Distributed Filter-and-Forward Beamforming for Two-Way Relaying Networks under Channel Uncertainties

Chen Luo†, Chengwen Xing†, Zesong Fei†, Shaodan Ma‡ and Jingming Kuang†
† RCDCT, School of Information and Electronics, Beijing Institute of Technology, Beijing, China.
Email: {luochen, feizesong}@bit.edu.cn, xingchengwen@gmail.com
‡Department of Electrical and Computer Engineering, University of Macau, Macau, China.
Email: {shaodanma}@umac.mo

***Abstract*—In this paper, we consider robust distributed filter-and-forward beamforming design for two-way relaying networks over the frequency selective fading channels, in which two users exchange information with the aid of multiple single-antenna relay nodes. In contrast to the prior works in which the channel state information (CSI) is available, in our work CSI for all relevant links between the users and relay nodes is not perfectly known with stochastic channel errors. Under stochastic channel errors, a robust beamforming design aiming at maximizing the total system signal-to-interference-plus-noise-ratio (SINR) under individual relay power constraints is proposed. Simulation results show that the proposed robust beamformer reduces the sensitivity of the two-way relaying systems to channel estimation errors, and performs better than the algorithm using estimated channels only.**

***Index Terms*—Two-way relaying, filter-and-forward, robust beamforming, stochastic channel uncertainty.**

I. Introduction

Recently, two-way relaying communication has been widely investigated, due to its capability to further enhance radio resource efficiency for information exchange compared with its counterpart one-way relaying [1]- [4]. In two-way relaying systems, the relays receive signals from two users and broadcast them to both terminals. Then each user can obtain its desired signals through self-information cancelation.

In most literatures, beamforming designs are considered for flat-fading channels, by assuming orthogonal frequency-division multiplexing (OFDM) technique [5] (i.e., a multi-carrier transmission) is applied to convert the frequency selective fading channels into a number of flat-fading channels. However, some systems still prefer the single-carrier transmission because of the legacy or disadvantages of OFDM such as a high peak-to-average power ratio. Taking GSM/EDGE mobile communication system as an example, the OFDM technique may be prohibitive due to high cost and power consumption of the linear power amplifiers in the terminals. Therefore, beamforming design under frequency selective fading channels is necessary and a filter-and-forward (FF) beamforming scheme for relaying systems over frequency selective channels has been recently discussed in [6]- [8].

Most of the existing beamforming designs have assumed that the channel state information (CSI) is perfectly known at the transceiver. Unfortunately, such an assumption generally does not hold in practice, because perfect channel estimation is impossible with a limited number of training data under noisy environment. Furthermore, quantization will also cause mismatches between the CSI at transceivers with limited feedback. To be practical, beamforming design for a two-way relaying system, in which all relevant channels between two users and relay nodes are not perfectly known, is investigated in this paper. The channel uncertainty is assumed following the widely used stochastic error (SE) model [9], in which the second-order statistics of the channel coefficients are available. With the objective to maximize the system signal-to-interference-plus-noise-ratio (SINR) subject to the individual relay transmitted power constraints [10], an efficient beamforming algorithm taking the channel uncertainties into account is proposed. The optimal solutions are then attained by solving a series of semi-definite programming (SDP) problems after semi-definite relaxation (SDR). Simulation results demonstrate that the proposed beamformer reduces the sensitivity of the two-way relaying systems to channel estimation errors, and outperforms the algorithm using estimated channels only.

The rest of the paper is organized as follows. In Section II, the system model of a distributed filter-and-forward two-way relaying network is introduced. The optimization problem for maximizing the total SINR and its optimal solution are described in the Section III. In Section IV, simulation results are provided to corroborate our proposed design. Finally, conclusions are drawn in Section V.

Notation: In this paper, the lowercase and uppercase bold letters are used to denote vectors and matrices, respectively. Transpose and Hermitian transpose of a matrix are denoted by $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$. $\mathrm{Tr}\{\cdot\}$ and $\mathrm{rank}(\cdot)$ are used to represent the trace and the rank of a matrix, respectively. A diagonal matrix $\mathbf{X}$ with the elements of $x_1, \ldots, x_N$ as diagonal entries is written

This work was supported by Ericsson, the Sino-Swedish IMT-Advanced and Beyond Cooperative Program under grant No. 2008DFA11780 and National Natural Science Foundation of China under Grant No. 61101130.

