Bonding

When there are multiple licensed channels, we first consider the case where the channel bonding/aggregation techniques are used by the transmitters and CR users [15], [16]. With channel bonding, a transmitter can utilize all the available channels in $\mathcal{A}(t)$ collectively to transmit the mixed signal. We assume that at the end of the sensing phase in each time slot, CR users tune their SDR transceiver to the common control channel to receive the set of available channels $\mathcal{A}(t)$ from the BS. Then each CR user can receive from all the channels in $\mathcal{A}(t)$ and decode its desired signal from the compound signal it receives.

This case is similar to the case of single licensed channel. Now all the active CR users receive from the same set of available channels $\mathcal{A}(t)$. We thus have $b_j^m=1$, for all $m\in\mathcal{A}(t)$, and $b_j^m=0$, for all $m\notin\mathcal{A}(t)$, $j=1,2,\cdots,N$. When all the $b_j^m$'s are determined this way, problem (16) is reduced to a nonlinear programming problem with constraints (11), (12), and (18). The distributed algorithm described in Section IV-A can be still applied to solve this reduced problem to obtain optimal solutions.

### C. Case of Multiple Channels without Channel Bonding

We finally consider the case of multiple channels without channel bonding, where each CR user has a narrow band SDR

TABLE IV
CHANNEL SELECTION ALGORITHM FOR THE CASE OF MULTIPLE CHANNELS WITHOUT CHANNEL BONDING

| | |
|---|---|
| 1: | Initialize $\vec{b}$ to a zero vector, user set $\mathcal{U} = \{1, \cdots, N\}$ and user-channel set $\mathcal{C} = \mathcal{U} \times \mathcal{A}(t)$; |
| 2: | WHILE ($\mathcal{C} \neq \emptyset$) |
| 3: | Find the user-channel pair $\{j', m'\}$, such that $\{j', m'\} = \arg\max_{\{\{j,m\} \in \mathcal{C}\}} \{\Phi(\vec{b} + \vec{v}_j^m) - \Phi(\vec{b})\}$; |
| 4: | Set $\vec{b} = \vec{b} + \vec{v}_{j'}^{m'}$ and remove $j'$ from $\mathcal{U}$; |
| 5: | IF ($\sum_{j=1}^{N} b_j^{m'} = K$) |
| 6: | Remove $m'$ from $\mathcal{A}(t)$; |
| 7: | END IF |
| 8: | Update user-channel set $\mathcal{C} = \mathcal{U} \times \mathcal{A}(t)$; |
| 9: | END WHILE |

transceiver and can only receive from one of the licensed channels. We present a greedy algorithm that leverages the optimal algorithm in Table III for near-optimal solutions, which has a lower bound for its performance.

When $M > 1$, the optimal solution to problem (16) also depends on the channel selection variables $b_j^m$'s, which determines whether user $j$ receives from channel $m$. Recall that there are two constraints for the $b_j^m$'s: (i) each user can use at most one channel (see (14)); (ii) the number of users on the same channel cannot exceed the number of transmitters $K$ (see (22)). Let $\vec{b}$ be the channel allocation vector with elements $b_j^m$'s, and $\Phi(\vec{b})$ the corresponding objective value for a given user channel allocation $\vec{b}$.

We take a two-step approach to solve problem (16). First, we apply the greedy algorithm in Table IV to choose one available channel in $\mathcal{A}(t)$ for each CR user (i.e., to determine $\vec{b}$). Second, we apply the algorithm in Table III to obtain a near-optimal solution for the given channel allocation $\vec{b}$.

In Table IV, $\vec{v}_j^m$ is a unit vector with 1 for the $[(j-1) \times M + m]$-th element and 0 for all other elements, and $\vec{b} = \vec{b} + \vec{v}_{j'}^{m'}$ indicates choosing channel $m'$ for user $j'$. In each iteration, the user-channel pair $(j', m')$ that can achieve the largest increase in the objective value is chosen, as in Step 3. The complexity of the greedy algorithm in the worst case is $O(K^2 M^2)$.

