2022 IEEE Wireless Communications and Networking Conference (WCNC) | 978-1-6654-4266-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/WCNC51071.2022.9771640

Joint Robust Relay Beamforming and Adaptive Channel Estimation using Cubature Kalman Filtering

Mohammad Amin Maleki Sadr
Department of Electrical and Computer Engineering
McGill University
Montreal, Canada
mohammadamin.malekisadr@mail.mcgill.ca

Benoit Champagne
Department of Electrical and Computer Engineering
McGill University
Montreal, Canada.
Benoit.champagne@mcgill.ca

***Abstract*—In this paper, an adaptive algorithm is proposed for the estimation and tracking of the channel coefficients in peer-to-peer communication through a network of relays. Using the observed signals at the relay and destination nodes, the channel state information (CSI) is estimated centrally by taking advantage of a Markov model for the source-relay and relay-destination channels, and employing the Cubature Kalman Filter (CKF). The estimated CSI is used for solving a robust relay beamforming problem, aiming to minimize the total transmitted power by the relays subject to signal-to-interference-plus-noise ratio (SINR) constraint at each one of the destination nodes. Through simulations, the proposed CSI estimation is shown to be unbiased and converge to the Cramer-Rao-Lower-Bound (CRLB) for low and moderate error levels. Furthermore, the ensuing beamformer design exhibits better performance compared to existing robust beamforming methods.**

***Index Terms*—Cramer-Rao-Lower-Bound, Cubature Kalman Filter, Channel State Information.**

I. Introduction

Cooperative communication has received significant interest as a promising way of achieving spatial diversity without using Multiple-Input-Multiple-Output (MIMO) technology [1], [2]. Peer-to-peer communication using a network of relays (also known as cooperative relaying) is a cooperative scheme in which each source in a pair wishes to communicate with its corresponding destination through the relays [3]. Employing the relay nodes between the source-destination pairs reduces the effect of channel degradation and mitigates the interference received from other nodes; thus, it can provide reliable communication between source-destination pairs. Among different cooperative relaying schemes, such as amplify-and-forward (AF) [3], decode-and-forward (DF) [4], and coded-cooperation-and-forward (CF) [5], AF is much more attractive due to its relatively low-complexity of implementation and higher security. In this scheme, a commonly adopted strategy to increase the SINR at the target destination is performing cooperative beamforming using the relay nodes.

One major problem in the practical implementation of a cooperative relaying system is that the Channel State Information (CSI) cannot be precisely estimated as it contains errors due to limited channel feedback, channel quantization errors, or feedback delays. Designing the beamformer requires accurate knowledge about CSI [6]; thus, establishing a reliable relay transmission is challenging under channel variations.

In the literature on cooperative robust relay beamforming, authors assume that they can somehow provide an initial estimation of the channel with error, while is unknown and does not decrease by receiving more measurements over time [6]–[8]. In the mentioned papers, the assumed channel error model does not change dynamically while, in the real communications world by observing new measurement data, the channel uncertainty can be decreased. Two different robust approaches have been proposed in the literature which are: 1) stochastic [6], [7], and 2) worst-case methods [8] whose applications depend on the way that the error is modelled. In this respect, in [6], authors proposed a stochastic-based robust method for tuning the relays' beamform weights. In [9], the authors proposed a method to model the channel error dynamically. While this method decreases the CSI error receiving new measurements, it suffers from non-convexity and ambiguity at the starting point, impacting the convergence.

In this paper, we assume there exists a processing center that has access to the destinations and relays observed signals and can fuse information of these two sources. The available information at the processing center, called measurements, used to estimate the channel coefficients (states). The measurements are a function of the beamforming weights which are optimized at each time step. In our designed method, the CSI be updated whenever we observe new data in the relays or destination nodes. Note that due to the low complexity of implementations we used AF relaying in this paper. Furthermore, the contributions of this paper are summarized as follows:

- The discussed robust approach is a real-time adaptive approach that mitigates the channel error at each time step, unlike the previous works in [6]–[8].
- In our approach, the convergence of the channel estimation method does not depend on the starting point, unlike the robust adaptive approach in [9].
- Due to the non-linear nature of measurements with respect to the states of the channels, we have utilized the Cubature Kalman Filter (CKF) [10], for tracking the CSI. The CKF is shown to have a better performance in highly non-linear measurements compared to other non-linear approximations of KF [10].
- A robust optimization approach for minimizing the total

This work was supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada

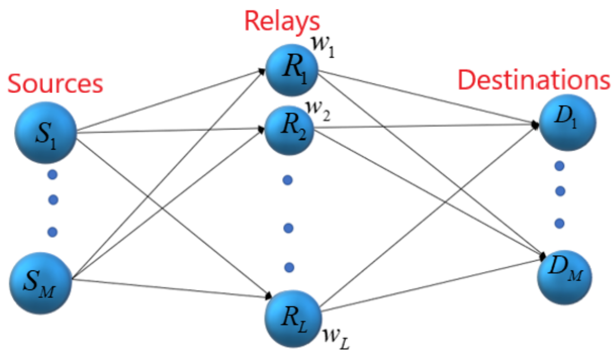

Figure 1: System Model: a cooperative relay beamforming with M source-destination pairs and L relays.

power transmission of the relays subject to probabilistic SINR constraint at each of the destinations is proposed.

- The results of simulations show that our proposed method is unbiased and achieves the Cramer-Rao-Lower-Bound (CRLB) for low and moderate error levels which shows the efficiency of our estimation.

Notations: In this paper, the notation $\odot$ stands for Schur-Hadamard (element-wise) multiplication and $\mathrm{BD}\,(.)$ denotes the block diagonalization of an array of matrices.

II. System Model and Problem Statement

A. System Model

We consider a network for point-to-point transmission between M source-destination pairs, through a layer of L relays operating in parallel, as illustrated in Fig. 1. The sources $\mathcal{S} = \{S_i\}_{i=1}^{M}$, relays $\mathcal{R} = \{R_r\}_{r=1}^{L}$, and destinations $\mathcal{D} = \{D_i\}_{i=1}^{M}$, are all assumed to be equipped with a single antenna. Assume that $f_{pr}(k)$, and $g_{rp}(k)$ are the channel coefficients between the pth source to rth relay at kth time step, and the channel coefficient between the rth relay to pth destination at kth time step, respectively. Assume that there exists no direct link between the source-destination pairs. For notation simplicity, assume that we drop the time step k from the channel coefficients (e.g. $f_{pr}(k) \rightarrow f_{pr}$, $g_{rp}(k) \rightarrow g_{rp}$), i.e., all the equations in this Section are for time step kth. Suppose that $\mathbf{v}_m^{\mathbf{x}}$ and $v_m^{y_j}$ are additive white Gaussian noise at relays and jth destination, respectively. The relays observed signal is