978-1-4673-0990-5/12/$31.00 ©2012 IEEE

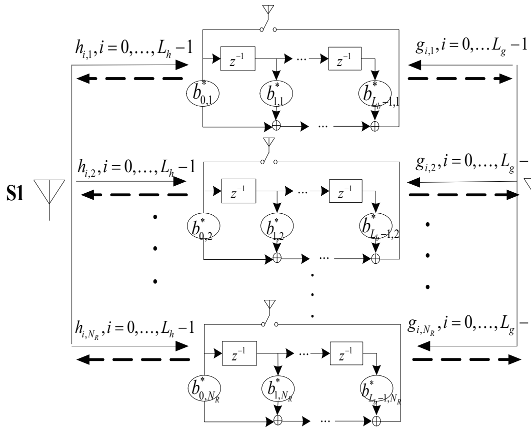

Fig. 1. The system model of filter-based-forward two-way relaying network.

as $\mathbf{X} = \mathrm{diag}\{x_1, \dots, x_N\}$, and the diagonal entries vector $\mathbf{x} = [x_1, \dots, x_N]^{\mathrm{T}}$ is expressed as $\mathbf{x} = \mathrm{diag}\{\mathbf{X}\}$. The symbol $\mathbb{E}\{\cdot\}$ represents the expectation operation. The ℓ_2 (Euclidean) norm is denoted by $\|\cdot\|$. I_M denotes the $M \times M$ identity matrix. For two Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} \succeq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semidefinite.

II. System Model

In this paper, we consider a distributed two-way relaying network under the frequency selective channels, as shown in Fig.1. All the network nodes (including two user terminals S_1, S_2 and N_R relay nodes) are equipped with a single-antenna. Due to deep fading, there is no direct link between the two users, and the information exchange between them is assisted by N_R relays. In the two-way relaying system, the information exchange is completed in only two transmission slots. In the first slots, both user terminals simultaneously transmit their data to the relays, and in the second slot, the relays linearly process the received signals and broadcast them to the two users.

In Fig.1, the discrete-time channel impulse responses (CIRs) between the user terminal S_1 and relay node R_j is denoted by $h_{i,j},\ i = 0, \dots, L_h - 1$ and that between the user terminal S_2 and relay node R_j is $g_{i,j},\ i = 0, \dots, L_g - 1$. L_h, L_g denote the lengths of the $S_1 - R_j$ and $S_2 - R_j$ CIRs, respectively. Furthermore, the channels between user terminals and the relay nodes are reciprocal. Because the terminals and relays generally obtain the CSI via channel estimation using the training sequences, the estimation errors are inevitable. When the estimation error is taken into account, we have

$$\begin{aligned} h_{i,j} &= \bar{h}_{i,j} + \Delta_{i,j}, \\ g_{i,j} &= \bar{g}_{i,j} + \Lambda_{i,j}, \end{aligned} \tag{1}$$

where $\bar{h}_{i,j}$ and $\bar{g}_{i,j}$ are the estimated CSI. Moreover, $\Delta_{i,j}$ and $\Lambda_{i,j}$ are the corresponding channel estimation errors with means and variances

$$\begin{aligned} \mathrm{E}\{\Delta_{\mathrm{i,j}}\} = 0,\ \mathrm{Var}\{\Delta_{\mathrm{i,j}}\} = \delta, \\ \mathrm{E}\{\Lambda_{\mathrm{i,j}}\} = 0,\ \mathrm{Var}\{\Lambda_{\mathrm{i,j}}\} = \lambda. \end{aligned} \tag{2}$$

The signal received at the relay nodes is read as

$$\mathbf{r}[k] = \sum_{l=0}^{L_h-1} \mathbf{h}_l x_1[k-l] + \sum_{l=0}^{L_g-1} \mathbf{g}_l x_2[k-l] + \mathbf{n}_r[k], \tag{3}$$

where the vectors $\mathbf{h}_l$ and $\mathbf{g}_l$ equal to

$$\begin{aligned} \mathbf{h}_l &= \bar{\mathbf{h}}_l + \boldsymbol{\Delta}_l = \left[\bar{h}_{l,1}, \dots, \bar{h}_{l,N_R}\right]^{\mathrm{T}} + [\Delta_{l,1}, \dots, \Delta_{l,N_R}]^{\mathrm{T}}, \\ \mathbf{g}_l &= \bar{\mathbf{g}}_l + \boldsymbol{\Lambda}_l = [\bar{g}_{l,1}, \dots, \bar{g}_{l,N_R}]^{\mathrm{T}} + [\Lambda_{l,1}, \dots, \Lambda_{l,N_R}]^{\mathrm{T}}. \end{aligned}$$

In (3), $x_1[k]$ and $x_2[k]$ are the source information from S_1 and S_2, respectively, and $\mathbf{n}_r[k]$ denotes the additive white Gaussian noise (AWGN) at the relay nodes with variance $\sigma^2_{n_r}$.