It can be shown that the greedy algorithm solution is lower bounded by $1/|\mathcal{A}(t)|$ of the global optimum $\Phi(\Omega)$, as

$$\frac{1}{|\mathcal{A}(t)|}\Phi(\Omega) \leq \Phi(\vec{b}) \leq \Phi(\Omega). \quad (31)$$

The proofs are omitted due to lack of space.

## V. PERFORMANCE EVALUATION

We evaluate the performance of the proposed algorithms with a MATLAB implementation and the JVSM 9.13 Video Codec. We present simulation results for the following two scenarios:

- Single licensed channel
- Multiple licensed channels without channel bonding,

since we observe similar performance for the case of multiple licensed channels with channel bonding.

For comparison purpose, we also developed and simulated the following two simpler heuristic schemes that do not incorporate interference alignment.

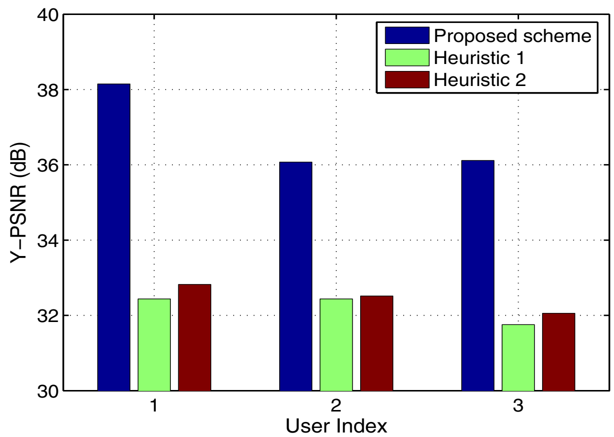


Fig. 2. Received video quality for each CR user with a single channel.

- *Heuristic 1*: each CR user selects the best channel in $\mathcal{A}(t)$ based on channel condition. The time slot is equally divided among the active users receiving from the same channel, to send their signals separately in each time slice.
- *Heuristic 2*: in each time slot, the active user with the best channel is selected for each available channel. The entire time slot is used to transmit this user's signal.

### A. Case of Single Licensed Channel

In the first scenario, there are $K = 4$ transmitters, i.e., one BS and three RNs. The channel utilization $\eta$ is set to 0.6 and the maximum allowable collision probability $\gamma$ is set to 0.2. There are three active CR users, each receives an MGS video stream from the BS: *Bus* to CR user 1, *Mobile* to CR user 2, and *Harbor* to CR user 3. The video sequences are in the Common Intermediate Format (CIF, 252×288). The GOP size of the videos is 16 and the delivery deadline $T$ is 10. The false alarm probability is $\epsilon_l^m = 0.3$ and the miss detection probability is $\delta_l^m = 0.3$ for all spectrum sensors. The channel bandwidth $B$ is 1 MHz. The peak power limit is 10 W for all the transmitters, unless otherwise specified.

We first plot the average Y-PSNRs of the three reconstructed MGS videos in Fig. 2, i.e., only the Y (Luminance) component of the original and reconstructed videos are used. Among three schemes, the proposed algorithm achieves the highest PSNR value, while the two heuristic algorithms have similar performance. Note that the proposed algorithm is optimal in the single channel case. It achieves significant improvements ranging from 3.1 dB to 5.25 dB over the two heuristic algorithms. Such PSNR gains are considerable, since in video coding and communications, a half dB gain is distinguishable and worth pursing.

We next examine the convergence rate of the distributed algorithm. In Section IV, it is shown that the distributed algorithm converges at a speed faster than $1/\sqrt{\tau}$ asymptotically. We compare the optimality gap of the proposed algorithm, i.e., $|q(\tau) - q^*|$, with series $10/\sqrt{\tau}$ in Fig. 3. Both curves converge to 0 as $\tau$ goes to infinity. It can be seen that the convergence speed, i.e., the slope of the curve, of the proposed scheme