$$\mathbf{x} = \sum_{p=1}^{M} \mathbf{f}_p s_p + \mathbf{v}_m^{\mathbf{x}} \tag{1}$$

where $\mathbf{x} = [x_1, x_2 ... x_L]^T$, $\mathbf{v}_m^{\mathbf{x}} = [v_{m1,}^{\mathbf{x}}, v_{m2}^{\mathbf{x}} ... v_{mL}^{\mathbf{x}}]^T$, $\mathbf{f}_p = [f_{p1}, f_{p2} ... f_{pL}]^T$ and $\{s_p\}_{p=1}^{M}$ are information symbols. The rth relay multiplies its received signal by a complex weight w_r^*, then transmits it to the target destination. The relays transmitted signal using AF approach can be written as

$$\mathbf{t} = \mathbf{W}^H \mathbf{x} \tag{2}$$

where $\mathbf{W} = \mathrm{diag}(\mathbf{w})$ and $\mathbf{w} = [w_1, w_2 ... w_L]^T$, the received signal at jth destination node is denoted as

$$\begin{aligned} y_j &= \mathbf{g}_j{}^T \mathbf{t} + \eta_j \qquad \forall j \in \mathcal{J} \\ &= \mathbf{g}_j^T \mathbf{W}^H \mathbf{f}_j s_j + \mathbf{g}_j{}^T \mathbf{W}^H \sum_{p \neq j}^{M} \mathbf{f}_p s_p + v_m^{y_j} \end{aligned} \tag{3}$$

where $\mathcal{J} = \{1 \dots M\}$, $v_m^{y_j} = \mathbf{g}_j^T \mathbf{W}^H \mathbf{v}_m^{\mathbf{x}} + \eta_j$. SINR at the jth destination node is written as

$$\mathrm{SINR}_j = \frac{P_s^j}{P_N^j + P_I^j} \qquad j \in \mathcal{J} \tag{4}$$

where P_s^j, P_I^j, P_N^j are the desired signal, interference, and noise power at the jth destination, respectively. By using (3), the total transmit power of the relays is

$$\begin{aligned} P_T &= \mathrm{E}\left\{\|\mathbf{t}\|^2\right\} = \mathrm{E}\left\{\mathrm{Tr}\left(\mathbf{t}^{\mathrm{H}}\mathbf{t}\right)\right\} = \mathrm{E}\left\{\mathbf{x}^H \mathbf{W}\mathbf{W}^H \mathbf{x}\right\} \\ &= \sum_{r=1}^{L} |w_r|^2 [\mathbf{R}_{\mathbf{x}}]_{r,r} = \mathbf{w}^H \mathbf{D} \mathbf{w} \end{aligned} \tag{5}$$

where P_T is the total power transmission of the relays and $\mathbf{D} = \mathrm{diag}\left([\mathbf{R}_{\mathbf{x}}]_{1,1} [\mathbf{R}_{\mathbf{x}}]_{2,2} \dots [\mathbf{R}_{\mathbf{x}}]_{L,L}\right)$ and $\mathbf{w} = \mathrm{diag}\,(\mathbf{W})$ and $\mathbf{R}_{\mathbf{x}} = \mathrm{E}\left\{\mathbf{x}\mathbf{x}^H\right\}$. The correlation matrix of the received signal at the relays is computed as

$$\begin{aligned} \mathbf{R}_{\mathbf{x}} &= \sum_{p,q=1}^{M} \mathrm{E}\left\{\mathbf{f}_p \mathbf{f}_q^H\right\} \mathrm{E}\left\{s_p s_q\right\} + \sigma_{N1}^2 \mathbf{I} \\ &= P_p \left(\bar{\mathbf{f}}_p \bar{\mathbf{f}}_p^H + \sigma_f^2 \mathbf{I}\right) + \sigma_{N1}^2 \mathbf{I} \end{aligned} \tag{6}$$

where $P_p = \mathrm{E}\{s_p s_p^*\}$. Using (3), the jth desired signal power can be written as

$$\begin{aligned} P_s^j &= \mathrm{E}\left\{\mathbf{g}_j^T \mathbf{W}^H \mathbf{f}_j \mathbf{f}_j^H \mathbf{W} \mathbf{g}_j^*\right\} \mathrm{E}\{s_j^2\} \\ &= P_j (\mathbf{w}^H (\mathbf{g}_j \odot \mathbf{f}_j)(\mathbf{g}_j \odot \mathbf{f}_j)^H \mathbf{w}) = \mathbf{w}^H \mathbf{R}_h^j \mathbf{w} \end{aligned} \tag{7}$$

where $\mathbf{h}_j^j$ and $\mathbf{R}_h^j$ are defined as

$$\mathbf{R}_h^j \triangleq P_j \mathbf{h}_j \mathbf{h}_j^H, \; \mathbf{h}_j^j \triangleq \mathbf{g}_j \odot \mathbf{f}_j$$

Similarly, the interference power at the jth destination can be written as

$$\begin{aligned} P_I^j &= \mathrm{E}\{\mathbf{g}_j^T \mathbf{W}^H (\sum_{p \neq q}^{M} \mathbf{f}_p \mathbf{f}_q^H s_p s_q^*) \mathbf{W} \mathbf{g}_k^*\} \\ &= \mathbf{w}^H \sum_{p \neq j} P_p ((\mathbf{g}_j \odot \mathbf{f}_p)(\mathbf{g}_j \odot \mathbf{f}_p)^H) \mathbf{w} = \mathbf{w}^H \mathbf{Q}_j \mathbf{w} \end{aligned} \tag{8}$$

where

$$\mathbf{h}_j^p \triangleq \mathbf{g}_j \odot \mathbf{f}_p, \;\; \mathbf{Q}_j \triangleq \sum_{p \neq j}^{M} P_j \mathbf{h}_j^p (\mathbf{h}_j^p)^H$$

Similarly, the noise power at the jth destination is

$$\begin{aligned} P_N^j &= \mathrm{E}\left\{(\mathbf{g}_j^T \mathbf{W}^H \mathbf{v}_m^{\mathbf{x}} + \eta_j)(\mathbf{g}_j^T \mathbf{W}^H \mathbf{v}_m^{\mathbf{x}} + \eta_j)^H\right\} \\ &= \sigma_{N_1}^2 \mathbf{w}^H \mathrm{diag}(\mathbf{g}_j \mathbf{g}_j^*) \mathbf{w} + \sigma_{N_2}^2 = \mathbf{w}^H \mathbf{D}_j \mathbf{w} + \sigma_{N_2}^2 \end{aligned} \tag{9}$$

where $\mathbf{D}_j \triangleq \sigma_{N_1}^2 \mathrm{diag}(\mathbf{g}_j^*) \mathrm{diag}(\mathbf{g}_j)$. Hence, the SINR at the jth destination is

$$\mathrm{SINR}_j = \frac{\mathbf{w}^H \mathbf{R}_h^j \mathbf{w}}{\mathbf{w}^H \left(\mathbf{Q}_j + \mathbf{D}_j\right) \mathbf{w} + \sigma_{N_2}^2} \qquad j \in \mathcal{J} \tag{10}$$