Introducing compact matrix variables

$$\begin{aligned} \mathbf{H} &\triangleq \bar{\mathbf{H}} + \boldsymbol{\Delta} = [\mathbf{h}_0, \dots, \mathbf{h}_{L_h-1}] + [\boldsymbol{\Delta}_0, \dots, \boldsymbol{\Delta}_{L_h-1}], \\ \mathbf{G} &\triangleq \bar{\mathbf{G}} + \boldsymbol{\Lambda} = [\mathbf{g}_0, \dots, \mathbf{g}_{L_g-1}] + [\boldsymbol{\Lambda}_0, \dots, \boldsymbol{\Lambda}_{L_g-1}], \\ \mathbf{X}_1[k] &\triangleq [x_1[k], \dots, x_1[k - L_h + 1]]^{\mathrm{T}}, \\ \mathbf{X}_2[k] &\triangleq [x_2[k], \dots, x_1[k - L_g + 1]]^{\mathrm{T}}, \end{aligned} \tag{4}$$

the received signal in (3) is reformulated as

$$\mathbf{r}[k] = \mathbf{H}\mathbf{X}_1[k] + \mathbf{G}\mathbf{X}_2[k] + \mathbf{n}_r[k]. \tag{5}$$

In order to compensate the frequency selective fading, at the relay nodes, the beamformers with the finite impulse response (FIR) filters are adopted. It is assumed that the length of the FIR is L_b, and when $L_b = 1$, they are the amplify-and-forward (AF) beamformers. The transmitted signals from the relays, $\mathbf{t}[k] = [t_1[k], \dots, t_{N_R}[k]]$, can be written as

$$\mathbf{t}[k] = \sum_{l=0}^{L_b-1} \mathbf{B}_l^{\mathrm{H}} \mathbf{r}[k-l], \tag{6}$$

where $\mathbf{B}_l$ is a diagonal matrix as $\mathbf{B}_l = \mathrm{diag}\{b_{l,1}, \dots, b_{l,N_R}\}$. $b_{l,j}$ is the FIR response corresponding to the l_{th} tap of the j_{th} FF beamformer. Using (3) and convolution expression, the transmitted relay signal in (6) can be rewritten as

$$\mathbf{t}[k] = \mathbf{B}\tilde{\mathbf{H}}\tilde{\mathbf{X}}_1[k] + \mathbf{B}\tilde{\mathbf{G}}\tilde{\mathbf{X}}_2[k] + \mathbf{B}\tilde{\mathbf{n}}_r[k], \tag{7}$$

where the components in (7) are defined as

$$\begin{aligned} \mathbf{B} &\triangleq [\mathbf{B}_0, \dots, \mathbf{B}_{L_b-1}]^{\mathrm{T}}, \\ \tilde{\mathbf{H}} &\triangleq \tilde{\bar{\mathbf{H}}} + \boldsymbol{\Delta}_{\tilde{H}} \\ &= [\bar{\mathbf{H}}_0^{\mathrm{T}}, \dots, \bar{\mathbf{H}}_{L_h-1}^{\mathrm{T}}]^{\mathrm{T}} + [\boldsymbol{\Delta}_{\mathbf{H}_0}^{\mathrm{T}}, \dots, \boldsymbol{\Delta}_{\mathbf{H}_{L_h-1}}^{\mathrm{T}}]^{\mathrm{T}}, \\ \bar{\mathbf{H}}_l &\triangleq [\mathbf{0}_{R\times l}, \bar{\mathbf{H}}, \mathbf{0}_{R\times(L_b-1-l)}], \\ \boldsymbol{\Delta}_{\mathbf{H}_l} &\triangleq [\mathbf{0}_{R\times l}, \boldsymbol{\Delta}, \mathbf{0}_{R\times(L_b-1-l)}], \\ \tilde{\mathbf{G}} &\triangleq \tilde{\bar{\mathbf{G}}} + \boldsymbol{\Lambda}_{\tilde{G}} \\ &= [\bar{\mathbf{G}}_0^{\mathrm{T}}, \dots, \bar{\mathbf{G}}_{L_b-1}^{\mathrm{T}}]^{\mathrm{T}} + [\boldsymbol{\Lambda}_{\mathbf{G}_0}^{\mathrm{T}}, \dots, \boldsymbol{\Lambda}_{\mathbf{G}_{L_b-1}}^{\mathrm{T}}]^{\mathrm{T}}, \\ \bar{\mathbf{G}}_l &\triangleq [\mathbf{0}_{R\times l}, \bar{\mathbf{G}}, \mathbf{0}_{R\times(L_b-1-l)}], \\ \boldsymbol{\Lambda}_{\mathbf{G}_l} &\triangleq [\mathbf{0}_{R\times l}, \boldsymbol{\Lambda}, \mathbf{0}_{R\times(L_b-1-l)}], \\ \tilde{\mathbf{X}}_1[k] &\triangleq [x_1[k], \dots, x_1[k - L_h - L_b + 2]]^{\mathrm{T}}, \\ \tilde{\mathbf{X}}_2[k] &\triangleq [x_2[k], \dots, x_2[k - L_g - L_b + 2]]^{\mathrm{T}}, \\ \tilde{\mathbf{n}}_r[k] &\triangleq [n_r[k], \dots, n_r[k - L_b + 1]]^{\mathrm{T}}. \end{aligned} \tag{8}$$