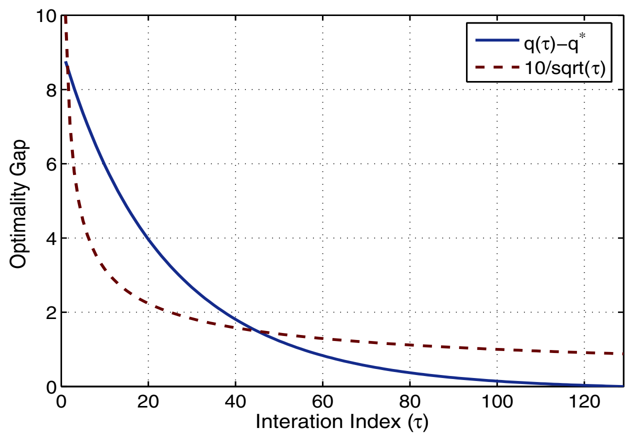


Fig. 3. Convergence rate of the distributed algorithm with a single channel.

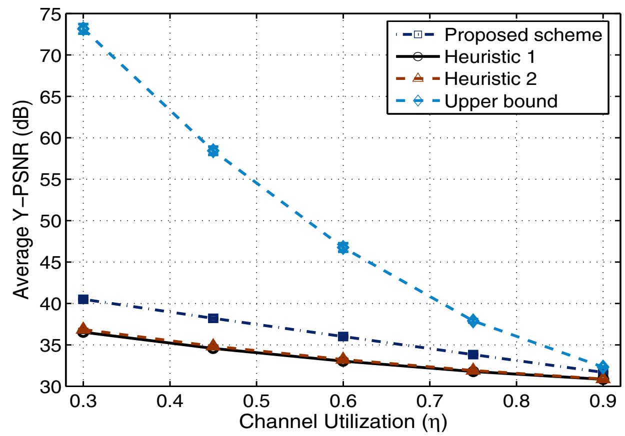


Fig. 4. Reconstructed video quality vs. channel utilization $\eta$ in the multi-channel without channel bonding case.

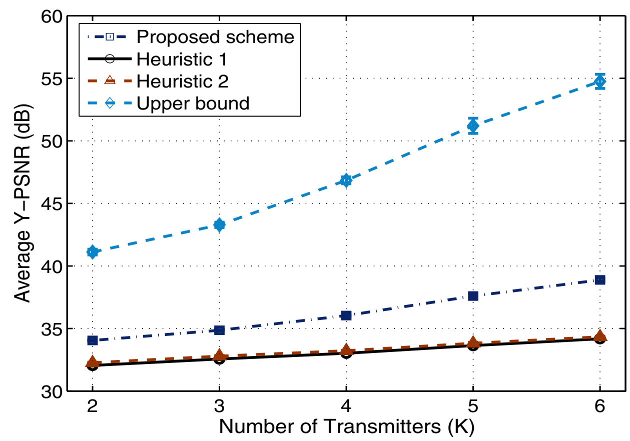


Fig. 5. Reconstructed video quality vs. number of transmitters $K$ in the multi-channel without channel bonding case.

is larger than that of $10/\sqrt{\tau}$ after about 10 iterations. The convergence of the optimality gap is much faster than $10/\sqrt{\tau}$, which exhibits a heavy tail.

In the case of multiple channels with channel bonding, the performance of the proposed algorithm is similar to that in the single channel case. We omit the results for lack of space.

### *B. Case of Multiple Channels without Channel Bonding*

The second scenario has six licensed channels and four transmitters. There are 12 CR users, each streaming one of the three different videos *Bus*, *Mobile*, and *Harbor*. The rest of the parameters are the same as those in the single channel case, unless otherwise specified. Eq. (31) can also be interpreted as an upper bound on the global optimal, i.e., $\Phi(\Omega) \leq |\mathcal{A}(t)|\Phi(\vec{b})$, which is also plotted in the figures. Each point in the following figures is the average of 10 simulation runs with different random seeds. The 95% confidence intervals are plotted as error bars, which are generally negligible.