B. Problem Statement

We want to design the beamforming weights at relays for all of the possible channel variations that guarantee the relays consume the minimum possible power while QoS at destination nodes preserves. The designed beamforming weights at relays and SINR at jth destination are the functions of CSI. More specifically, the CSI needs to be estimated in order to solve the following optimization problem

$$\min_{\mathbf{w}} \quad \mathbf{w}^H \mathbf{D} \mathbf{w} \tag{11a}$$

$$\text{s.t.} \quad \Pr_{\mathbf{f},\mathbf{g}_j} (\mathrm{SINR}_j \geq \gamma_j) \leq \epsilon_j \qquad j \in \mathcal{J} \tag{11b}$$

where γ_j is a preselected SINR threshold value at the jth destination node and ϵ_j is a maximum threshold outage probability at the jth destination node. We intend to solve optimization problem (11) the in the real-time by considering CSI rapid changes. This optimization problem needs to be solved at each specific time step. In our previous research study in [6], we solved the stochastic version of the aforementioned robust beamforming problem supposing the perfect knowledge of the distribution of the channel error and the mean and covariance of this distribution to be known. However, we do not have such information about the channel in many situations. Also, in [9], we discussed the joint prediction of beamforming and CSI, we reformulated the non-convex optimization problem as a constrained least square problem. The major drawback of this solution is that it has high complexity and requires the initial estimation of beamforming weights which may not exist in practical scenarios. In Section III, an optimization procedure for solving (11) is proposed.

III. Stochastic Robust Relays Beamforming

In this section, we propose a probabilistic method for making the problem robust against CSI uncertainty. It has been proved that considering the exact CSI in the situation that the CSI has uncertainty, severely degraded the performance [6]. By considering the uncertainty in the channel coefficients, the channel coefficient between sources and relays can be written as

$$\mathbf{f} = \bar{\mathbf{f}} + \Delta\mathbf{f} \tag{12}$$

where $\mathbf{f}$, $\bar{\mathbf{f}}$ and $\Delta\mathbf{f}$ are the estimated, the actual value, and the corresponding complex Gaussian error vector, respectively. Similarly, for the relay to the destination counterpart, the channel coefficients between the relays and the jth destination is

$$\mathbf{g}_j = \bar{\mathbf{g}}_j + \Delta\mathbf{g}_j \qquad j \in \mathcal{J} \tag{13}$$

where $\mathbf{g}_j$, $\bar{\mathbf{g}}_j$ and $\Delta\mathbf{g}_j$ are the estimation, the actual value, and the corresponding complex Gaussian error vector, respectively. Let us define an auxiliary variable Z_j as follows:

$$Z_j = \mathbf{w}^H \left(\mathbf{R}_h^j - \gamma_j \mathbf{Q}_j - \gamma_j \mathbf{D}_j \right) \mathbf{w} \qquad j \in \mathcal{J} \tag{14}$$

As a result, the constraints of our optimization problem in (11) can be rewritten as

$$\Pr \left(Z_j \leq \gamma_j \sigma_{N2}^2 \right) \leq \varepsilon_j \qquad j \in \mathcal{J} \tag{15}$$

By considering the statistics of uncertainty of the sources-relays channel coefficients, the error vector is modelled as

$$\Delta\mathbf{f} \sim \mathcal{CN}(0, \sigma_f^2 \mathbf{I}) \qquad j \in \mathcal{J} \tag{16}$$

Similarly, for the relays to the jth destination counterpart, the error vector is modelled as

$$\Delta\mathbf{g}_j \sim \mathcal{CN}\left(0, \sigma_{\mathbf{g}}^2 \mathbf{I}\right) \qquad j \in \mathcal{J} \tag{17}$$

It has been shown in [6] that (14) has the terms in first, second, third and the forth order terms of $\Delta\mathbf{f}$, and $\Delta\mathbf{g}_j$. In [6], we showed that the third and forth order error terms can be neglected. The following theorem shows that a certain set of second-order cone constraints can serve as a tractable convex approximation of the chance constraints (15).

Theorem 1. *Let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, ..., \boldsymbol{\xi}_m$ be independent standard Gaussian random variables. Consider the function $Q : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ defined via*

$$\begin{aligned} Q(\boldsymbol{x}, \boldsymbol{\xi}) = &-a_0(\mathbf{x}) + \sum_{i=1}^{m} \boldsymbol{\xi}_i a_i(\mathbf{x}) + \sum_{i=1}^{m} \boldsymbol{\xi}_i \boldsymbol{\xi}_j a_{i,j}(\mathbf{x}) \\ &+ \sum_{i=1}^{m} \boldsymbol{\xi}_i \boldsymbol{\xi}_j \boldsymbol{\xi}_k a_{i,j,k}(\mathbf{x}) + \sum_{i=1}^{m} \boldsymbol{\xi}_i \boldsymbol{\xi}_j \boldsymbol{\xi}_k \boldsymbol{\xi}_l a_{i,j,k,l}(\mathbf{x}) \end{aligned} \tag{18}$$

where $a_0(.)$ is affine and $a_i(.)$, $a_{i,j}(.)$, $a_{i,j,k}(.)$, $a_{i,j,k,l}(.)$ are linear in their arguments. Consider the chance constraint

$$\Pr\left(\mathrm{Q}\left(\mathrm{x}, \boldsymbol{\xi}\right) \geq 0\right) \leq \varepsilon \tag{19}$$

where $\varepsilon \geq 0$ is given. Set

$$\bar{q}(\varepsilon) = \begin{cases} \frac{-\ln\varepsilon + \sqrt{(\ln\varepsilon)^2 - 8\ln\varepsilon}}{4} & \varepsilon \leq \exp(-8) \\ 2 & else \end{cases} \tag{20}$$

and $\bar{Q}(\boldsymbol{x}, \boldsymbol{\xi}) = Q(\boldsymbol{x}, \boldsymbol{\xi}) + a_0(\mathbf{x})$. Then, the following hold:
(a) For each $\mathbf{x} \in \mathbb{R}^n$, $\boldsymbol{\xi} \in \mathbb{R}^m$, we have $\bar{Q}(\boldsymbol{x}, \boldsymbol{\xi})^2 = \mathbf{x}^T \mathbf{U}(\boldsymbol{\xi}) \mathbf{x}$
(b) Let $\mathbf{U} = \mathrm{E}\{\mathbf{U}(\boldsymbol{\xi})\} \succ 0$ and

$$c(\varepsilon) = \begin{cases} (\bar{q}(\varepsilon) - 1)^2 \exp(\frac{2\bar{q}(\varepsilon)}{\bar{q}(\varepsilon)-1}) & \bar{q}(\varepsilon) > 2 \\ \frac{1}{\sqrt{\varepsilon}} & \bar{q}(\varepsilon) = 2 \end{cases} \tag{21}$$

The second-order cone constraint

$$a_0(\mathbf{x}) \geq c(\varepsilon) \|\mathbf{U}^{\frac{1}{2}} \mathbf{x}\| \tag{22}$$

serves as a tractable approximation of the chance constraint.