At the user terminals, the received signals $\mathbf{y}_i[k], i = 1,2$ are given by

$$
\begin{aligned}
\mathbf{y}_1[k] =& \sum_{l=0}^{L_h-1} \mathbf{h}_l^{\mathrm{T}} \mathbf{B} \left(\tilde{\mathbf{H}} \tilde{\mathbf{X}}_1[k-l] + \tilde{\mathbf{G}} \tilde{\mathbf{X}}_2[k-l] \right) \\
&+ \sum_{l=0}^{L_h-1} \mathbf{h}_l^{\mathrm{T}} \mathbf{B} \tilde{\mathbf{n}}_r[k-l] + \mathbf{n}_1[k], \\
\mathbf{y}_2[k] =& \sum_{l=0}^{L_g-1} \mathbf{g}_l^{\mathrm{T}} \mathbf{B} \left(\tilde{\mathbf{H}} \tilde{\mathbf{X}}_1[k-l] + \tilde{\mathbf{G}} \tilde{\mathbf{X}}_2[k-l] \right) \\
&+ \sum_{l=0}^{L_g-1} \mathbf{g}_l^{\mathrm{T}} \mathbf{B} \tilde{\mathbf{n}}_r[k-l] + \mathbf{n}_2[k],
\end{aligned} \tag{9}
$$

where $\mathbf{n}_1[k]$ and $\mathbf{n}_2[k]$ are AWGN noises at the destinations S_1 and S_2 with variances $\sigma_{n_1}^2$ and $\sigma_{n_2}^2$, respectively.

Since the matrix $\mathbf{B}_l$ is diagonal and according to the properties of the Kronecker matrix product, we have:

$$
\begin{aligned}
\mathbf{h}_l^{\mathrm{T}} \mathbf{B}^{\mathrm{H}} &= \mathbf{b}^{\mathrm{H}} \left(\mathbf{I}_{L_b} \otimes \bar{\boldsymbol{\mathcal{H}}}_l^{\mathrm{T}} \right) + \mathbf{b}^{\mathrm{H}} \left(\mathbf{I}_{L_b} \otimes \boldsymbol{\Delta}_{\boldsymbol{\mathcal{H}}_l}^{\mathrm{T}} \right), \\
\mathbf{g}_l^{\mathrm{T}} \mathbf{B}^{\mathrm{H}} &= \mathbf{b}^{\mathrm{H}} \left(\mathbf{I}_{L_b} \otimes \bar{\boldsymbol{\mathcal{G}}}_l^{\mathrm{T}} \right) + \mathbf{b}^{\mathrm{H}} \left(\mathbf{I}_{L_b} \otimes \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{G}}_l}^{\mathrm{T}} \right),
\end{aligned} \tag{10}
$$

which are under the definitions that

$$
\begin{aligned}
\mathbf{b}_l &\triangleq \mathrm{diag}\{\mathbf{B}_l\}, \quad \mathbf{b} = \left[\mathbf{b}_0^{\mathrm{T}}, \ldots, \mathbf{b}_{L_b-1}^{\mathrm{T}} \right]^{\mathrm{T}}, \\
\bar{\boldsymbol{\mathcal{H}}}_l &\triangleq \mathrm{diag}\{\bar{\mathbf{h}}_l\}, \quad \boldsymbol{\Delta}_{\boldsymbol{\mathcal{H}}_l} \triangleq \mathrm{diag}\{\boldsymbol{\Delta}_l\}, \\
\bar{\boldsymbol{\mathcal{G}}}_l &\triangleq \mathrm{diag}\{\bar{\mathbf{g}}_l\}, \quad \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{G}}_l} \triangleq \mathrm{diag}\{\boldsymbol{\Lambda}_l\}.
\end{aligned}
$$