The impact of channel utilization $\eta$ on received video quality is presented in Fig. 4. We increase $\eta$ from 0.3 to 0.9 in steps of 0.15, and plot the Y-PSNRs of reconstructed videos averaged over all the 12 CR users. Intuitively, a smaller $\eta$ allows more transmission opportunities for CR users, thus allowing the CR users to achieve higher video rates and better video quality. This is shown in the figure, in which all four curves decrease as $\eta$ is increased. We also observe that the gap between the upper bound and proposed schemes becomes smaller as $\eta$ gets larger, from 32.65 dB when $\eta = 0.3$ to 0.63 dB when $\eta = 0.9$. The proposed scheme outperforms the two heuristic schemes with considerable gains, ranging from 0.8 dB to 3.65 dB.

Finally, we investigate the impact of the number of transmitters $K$ on the video quality. In this simulation, we increase $K$ from 2 to 6 with step size 1. The average Y-PSNRs of all the 12 CR users are plotted in Fig. 5. As expected, the more transmitters, the more effective the interference alignment technique, and thus the better the video quality. The proposed algorithm achieves gains ranging from 1.78 dB (when $K = 2$) to 4.55 dB (when $K = 6$) over the two heuristic schemes.

## VI. Related Work

This work is closely related to the prior work on cooperative communications [3], [4], where relays are used to achieve cooperative diversity, and that on CR networking [2], where spectrum opportunities in licensed channels are exploited for the benefit of unlicensed users. There have been significant advances in these areas, which laid out the foundation for this work. In particular, researchers have been exploring the idea of combining these two techniques [5]–[7]. In [5], an overview of cooperative relay scenarios and related issues was presented, along with a GNU Radio implementation of a MAC protocol. In [6], a centralized heuristic was presented to address the relay selection and spectrum allocation problem in CR networks.

The problem of video over CR networks has only been studied in a few recent papers [17], [18], [20], [22]–[26]. In [22], a dynamic channel selection scheme was proposed for CR users to transmit videos over multiple channels. In [23], a distributed joint routing and spectrum sharing algorithm for video streaming over CR ad hoc networks was described and evaluated with simulations. In our prior work, we considered video multicast in an infrastructure-based CR network [17],

[24], unicast video streaming over multihop CR networks [20] and CR femtocell networks [18], [25]. In [26], the impact of system parameters residing in different network layers are jointly considered to achieve the best possible video quality for CR users. Unlike the heuristic approaches in [22], [23], the analytical and optimization approach taken in this paper yields algorithms with optimal or bounded performance. The cooperative relay and interference alignment techniques also distinguish this paper from prior work on this topic.

As point-to-point link capacity approaches the Shannon limit, there has been considerable interest on exploiting interference to improve wireless network capacity [8]–[12]. In addition to information theoretic work on asymptotic capacity [8], [9], practical issues have been addressed in [10]–[12]. In [11], the authors presented a practical design of analog network coding to exploit interference and allow concurrent transmissions, which does not make any synchronization assumptions. In [12], interference alignment and cancellation is incorporated in MIMO LANs, and the network capacity is shown, analytically and experimentally, to be almost doubled. In [10], the authors presented a general algorithm for identifying interference alignment and cancellation opportunities in practical multi-hop mesh networks. The impact of synchronization and channel estimation was evaluated through a GNU Radio implementation. Our work was motivated by these interesting papers, and we incorporate interference alignment in cooperative CR networks and exploit the enhanced capacity for wireless video streaming.

## VII. Conclusion

In this paper, we investigated the problem of interference alignment for MGS video streaming in a cooperative relay enhanced CR network. We presented a stochastic programming formation, and derived a reformulation that leads to considerable reduction in computational complexity. A distributed optimal algorithm was developed for the case of a single channel and the case of multi-channel with channel bonding, with proven convergence and convergence speed. We also presented a greedy algorithm for the multi-channel without channel bonding case, with a proven performance bound. The proposed algorithms are evaluated with simulations and are shown to outperform two heuristic schemes without interference alignment with considerable gains.

## Acknowledgment

This work is supported in part by the US National Science Foundation (NSF) under Grants CNS-0953513, ECCS-0802113, CNS-1145446 and IIP-1127952, and DUE-1044021, and through the NSF Wireless Internet Center for Advanced Technology. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

## References

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2010-2015," Feb. 2010, [online] Available: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html.

[2] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.