Proof. By assumption, for each $\boldsymbol{\xi} \in \mathbb{R}^m$, the function $\bar{Q}(\mathbf{x}, \boldsymbol{\xi})$ is linear in $\mathbf{x} \in \mathbb{R}^n$. This implies that $\bar{Q}(\mathbf{x}, \boldsymbol{\xi})^2$ is a non-negative homogeneous quadratic polynomial in $\mathbf{x} \in \mathbb{R}^n$. This establishes (a). To prove (b), we invoke [11] (Theorem 5.10), which states that for any $q \geq 2$:

$$\mathrm{E}\left[\left|\bar{Q}(\mathbf{x}, \boldsymbol{\xi})\right|^q\right]^{\frac{1}{q}} \leq (q-1)^2 \mathrm{E}\left[\left|\bar{Q}(\mathbf{x}, \boldsymbol{\xi})\right|^2\right]^{\frac{1}{2}} \tag{23}$$

This, together with Markov's inequality and the result in (a), implies that for any $q \geq 2$

$$\begin{aligned} \Pr\left(\left|\bar{Q}(\mathbf{x}, \boldsymbol{\xi})\right| \geq t\right) &\leq t^{-q} \mathrm{E}\left[\left|\bar{Q}(\mathbf{x}, \boldsymbol{\xi})\right|^q\right] \\ &= [t^{-1}(q-1)^2 \|\mathbf{U}^{0.5}\mathbf{x}\|]^q \end{aligned} \tag{24}$$

By setting $q = \bar{q}(\varepsilon)$, we have $q \geq 2$. Moreover, whenever $t \geq c(\varepsilon) \|\mathbf{U}^{0.5}\mathbf{x}\|$, we have $\Pr\left(\left|\bar{Q}(\mathbf{x}, \boldsymbol{\xi})\right| \geq t\right) \leq \varepsilon$. It follows that

$$\begin{aligned} &\Pr\left(Q(\mathbf{x}, \boldsymbol{\xi}) \geq 0\right) = \Pr\left(\bar{Q}(\mathbf{x}, \boldsymbol{\xi}) \geq a_0(\mathbf{x})\right) \\ &\leq \Pr\left(\left|\bar{Q}(\mathbf{x}, \boldsymbol{\xi})\right| \geq a_0(\mathbf{x})\right) \\ &\leq \varepsilon \end{aligned} \tag{25}$$

whenever (22) holds. In particular, the second-order cone constraint (22) is a tractable safe approximation of (19). □

Here, we aim to apply the result of Theorem 1 to the non-convex problem (11). The auxiliary variable Z_j can be simplified as

$$\begin{aligned} Z_j = &\Delta\mathbf{f}^H \mathbf{Q}_{\mathbf{f}} \Delta\mathbf{f} + \Delta\mathbf{g}_j^H \mathbf{Q}_{\mathbf{f},\mathbf{g}} \Delta\mathbf{f} + \Delta\mathbf{f}^H \mathbf{Q}_{\mathbf{f},\mathbf{g}}^H \Delta\mathbf{g}_j \\ &+ \Delta\mathbf{g}_j^T \widehat{\mathbf{Q}}_{\mathbf{f},\mathbf{g}} \Delta\mathbf{f} + \Delta\mathbf{f}^H \widehat{\mathbf{Q}}_{\mathbf{f},\mathbf{g}}^H \Delta\mathbf{g}_j^* + \Delta\mathbf{g}_j^H \mathbf{Q}_{\mathbf{g}} \Delta\mathbf{g}_j \\ &+ \mathbf{c}_{\mathbf{g}}^H \Delta\mathbf{g}_j + \Delta\mathbf{g}_j^H \mathbf{c}_{\mathbf{g}} + \mathbf{c}_{\mathbf{f}}^H \Delta\mathbf{f} + \Delta\mathbf{f}^H \mathbf{c}_{\mathbf{f}} + d_j + \text{h.o.t.} \end{aligned} \tag{26}$$

By assuming the third and fourth order of the cross product of channel errors to be negligible, we can reformulate Z_j as

$$Z_j = \Delta\mathbf{H}_j^H \tilde{\mathbf{Q}} \Delta\mathbf{H}_j + 2\mathrm{Re}\left\{\tilde{\mathbf{r}}_j \Delta\mathbf{H}_j^H\right\} + d_j \tag{27}$$

where $\tilde{\mathbf{Q}} \triangleq \begin{bmatrix} \mathbf{R}_I & \mathbf{R}_{IQ} \\ \mathbf{R}_{IQ} & \mathbf{R}_Q \end{bmatrix}$, $\mathbf{R}_I \triangleq \begin{bmatrix} \mathbf{Q}_{\mathbf{f}} & \mathbf{Q}_{\mathbf{f},\mathbf{g}}^H \\ \mathbf{Q}_{\mathbf{f},\mathbf{g}} & \mathbf{Q}_{\mathbf{g}_j} \end{bmatrix}$, $\mathbf{R}_{I,Q} \triangleq$

$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \widehat{\mathbf{Q}}_{\mathbf{f},\mathbf{g}}^H & \mathbf{0} \end{bmatrix}$, $\mathbf{R}_Q \triangleq \mathbf{0}$, $\Delta\mathbf{H}_j^H \triangleq [\Delta\mathbf{h}_j^H \, \Delta\mathbf{h}_j^T]$, $\Delta\mathbf{h}_j^T \triangleq [\Delta\mathbf{f}^T \;\; \Delta\mathbf{g}_j^T]$, $\tilde{\mathbf{r}}_j \triangleq \frac{1}{2}[\mathbf{c}^H\mathbf{c}^T]$, $\mathbf{c}^H \triangleq [\mathbf{c_f}^H\mathbf{c_g}^H]$.