In order to have a compact form for (9), defining

$$
\begin{aligned}
\breve{\mathbf{X}}_1[k] &\triangleq \left[x_1[k], \ldots, x_1[k - L_h - L_b - L_g + 3] \right]^{\mathrm{T}}, \\
\breve{\mathbf{X}}_2[k] &\triangleq \left[x_2[k], \ldots, x_2[k - L_g - L_b - L_h + 3] \right]^{\mathrm{T}}, \\
\breve{\mathbf{n}}_r[k] &\triangleq \left[n_r[k], \ldots, n_r[k - L_h - L_b + 2] \right]^{\mathrm{T}}, \\
\hat{\mathbf{n}}_r[k] &\triangleq \left[n_r[k], \ldots, n_r[k - L_g - L_b + 2] \right]^{\mathrm{T}}, \\
\breve{\boldsymbol{\mathcal{H}}} &\triangleq \bar{\breve{\boldsymbol{\mathcal{H}}}} + \boldsymbol{\Delta}_{\breve{\boldsymbol{\mathcal{H}}}} \\
&= \left[\mathbf{I}_{L_b} \otimes \bar{\boldsymbol{\mathcal{H}}}_0^{\mathrm{T}}, \ldots, \mathbf{I}_{L_b} \otimes \bar{\boldsymbol{\mathcal{H}}}_{L_h-1}^{\mathrm{T}} \right] \\
&\quad + \left[\mathbf{I}_{L_b} \otimes \boldsymbol{\Delta}_{\boldsymbol{\mathcal{H}}_0}^{\mathrm{T}}, \ldots, \mathbf{I}_{L_b} \otimes \boldsymbol{\Delta}_{\boldsymbol{\mathcal{H}}_{L_h-1}}^{\mathrm{T}} \right], \\
\breve{\mathbf{H}} &\triangleq \bar{\breve{\mathbf{H}}} + \boldsymbol{\Delta}_{\breve{\mathbf{H}}} \\
&= \left[\bar{\tilde{\mathbf{H}}}_0^{\mathrm{T}}, \ldots, \bar{\tilde{\mathbf{H}}}_{L_h-1}^{\mathrm{T}} \right]^{\mathrm{T}} + \left[\boldsymbol{\Delta}_{\tilde{\mathbf{H}}_0}^{\mathrm{T}}, \ldots, \boldsymbol{\Delta}_{\tilde{\mathbf{H}}_{L_h-1}}^{\mathrm{T}} \right]^{\mathrm{T}}, \\
\breve{\mathbf{G}} &\triangleq \bar{\breve{\mathbf{G}}} + \boldsymbol{\Lambda}_{\breve{\mathbf{G}}} \\
&= \left[\bar{\tilde{\mathbf{G}}}_0^{\mathrm{T}}, \ldots, \bar{\tilde{\mathbf{G}}}_{L_h-1}^{\mathrm{T}} \right]^{\mathrm{T}} + \left[\boldsymbol{\Lambda}_{\tilde{\mathbf{G}}_0}^{\mathrm{T}}, \ldots, \boldsymbol{\Lambda}_{\tilde{\mathbf{G}}_{L_h-1}}^{\mathrm{T}} \right]^{\mathrm{T}}, \\
\breve{\mathbf{I}} &\triangleq \left[\tilde{\mathbf{I}}_0^{\mathrm{T}}, \ldots, \tilde{\mathbf{I}}_{L_h-1}^{\mathrm{T}} \right]^{\mathrm{T}}, \\
\hat{\boldsymbol{\mathcal{G}}} &\triangleq \bar{\hat{\boldsymbol{\mathcal{G}}}} + \boldsymbol{\Lambda}_{\hat{\boldsymbol{\mathcal{G}}}} \\
&= \left[\mathbf{I}_{L_b} \otimes \bar{\boldsymbol{\mathcal{G}}}_0^{\mathrm{T}}, \ldots, \mathbf{I}_{L_b} \otimes \bar{\boldsymbol{\mathcal{G}}}_{L_g-1}^{\mathrm{T}} \right] \\
&\quad + \left[\mathbf{I}_{L_b} \otimes \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{G}}_0}^{\mathrm{T}}, \ldots, \mathbf{I}_{L_b} \otimes \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{G}}_{L_g-1}}^{\mathrm{T}} \right], \\
\hat{\mathbf{H}} &\triangleq \bar{\hat{\mathbf{H}}} + \boldsymbol{\Delta}_{\hat{\mathbf{H}}} \\
&= \left[\bar{\tilde{\mathbf{H}}}_0^{\mathrm{T}}, \ldots, \bar{\tilde{\mathbf{H}}}_{L_g-1}^{\mathrm{T}} \right]^{\mathrm{T}} + \left[\boldsymbol{\Delta}_{\tilde{\mathbf{H}}_0}^{\mathrm{T}}, \ldots, \boldsymbol{\Delta}_{\tilde{\mathbf{H}}_{L_g-1}}^{\mathrm{T}} \right]^{\mathrm{T}}, \\
\hat{\mathbf{G}} &\triangleq \left[\bar{\tilde{\mathbf{G}}}_0^{\mathrm{T}}, \ldots, \bar{\tilde{\mathbf{G}}}_{L_g-1}^{\mathrm{T}} \right]^{\mathrm{T}} + \left[\tilde{\boldsymbol{\Lambda}}_0^{\mathrm{T}}, \ldots, \tilde{\boldsymbol{\Lambda}}_{L_g-1}^{\mathrm{T}} \right]^{\mathrm{T}}, \\
\hat{\mathbf{I}} &\triangleq \left[\bar{\mathbf{I}}_0^{\mathrm{T}}, \ldots, \bar{\mathbf{I}}_{L_g-1}^{\mathrm{T}} \right]^{\mathrm{T}},
\end{aligned} \tag{11}
$$