[3] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity – Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[4] J. N. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[5] Q. Zhang, J. Jia, and J. Zhang, "Cooperative relay to improve diversity in cognitive radio networks," *IEEE Commun. Mag.*, vol. 47, no. 2, pp. 111–117, Feb. 2009.

[6] J. Jia, J. Zhang, and Q. Zhang, "Cooperative relay for cognitive radio networks," in *Proc. IEEE INFOCOM'09*, Apr. 2009, pp. 2304–2312.

[7] D. Hu and S. Mao, "Cooperative relay in cognitive radio networks: Decode-and-forward or amplify-and-forward?" in *Proc. IEEE GLOBECOM 2010*, Miami, FL, Dec. 2010, pp. 1–5.

[8] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, UK: Cambridge University Press, 2005.

[9] V. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom for the $k$ user ntererence channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, May 2008.

[10] L. E. Li, R. Alimi, D. Shen, H. Viswanathan, and Y. R. Yang, "A general algorithm for interference alignment and cancellation in wireless networks," in *Proc. IEEE INFOCOM 2010*, San Diego, CA, Mar. 2010, pp. 1–9.

[11] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," in *Proc. ACM SIGCOMM'07*, Kyoto, Japan, Aug. 2007, pp. 397–408.

[12] S. Gollakota, S. David, and D. Katabi, "Interference alignment and cancellation," in *Proc. ACM SIGCOMM'09*, Barcelona, Spain, Aug. 2009, pp. 159–170.

[13] G. Strang, *Introduction to Linear Algebra*, 4th ed. Wellesley, MA: Wellesley Cambridge Press, 2009.

[14] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Nashua, NH: Athena Scientific, 1999.

[15] H. Mahmoud, T. Yucek, and H. Arslan, "OFDM for cognitive radio: Merits and challenges," *IEEE Wireless Commun. Mag.*, vol. 16, no. 2, pp. 6–14, Apr. 2009.

[16] C. Corderio, K. Challapali, D. Birru, and S. Shankar, "IEEE 802.22: An introduction to the first wireless standard based on cognitive radios," *J. Commun.*, vol. 1, no. 1, pp. 38–47, Apr. 2006.

[17] D. Hu and S. Mao, "Streaming scalable videos over multi-hop cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3501–3511, Nov. 2010.

[18] ——, "Resource allocation for medium grain scalable videos over femtocell cognitive radio networks," in *Proc. IEEE ICDCS 2011*, Minneapolis, MN, June 2011, pp. 258–267.

[19] I. K. Son and S. Mao, "Design and optimization of a tiered wireless access network," in *Proc. IEEE INFOCOM 2010*, San Diego, CA, Mar. 2010, pp. 1–9.

[20] D. Hu, S. Mao, Y. T. Hou, and J. H. Reed, "Fine grained scalability video multicast in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 334–344, Apr. 2010.

[21] M. Wien, H. Schwarz, and T. Oelbaum, "Performance analysis of SVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1194–1203, 2007.

[22] H.-P. Shiang and M. van der Schaar, "Dynamic channel selection for multi-user video streaming over cognitive radio networks," in *Proc. IEEE ICIP'08*, San Diego, CA, Oct. 2008, pp. 2316–2319.

[23] L. Ding, S. Pudlewski, T. Melodia, S. Batalama, J. Matyjas, and M. Medley, "Distributed spectrum sharing for video streaming in cognitive radio ad hoc networks," in *Intl. Workshop on Cross-layer Design in Wireless Mobile Ad Hoc Networks*, Niagara Falls, Canada, Sept. 2009, pp. 1–13.

[24] D. Hu, S. Mao, and J. H. Reed, "On video multicast in cognitive radio networks," in *Proc. IEEE INFOCOM 2009*, Rio de Janeiro, Brazil, Apr. 2009, pp. 2222–2230.

[25] D. Hu and S. Mao, "On medium grain scalable video streaming over cognitive radio femtocell networks," *IEEE J. Sel. Areas Commun.*, vol. 30, p. 4, Apr. 2012.

[26] H. Luo, S. Ci, and D. Wu, "A cross-layer design for the performance improvement of real-time video transmission of secondary users over cognitive radio networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 8, pp. 1040–1048, Aug. 2011.