The matrices $\mathbf{Q_f}, \mathbf{Q_g}, \mathbf{Q_{f,g}}, \widehat{\mathbf{Q}}_{\mathbf{f},\mathbf{g}}, \mathbf{c_f}, \mathbf{c_g}$, and d_j are defined in Appendix 1. We have

$$\begin{aligned} \mathrm{E}\{|Z_j|^2\} &= \mathrm{E}\{Z_j Z_j^H\} \\ &= \mathrm{E}\{\Delta\mathbf{H}_j^H\tilde{\mathbf{Q}}\Delta\mathbf{H}_j\Delta\mathbf{H}_j^H\tilde{\mathbf{Q}}\Delta\mathbf{H}_j\} \\ &\quad + 4\mathrm{E}\{\Delta\mathbf{H}_j^H\mathbf{r}_j\mathbf{r}_j^H\Delta\mathbf{H}_j\} \end{aligned} \tag{28}$$

we have

$$2\mathrm{Tr}[\tilde{\mathbf{Q}}\boldsymbol{\Sigma}\tilde{\mathbf{Q}}\boldsymbol{\Sigma}] + (\mathrm{Tr}(\tilde{\mathbf{Q}}\boldsymbol{\Sigma}))^2 + 4\mathrm{Tr}(\mathbf{r}_j\boldsymbol{\Sigma}) \tag{29}$$

After some algebraic manipulations, it can be represented as

$$\mathrm{E}\{|Z_j|^2\} = \mathrm{vec}(\mathbf{X})^H\mathbf{U}_j\mathrm{vec}(\mathbf{X}) \tag{30}$$

The Affine term is also

$$\begin{aligned} d_j(\mathbf{X}) &= \mathrm{vec}(\mathbf{X})^H\mathrm{vec}(P_j\bar{\mathbf{h}}_j^j(\bar{\mathbf{h}}_j^j)^H) \\ &\quad - \gamma_j\sigma_{N_1}^2\mathrm{vec}(\mathbf{X})^H\mathrm{diag}(\mathrm{vec}(\mathbf{I}))\mathrm{vec}(\mathbf{g}_j\mathbf{g}_j^H) \\ &\quad - \gamma_j\mathrm{vec}(\mathbf{X})^H\mathrm{vec}(\sum_{p\neq j}^{M} P_p\bar{\mathbf{h}}_j^p(\bar{\mathbf{h}}_j^p)^H) \end{aligned} \tag{31}$$

Using Theorem 1, the following set of second-order cone constraints serves as a convex approximation of the chance constraint (15):

$$d_j(\mathbf{X}) \geq c(\varepsilon)\|\mathbf{U}_j^{\frac{1}{2}}\mathbf{X}\| \quad j \in \mathcal{J} \tag{32}$$

Hence, our optimization problem in (11), would be

$$\min_{\mathbf{X}} \quad \mathrm{Tr}(\mathbf{DX}) \tag{33a}$$

$$\text{s.t.} \;\; d_j(\mathbf{X}) \geq c(\varepsilon)\|\mathbf{U}_j^{\frac{1}{2}}\mathbf{X}\| \quad j \in \mathcal{J} \tag{33b}$$

$$\mathrm{Rank}(\mathbf{X}) = 1 \tag{33c}$$

The rank constraint in (33) is not convex. By dropping this constraint, and using Semi-Definite Program (SDP) relaxation the problem would be convex and can be solved efficiently using CVX software [12]. The relaxed SDP problem is

$$\min_{\mathbf{X}} \quad \mathrm{Tr}(\mathbf{DX}) \tag{34a}$$

$$\text{s.t.} \;\; d_j(\mathbf{X}) \geq c(\varepsilon)\|\mathbf{U}_j^{\frac{1}{2}}\mathbf{X}\| \quad j \in \mathcal{J} \tag{34b}$$

Note that, we want to extract the beamforming weights $\mathbf{w}$ from the solution of (34). The solution of (34) is a lower bound to the solution of (33), because the feasibility region of the non-convex problem is a subset of the relaxed problem. In general, the solution of the relax problem may have a general rank. Let $\mathbf{X}_{opt}$ denote the optimal solution of the relaxed problem (34). If the rank of $\mathbf{X}_{opt}$ is one, the principal eigenvector of $\mathbf{X}_{opt}$ is the optimal solution to the original problem. Otherwise, if the rank of the matrix $\mathbf{X}_{opt}$ is higher than one, an approximation technique is needed to obtain a rank one solution from the relaxed problem. This method is called randomization in literature (e.g. [3], [6]). The idea behind this technique is to generate a candidate set of beamforming vectors from the optimal solution of (34).

To design a randomization procedure for our problem, let $\mathbf{X}_{opt} = \mathbf{UVU}^H$ denotes eigen-decomposition of $\mathbf{X}_{opt}$. The candidate vector $\mathbf{w}_c$ can be chosen as $\mathbf{w}_c = \mathbf{UV}^{\frac{1}{2}}\boldsymbol{\xi}$, where $\boldsymbol{\xi} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, so that $E(\mathbf{w}_c\mathbf{w}_c^H) = \mathbf{X}_{opt}$. A feasible solution can be obtained by generating a sufficient number of realizations of $\mathbf{w}_c$, then easily choosing the best feasible solution. One way to generate the candidate solution of the problem (33) is to find the scaling factor $\sqrt{\alpha}$ to scale $\mathbf{w}_c$. This scaling factor is obtained by solving the following linear optimization problem:

$$\begin{aligned} \min_{\alpha} \quad & \alpha \\ \text{s.t.} \quad & \alpha a_j - c(\varepsilon)b_j \geq 0 \quad j \in \mathcal{J} \end{aligned} \tag{35}$$

where

$$\begin{aligned} a_j &= \mathrm{vec}(\mathbf{w}_c\mathbf{w}_c^H)^H\mathrm{vec}(P_j\bar{\mathbf{h}}_j^j(\bar{\mathbf{h}}_j^j)^H) \\ &\quad -\gamma_j\sigma_{N_1}^2\mathrm{vec}(\mathbf{w}_c\mathbf{w}_c^H)\mathrm{diag}(\mathrm{vec}(\mathbf{I}))\mathrm{vec}(\mathbf{g}_j\mathbf{g}_j^H) \\ &\quad -\gamma_j\mathrm{vec}(\mathbf{w}_c\mathbf{w}_c^H)^H\mathrm{vec}(\sum_{p\neq j}^{M} P_p\bar{\mathbf{h}}_j^p(\bar{\mathbf{h}}_j^p)^H) \\ b_j &= \|\mathbf{U}_j^{\frac{1}{2}}\mathbf{w}_c\mathbf{w}_c^H\| \end{aligned}$$

This random vector generation process is performed over a predefined value N_{max} times and, then, we select the appropriate vector which has the minimum objective value in the optimization problem (33). Algorithm 1 shows the different steps of randomization method.