the received signal can be reformulated as

$$
\begin{aligned}
\mathbf{y}_1 &= \mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{H}} \breve{\mathbf{X}}_1 + \mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{G}} \breve{\mathbf{X}}_2 + \mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{I}} \breve{\mathbf{n}}_r + \mathbf{n}_1, \\
\mathbf{y}_2 &= \mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{H}} \breve{\mathbf{X}}_1 + \mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{G}} \breve{\mathbf{X}}_2 + \mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{I}} \hat{\mathbf{n}}_r + \mathbf{n}_2,
\end{aligned} \tag{12}
$$

where we omit the time index k for simplicity.

From the definition of the source signal $\breve{\mathbf{X}}_i, i = 1,2$ in (11), $\breve{\mathbf{X}}_i$ can be decomposed into two parts, the desired signal $x_i[k]$ and the inter-symbol interference (ISI) $\breve{\mathbf{X}}_{i,r}[k]$ caused by the frequency selective channel. Let $\hat{\mathbf{h}}$ and $\breve{\mathbf{g}}$ denote the first columns of the $\hat{\mathbf{H}}$ and $\breve{\mathbf{G}}$, while $\hat{\mathbf{H}}_r$ and $\breve{\mathbf{G}}_r$ denote the residues of them, respectively, i.e. $\hat{\mathbf{H}} = [\ \hat{\mathbf{h}}, \quad \hat{\mathbf{H}}_r\ ]$ and $\breve{\mathbf{G}} = [\ \breve{\mathbf{g}}, \quad \breve{\mathbf{G}}_r\ ]$. Then, (12) can be rewritten as

$$
\begin{aligned}
\mathbf{y}_1 =& \mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \left[\breve{\mathbf{g}}\ \breve{\mathbf{G}}_r \right] \begin{bmatrix} x_2 \\ \breve{\mathbf{X}}_{2,r} \end{bmatrix} + \mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{H}} \breve{\mathbf{X}}_1 + \mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{I}} \breve{\mathbf{n}}_r + \mathbf{n}_1 \\
=& \underbrace{\mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{g}} x_2}_{Signal} + \underbrace{\mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{G}}_r \breve{\mathbf{X}}_{2,r} + \mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{H}} \breve{\mathbf{X}}_1}_{Interference} + \underbrace{\mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{I}} \breve{\mathbf{n}}_r + \mathbf{n}_1}_{Noise}, \\
\mathbf{y}_2 =& \mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \left[\hat{\mathbf{h}}\ \hat{\mathbf{H}}_r \right] \begin{bmatrix} x_1 \\ \breve{\mathbf{X}}_{1,r} \end{bmatrix} + \mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{G}} \breve{\mathbf{X}}_2 + \mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{I}} \hat{\mathbf{n}}_r + \mathbf{n}_2 \\
=& \underbrace{\mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{h}} x_1}_{Signal} + \underbrace{\mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{H}}_r \breve{\mathbf{X}}_{1,r} + \mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{G}} \breve{\mathbf{X}}_2}_{Interference} + \underbrace{\mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{I}} \hat{\mathbf{n}}_r + \mathbf{n}_2}_{Noise}.
\end{aligned} \tag{13}
$$

In (13), the received signal consists of four parts, the desired information, the inter-symbol interference, the self interference and the channel noise.

At each user terminal, with perfect knowledge of the self input signals and perfect CSI, it is achievable to cancel out the harmful effects of the self interference. However, because of the incomplete knowledge of the CSI, the self cancelation filters are not able to cancel the self signals completely in each receiver. Therefore, the residual signals after the self cancelation at the destinations are given by