Algorithm 1 Randomization Method

Input: $\mathbf{X_{opt}}$

Output: The rank one solution of $\mathbf{X_{opt}}$, $\mathbf{w}$

1: Compute SVD $\mathbf{X} = \mathbf{UVU}^H$ after solving (34)
2: **for** $i < N_{max}$ **do**
3: Generate a complex Gaussian random vector $\boldsymbol{\xi} \sim \mathcal{N}(0, \mathbf{I})$
4: Generate a candidate vector as $\mathbf{w}_c^i = \mathbf{UV}^{1/2}\boldsymbol{\xi}$
5: Solve the optimization problem (35) and obtain α^i
6: **if** the optimization problem is infeasible **then**
7: discard and return to step 4
8: **else**
9: Store $\mathbf{w}_c^i$ and the corresponding α^i and relays power transmission $\alpha^i\left(\mathbf{w}_c^i\mathbf{D}\mathbf{w}_c^i\right)$
10: **end if**
11: **end for**
12: Select $\alpha^{opt} = \alpha^i$ and $\mathbf{w}_c^{opt} = \mathbf{w}_c^i$ in which α^i and $\mathbf{w}_c^i$ correspond to the minimum relay transmit power
13: Output best candidate vector is $\mathbf{w}^{opt} = \sqrt{\alpha^{opt}}\mathbf{w}_c^{opt}$ and the minimum objective function $\alpha^{opt}\left(\mathbf{w}_c^{opt}\mathbf{D}\mathbf{w}_c^{opt}\right)$

IV. Channel Estimation

To solve the optimization problem (33), the means and variances of CSI for sources-relays and relays-destinations channels are required. The observations (measurements) at relays and jth destination at time step k is denoted as

$$\mathbf{z}_m(k) = \mathbf{h}\left(\mathbf{f}(k), \mathbf{g}_j(k)\right) + \mathbf{v}_m^{\mathbf{z}}(k) \tag{36}$$

where $\mathbf{z}_m(k) = [\mathbf{x}^T(k), y_j(k)]^T$, and $\mathbf{v}_m^{\mathbf{z}}(k) = [\mathbf{v}_m^{\mathbf{x}}(k)^T, \mathbf{v}_m^{y_j}(k)]^T$, $j \in \mathcal{J}$ is measurement error with variance $\mathbf{R}_k$, and $\mathbf{h}$ is a non-linear function of states, the footnote m denotes the measured values. Moreover, by using the Markov model, the sources-relays and the relays-jth destination channel coefficients are respectively formulated as

$$\mathbf{f}\,(k+1) = \alpha\mathbf{f}\,(k) + \mathbf{v}_s^{\mathbf{f}}(k) \tag{37}$$

$$\mathbf{g}_j\,(k+1) = \beta\mathbf{g}_j\,(k) + \mathbf{v}_s^{\mathbf{g}_j}(k) \qquad \forall j \in \mathcal{J} \tag{38}$$

where α, β are the temporal correlation coefficients. Note that α, β should be chosen less than one. Otherwise, the system will not be stable. The state variable at kth time step is $\boldsymbol{\xi}_j^T(k) =$

$\left[\mathbf{f}_1^T(k)\ \mathbf{f}_2^T(k)\ldots\mathbf{f}_M^T(k)\ \ \mathbf{g}_j^T(k)\right]^T$ which is a $(LM+L)\times 1$ vector. It can be written as

$$\boldsymbol{\xi}_j\left(k+1\right)=\mathbf{F}\boldsymbol{\xi}_j\left(k\right)+\mathbf{v}_s(k)\quad\forall j\in\mathcal{J}\tag{39}$$

where $\mathbf{F}=BD\left(\alpha\mathbf{I}_{LM},\beta\mathbf{I}_L\right)$, and $\mathbf{v}_s(k)=[\mathbf{v}_s^{\mathbf{f}}(k)^T,\mathbf{v}_s^{\mathbf{g}_j}(k)^T]^T$ is the state noise which has a Gussian distribution with variance $\mathbf{Q}=\sigma_s^2\mathbf{I}$. Here, the footnote s denotes the state equations. In (36), the observation equations are not linear functions of CSI, so we need to utilize the non-linear versions of KF (i.e. EKF, or CKF) for CSI estimation. Here, we have utilized the CKF for overcoming the non-linearity of the observation equations. The CKF method is an approximation of a Bayesian filter which has shown to have a better performance than Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) [10]. A single step of CKF algorithm is presented in Algorithm 2. The worst case complexity of CKF is given by $O((LM+L)^2)$ [13], Note that, the complexity order of channel estimation is nearly negligible compared to the complexity of solving optimization problem (34) and randomization procedure in Algorithm 1. A complete complexity analysis of SDP-based approaches can be found in [6].

Algorithm 2 Single Step of CKF Algorithm

Assume that we have $2n$ samples for state. For ease of notation, we denote $l=k-1$ and we show the time index by a subscript (e.g. $\boldsymbol{\xi}^j\left(k|k\right)\rightarrow\boldsymbol{\xi}^j_{k|k}$).

1: **for** $\leq i\leq 2n$ **do**

2: Evaluate the cubature points:
$$\boldsymbol{\chi}^i_{l|l}=\boldsymbol{\xi}_{l|l}+\sqrt{\mathbf{P}_{l|l}}\zeta_i$$

3: Evaluate the propagated cubature points:
$$\boldsymbol{\chi}^i_{k|l}=\mathbf{F}\boldsymbol{\chi}^i_{l|l}+\mathbf{v}_s$$

4: Estimate the predicted state and error covariance:
$$\boldsymbol{\xi}_{k|l}=\frac{1}{2n}\sum_{m=1}^{2n}\boldsymbol{\chi}^m_{k|l}$$
$$\mathbf{P}_{k|l}=\tfrac{1}{2n}\sum_{m=1}^{2n}((\boldsymbol{\chi}^m_{k|l}-\boldsymbol{\xi}_{k|l})(\boldsymbol{\chi}^m_{k|l}-\boldsymbol{\xi}_{k|l})^H)$$

5: Form the cubature points:
$$\boldsymbol{\chi}^i_{k|l}=\sqrt{\mathbf{P}_{k|l}}\mathbf{F}\boldsymbol{\zeta_i}+\boldsymbol{\xi}_{k|l}$$

6: Propagate cubature points:
$$\mathbf{Z}^i_{k|l}=\boldsymbol{h}(\boldsymbol{\chi}^i_{k|l})$$

7: **end for**

8: Estimate predicted measurement:
$$\mathbf{z}\left(k|l\right)=\frac{1}{2n}\sum_{i=1}^{2n}\mathbf{Z}^i{}_{k|l}$$

9: Estimate innovations covariance matrix:
$$\mathbf{S}_k=\tfrac{1}{2n}\sum_{i=1}^{2n}(\mathbf{Z}^i{}_{k|l}-\mathbf{z}_{k|l})(\mathbf{Z}^i{}_{k|l}-\mathbf{z}_{k|l})^H+\mathbf{R}_k$$

10: Estimate the cross-covariance matrix:
$$\mathbf{C}_k=\tfrac{1}{2n}\sum_{i=1}^{2n}(\boldsymbol{\chi}^i_{k|l}-\boldsymbol{\xi}_{k|l})(\mathbf{Z}^i{}_{k|l}-\mathbf{z}_{k|l})^H$$

11: Estimate the updated state:
$$\boldsymbol{\xi}_{k|k}=\boldsymbol{\xi}_{k|l}+\mathbf{C}_k\mathbf{S}_k^{-1}(\mathbf{z}_m-\mathbf{z}_{k|l})$$

12: Estimate the corresponding error covariance:
$$\mathbf{P}_{k|k}=\mathbf{P}_{k|l}-\boldsymbol{\Omega}_k\mathbf{S}_k\boldsymbol{\Omega}_k^T$$

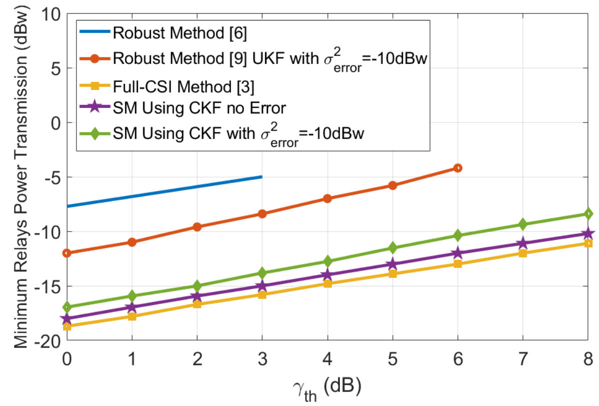

Figure 2: Minimum relays transmission power versus the SINR threshold

V. SIMULATION RESULTS

In this section, we present numerical results to compare the performance of the proposed robust beamforming design with other approaches in terms of power consumption. Also, the convergence of CKF method for estimating CSI is discussed and compared with CRLB.

A. Methodology

In this section, we use the Monte-Carlo simulations to generate the results. We consider a network with two source-destination pairs ($M=2$) and $L=20$ relays. We have the following assumptions: the noise variance is -10 dBw, each channel has 0 dBw power and that the error variance of each channel is -20 dBw and $\alpha=\beta=0.99$ and the measurements error matrix is $\mathbf{R}=\sigma^2_{error}\mathbf{I}=0.1\mathbf{I}$. Unless another value is specified, the SINR threshold is assumed to be 5 dB. The following methods are compared:

- The robust method in [6];
- The non-robust method in [3] denoted by "Full-CSI";
- The robust method in [9];
- Stochastic method in (33) which is denoted by "SM".

B. Results and Discussion

We first study the average transmit power of the relays versus the SINR threshold. To this end, we compare different robust and non-robust methods in terms of power transmission. In Fig. 2, we plot the minimum power required versus the target SINR. It can be seen that, as the measurement error covariance increases, more power is required. Also, the proposed method outperforms the methods in [9] and [6] in terms of power consumption and feasibility region. Furthermore, this figure shows that the performance of the proposed real-time channel estimation using CKF for our beamforming method is comparable with that of the Full-CSI method in [3].

Next, in Fig.3, we study the convergence behavior of the proposed algorithm. For this purpose, we consider the first 100 iterations of CKF. The root mean square error (RMSE) for the proposed CKF method is compared with CRLB and depicted in Fig.3. We can see that the CKF method almost achieves the CRLB after some iterations which shows the efficiency of the proposed method. Fig.4 evaluates the performance of proposed

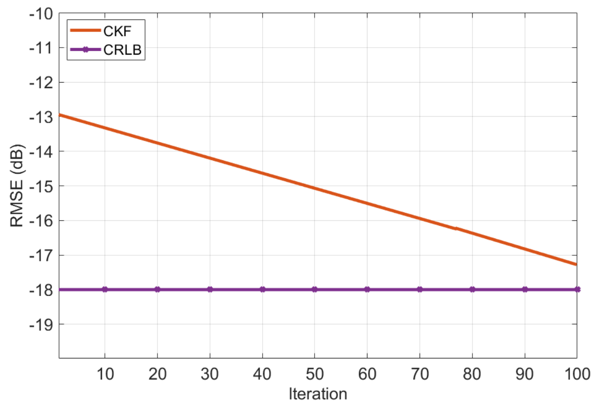

Figure 3: RMSE versus the number iterations

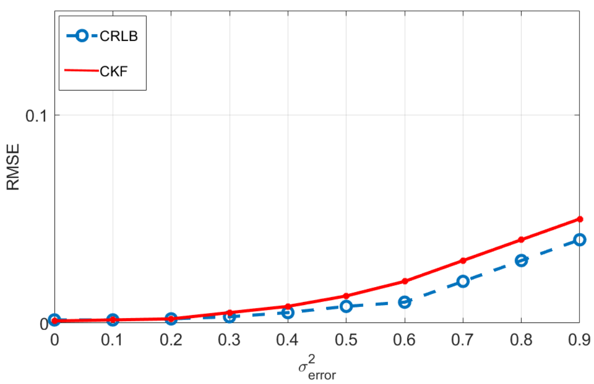

Figure 4: RMSE versus the measurement error variance

CKF method versus measurement error level. As shown in this figure, the proposed CKF method achieves the CRLB for low and moderate error levels. Moreover, the performance of CKF method in higher levels of error is also close to CRLB.

VI. Conclusion

In this paper, we proposed an adaptive CKF method for estimation and tracking the sources-relays and relays-destinations channels for a peer-to-peer communications system using a network of relays. The estimated CSI were used for solving a stochastic robust beamforming optimization problem to minimize the total power transmission of the relays subject to probabilistic SINR constraint at each of the destination nodes. Simulation results showed our stochastic beamforming method have a better performance than the existing methods. Moreover, the CSI estimation methods were shown to converge to CRLB for low and moderate error levels which confirms the efficiency of our estimation.

Appendix1: Definition of matrices

Assume that $\bar{\mathbf{F}}_p = \operatorname{diag}\left(\bar{\mathbf{f}}_p\right)$, $\bar{\mathbf{G}}_j = \operatorname{diag}\left(\bar{\mathbf{g}}_j\right)$, $\bar{\mathbf{H}}^p_j = \operatorname{diag}\left(\bar{\mathbf{h}}^p_j\right)$, $\mathbf{X} = \mathbf{w}\mathbf{w}^H$, $\mathcal{W} = \mathbf{W}\mathbf{W}^H$, $\mathcal{H}^p_j = \bar{\mathbf{h}}^p_j \bar{\mathbf{h}}^{p\,H}_j$, $j, p \in \mathcal{J}$ and $j \neq p$. The matrices $\mathbf{Q}_\mathbf{f}, \mathbf{Q}_\mathbf{g}, \mathbf{Q}_{\mathbf{f},\mathbf{g}}, \widehat{\mathbf{Q}}_{\mathbf{f},\mathbf{g}}, \mathbf{c}_\mathbf{f}, \mathbf{c}_\mathbf{g}$, and d_j which have been derived in [6], are defined as:

$$\mathbf{Q}_\mathbf{f} = \operatorname{BD}([\mathbf{Q}_\mathbf{f}]_{11}, [\mathbf{Q}_\mathbf{f}]_{22}, \ldots, [\mathbf{Q}_\mathbf{f}]_{jj}, \ldots, [\mathbf{Q}_\mathbf{f}]_{MM}) \tag{40}$$

$$[\mathbf{Q}_\mathbf{f}]_{jj} = P_j(\bar{\mathbf{G}}_j \mathbf{X} \bar{\mathbf{G}}^H_j) - \gamma_j \sum_{p \neq j}^{M} P_p \bar{\mathbf{G}}_p \mathbf{X} \bar{\mathbf{G}}^H_p$$

$$\mathbf{Q}_{\mathbf{g}_j} = \frac{1}{d}\operatorname{BD}(\mathbf{Q}'_{\mathbf{g}_j}, \mathbf{Q}'_{\mathbf{g}_j}, \ldots, \mathbf{Q}'_{\mathbf{g}_j}) \tag{41}$$

$$\mathbf{Q}'_{\mathbf{g}_j} = P_j \bar{\mathbf{F}}^H_j \mathbf{X} \bar{\mathbf{F}}_j - \gamma_j \sigma^2_{N_1} \mathcal{W} - \gamma_j \sum_{p \neq j}^{M} P_p \bar{\mathbf{F}}^H_p \mathbf{X} \bar{\mathbf{F}}_p$$

$$\mathbf{Q}_{\mathbf{f},\mathbf{g}} = \operatorname{BD}\left([\mathbf{Q}_{\mathbf{f},\mathbf{g}}]_{11}, [\mathbf{Q}_{\mathbf{f},\mathbf{g}}]_{22}, \ldots, [\mathbf{Q}_{\mathbf{f},\mathbf{g}}]_{MM}\right) \tag{42}$$

$$[\mathbf{Q}_{\mathbf{f},g}]_{jj} = P_j \bar{\mathbf{F}}^H_j \mathbf{X} \bar{\mathbf{G}}_j - \gamma_j \sum_{p \neq j}^{M} P_p \bar{\mathbf{F}}^H_p \mathbf{X} \bar{\mathbf{G}}_p$$

$$\widehat{\mathbf{Q}}^H_{\mathbf{f},\mathbf{g}} = \operatorname{BD}([\widehat{\mathbf{Q}}^H_{\mathbf{f},\mathbf{g}}]_{11}, [\widehat{\mathbf{Q}}^H_{\mathbf{f},\mathbf{g}}]_{22}, \ldots, [\widehat{\mathbf{Q}}^H_{\mathbf{f},\mathbf{g}}]_{MM}) \tag{43}$$

$$[\widehat{\mathbf{Q}}^H_{\mathbf{f},\mathbf{g}}]_{jj} = P_j \operatorname{diag}(\mathbf{w})^H \bar{\mathbf{H}}^j_j \operatorname{diag}(\mathbf{w}) - \gamma_j \sum_{p \neq j}^{M} P_p \mathbf{W}^H \bar{\mathbf{H}}^p_p \mathbf{W}$$

$$\mathbf{c}_\mathbf{f} = \operatorname{BD}[[\mathbf{c}_\mathbf{f}]^{\mathrm{H}}_{11} \ldots [\mathbf{c}_\mathbf{f}]^{\mathrm{H}}_{\mathrm{jj}} \ldots [\mathbf{c}_\mathbf{f}]^{\mathrm{H}}_{\mathrm{MM}}] \tag{44}$$

$$[\mathbf{c}_\mathbf{f}]^H_{jj} = P_j (\bar{\mathbf{h}}^j_j)^H \mathbf{X} \bar{\mathbf{G}}_j - \gamma_i \sum_{p \neq j}^{M} P_j \left(\bar{\mathbf{h}}^p_j\right)^H \mathbf{X} \bar{\mathbf{G}}_j$$

$$\mathbf{c}_\mathbf{g} = \frac{1}{d}(P_j \mathbf{F}^H_j \mathbf{X} \bar{\mathbf{h}}^j_j - \gamma_j \sum_{p \neq j}^{M} P_p \mathbf{F}^H_p \mathbf{X} \bar{\mathbf{h}}^p_j - \gamma_j \sigma^2_{N_1} \mathcal{W} \bar{\mathbf{g}}_j) \tag{45}$$

$$d_j = \mathbf{w}^H P_j \mathcal{H}^j_j \mathbf{w} - \gamma_j \sigma^2_{N_1} \bar{\mathbf{g}}^H_j \mathcal{W} \bar{\mathbf{g}}_j - \gamma_j \mathbf{w}^H \sum_{p \neq j}^{M} P_p \mathcal{H}^p_j \mathbf{w} \tag{46}$$

References

[1] Y. Xiao, X. Jin, Y. Shen, and Q. Guan, “Joint relay selection and adaptive modulation and coding for wireless cooperative communications,” *IEEE Sens. J.*, 2021 (Early-Access).

[2] A. K. Shukla, V. Singh, P. K. Upadhyay, A. Kumar, and J. M. Moualeu, “Performance analysis of energy harvesting-assisted overlay cognitive NOMA systems with incremental relaying,” *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1558–1576, Jun. 2021.

[3] S. Fazeli-Dehkordy, S. Shahbazpanahi, and S. Gazor, “Multiple peer-to-peer communications using a network of relays,” *IEEE Trans. Signal. Process.*, vol. 57, no. 8, pp. 3053–3062, 2009.

[4] F. Gao and Y. Chen, “Performance analysis for an opportunistic DF based WPC system with interferers over log-normal fading channels,” in *13th ICCSN*, Jun. 2021, pp. 6–10.

[5] A. R. Heidarpour, M. Ardakani, and C. Tellambura, “Network-coded cooperation with outdated CSI,” *IEEE Commun. Lett.*, vol. 22, no. 8, pp. 1720–1723, Jun. 2018.

[6] M. A. M. Sadr, B. Mahboobi, S. Mehrizi, M. A. Attari, and M. Ardebilipour, “Stochastic robust collaborative beamforming: Non-regenerative relay,” *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 947–958, March 2016.

[7] M. Maleki Sadr, M. Ahmadian and B. Mahboobi, “Low-complexity robust relay optimisation for multiple peer-to-peer beamforming: a safe tractable approximation approach,” *IET Commun.*, vol. 9, no. 16, pp. 1968–1979, Nov. 2015.

[8] B. Chalise and L. Vandendorpe, “MIMO Relay Design for Multipoint-to-Multipoint Communications With Imperfect Channel State Information,” *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2785–2796, July 2009.

[9] M. A. Maleki Sadr, M. Ahmadian-Attari, and R. Amiri, “Real-time cooperative adaptive robust relay beamforming based on kalman filtering channel estimation,” *IEEE Trans. Wireless. Commun.*, vol. 18, no. 12, pp. 5600–5612, Sep. 2019.

[10] I. Arasaratnam, “Cubature Kalman filtering theory & applications,” Ph.D. dissertation, 2009.

[11] S. Janson *et al.*, *Gaussian hilbert spaces*. Cambridge university press, 1997, no. 129.

[12] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1,” http://cvxr.com/cvx, Mar. 2014.

[13] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.