$$
\begin{aligned}
\mathbf{y}_1 =& \underbrace{\mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{g}} x_2}_{Signal} + \underbrace{\mathbf{w}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{G}}_r \breve{\mathbf{X}}_{2,r} + (\mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{H}} - \mathbf{D}_1) \breve{\mathbf{X}}_1}_{Interference} \\
&+ \underbrace{\mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{I}} \breve{\mathbf{n}}_r + \mathbf{n}_1}_{Noise}, \\
\mathbf{y}_2 =& \underbrace{\mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{h}} x_1}_{Signal} + \underbrace{\mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{H}}_r \breve{\mathbf{X}}_{1,r} + (\mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{G}} - \mathbf{D}_2) \breve{\mathbf{X}}_2}_{Interference} \\
&+ \underbrace{\mathbf{w}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{I}} \hat{\mathbf{n}}_r + \mathbf{n}_2}_{Noise},
\end{aligned} \tag{14}
$$

where $\mathbf{D}_1 = \mathbf{b}^{\mathrm{H}} \bar{\breve{\boldsymbol{\mathcal{H}}}} \bar{\breve{\mathbf{H}}}$ and $\mathbf{D}_2 = \mathbf{b}^{\mathrm{H}} \bar{\hat{\boldsymbol{\mathcal{G}}}} \bar{\hat{\mathbf{G}}}$ are self cancelation filters at S_1 and S_2. Based on (14), the powers of the desired signal $P_{sig,i}, i = 1,2$, interference $P_{intf,i}, i = 1,2$ and noise $P_{nois,i}, i = 1,2$ are expressed as follows:

$$
\begin{aligned}
P_{sig,1} &\triangleq \mathbb{E}\{\|\mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{g}} x_2\|^2\}, \\
P_{sig,2} &\triangleq \mathbb{E}\{\|\mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{h}} x_1\|^2\}, \\
P_{intf,1} &\triangleq \mathbb{E}\{\|\mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{G}}_r \breve{\mathbf{X}}_{2,r} + (\mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{H}} - \mathbf{D}_1) \breve{\mathbf{X}}_1\|^2\}, \\
P_{intf,2} &\triangleq \mathbb{E}\{\|\mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{H}}_r \breve{\mathbf{X}}_{1,r} + (\mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{G}} - \mathbf{D}_2) \breve{\mathbf{X}}_2\|^2\}, \\
P_{nois,1} &\triangleq \mathbb{E}\{\|\mathbf{b}^{\mathrm{H}} \breve{\boldsymbol{\mathcal{H}}} \breve{\mathbf{I}} \breve{\mathbf{n}}_r + \mathbf{n}_1\|^2\}, \\
P_{nois,2} &\triangleq \mathbb{E}\{\|\mathbf{b}^{\mathrm{H}} \hat{\boldsymbol{\mathcal{G}}} \hat{\mathbf{I}} \hat{\mathbf{n}}_r + \mathbf{n}_2\|^2\}.
\end{aligned} \tag{15}
$$

III. Filter-and-Forward beamforming under channel uncertainties

In this section, we discuss in details the robust beamforming design with the imperfect CSI. As the discussion in Section II, the destinations only obtain partial channel information due to the estimation errors. The robust design aims at maximizing the total system signal-to-interference-plus-noise-ratio (SINR) under the individual relay power constraints. The optimization problem can be formulated as

$$\begin{aligned} \max_{\mathbf{w}} \; & \rho \mathrm{SINR}_1 + (1-\rho)\mathrm{SINR}_2 \\ s.t. \; & p_m \le P_{r_m max}, \; m = 1, \ldots, N_R, \end{aligned} \tag{16}$$

where $0 \leqslant \rho \leqslant 1$ is a constant. p_m is the power of the m_{th} relay and $P_{r_m max}$ is the maximum transmitted power of the m_{th} relay . SINR_1 and SINR_2 are given by

$$\begin{aligned} \mathrm{SINR}_1 &= \frac{P_{sig,1}}{P_{intf,1} + P_{nois,1}}, \\ \mathrm{SINR}_2 &= \frac{P_{sig,2}}{P_{intf,2} + P_{nois,2}}. \end{aligned} \tag{17}$$

The powers of the desired signals $P_{sig,1}$ and $P_{sig,2}$ in (15) can be expressed as

$$\begin{aligned} P_{sig,1} &= P_{s_2} \mathbf{b}^{\mathrm{H}} \boldsymbol{\Theta}_{11} \mathbf{b}, \\ P_{sig,2} &= P_{s_1} \mathbf{b}^{\mathrm{H}} \boldsymbol{\Theta}_{21} \mathbf{b}, \end{aligned} \tag{18}$$

with the definitions as follows:

$$\begin{aligned} \boldsymbol{\Theta}_{11} \triangleq\; & \bar{\breve{\boldsymbol{\mathcal{H}}}} \bar{\breve{\mathbf{g}}} \bar{\breve{\mathbf{g}}}^{\mathrm{H}} \bar{\breve{\boldsymbol{\mathcal{H}}}}^{\mathrm{H}} + \bar{\breve{\boldsymbol{\mathcal{H}}}} \mathbb{E}\{\boldsymbol{\Lambda}_{\breve{g}} \boldsymbol{\Lambda}_{\breve{g}}^{\mathrm{H}}\} \bar{\breve{\boldsymbol{\mathcal{H}}}}^{\mathrm{H}} \\ & + \mathbb{E}\{\boldsymbol{\Delta}_{\breve{\mathcal{H}}} \bar{\breve{\mathbf{g}}} \bar{\breve{\mathbf{g}}}^{\mathrm{H}} \boldsymbol{\Delta}_{\breve{\mathcal{H}}}^{\mathrm{H}}\} + \mathbb{E}\{\boldsymbol{\Delta}_{\breve{\mathcal{H}}} \boldsymbol{\Lambda}_{\breve{g}} \boldsymbol{\Lambda}_{\breve{g}}^{\mathrm{H}} \boldsymbol{\Delta}_{\breve{\mathcal{H}}}^{\mathrm{H}}\}, \\ \boldsymbol{\Theta}_{21} \triangleq\; & \bar{\hat{\boldsymbol{\mathcal{G}}}} \bar{\hat{\mathbf{h}}} \bar{\hat{\mathbf{h}}}^{\mathrm{H}} \bar{\hat{\boldsymbol{\mathcal{G}}}}^{\mathrm{H}} + \bar{\hat{\boldsymbol{\mathcal{G}}}} \mathbb{E}\{\boldsymbol{\Delta}_{\hat{h}} \boldsymbol{\Delta}_{\hat{h}}^{\mathrm{H}}\} \bar{\hat{\boldsymbol{\mathcal{G}}}}^{\mathrm{H}} \\ & + \mathbb{E}\{\boldsymbol{\Lambda}_{\hat{\mathcal{G}}} \bar{\hat{\mathbf{h}}} \bar{\hat{\mathbf{h}}}^{\mathrm{H}} \boldsymbol{\Lambda}_{\hat{\mathcal{G}}}^{\mathrm{H}}\} + \mathbb{E}\{\boldsymbol{\Lambda}_{\hat{\mathcal{G}}} \boldsymbol{\Delta}_{\hat{h}} \boldsymbol{\Delta}_{\hat{h}}^{\mathrm{H}} \boldsymbol{\Lambda}_{\hat{\mathcal{G}}}^{\mathrm{H}}\}. \end{aligned} \tag{19}$$

Note that $\bar{\boldsymbol{\Upsilon}}$ is the estimated value of $\boldsymbol{\Upsilon}$ via channel estimation, and $\boldsymbol{\Delta}_{\boldsymbol{\Upsilon}}(\boldsymbol{\Lambda}_{\boldsymbol{\Upsilon}})$ is the estimation error. Using the statistic properties and the error expressions in (2) and (11), the last three terms with respect to the uncertainties in (19) can be calculated, which are shown in Appendix.

The powers of the interferences and the noises can be derived as

$$\begin{aligned} P_{intf,1} =& \mathbf{b}^{\mathrm{H}} \boldsymbol{\Theta}_{12} \mathbf{b}, \\ P_{intf,2} =& \mathbf{b}^{\mathrm{H}} \boldsymbol{\Theta}_{22} \mathbf{b}, \end{aligned} \tag{20}$$

and

$$\begin{aligned} P_{nois,1} =& \mathbf{b}^{\mathrm{H}} \boldsymbol{\Theta}_{13} \mathbf{b} + \sigma_{n_1}^2, \\ P_{nois,2} =& \mathbf{b}^{\mathrm{H}} \boldsymbol{\Theta}_{23} \mathbf{b} + \sigma_{n_2}^2, \end{aligned} \tag{21}$$

under the definitions that

$$\begin{aligned} \boldsymbol{\Theta}_{12} \triangleq & P_{s_2} (\bar{\breve{\boldsymbol{\mathcal{H}}}} \bar{\breve{\mathbf{G}}}_r \bar{\breve{\mathbf{G}}}_r^{\mathrm{H}} \bar{\breve{\boldsymbol{\mathcal{H}}}}^{\mathrm{H}} + \bar{\breve{\boldsymbol{\mathcal{H}}}} \mathbb{E}\{\boldsymbol{\Lambda}_{\breve{G}_r} \boldsymbol{\Lambda}_{\breve{G}_r}^{\mathrm{H}}\} \bar{\breve{\boldsymbol{\mathcal{H}}}}^{\mathrm{H}} \\ & + \mathbb{E}\{\boldsymbol{\Delta}_{\breve{\mathcal{H}}} \bar{\breve{\mathbf{G}}}_r \bar{\breve{\mathbf{G}}}_r^{\mathrm{H}} \boldsymbol{\Delta}_{\breve{\mathcal{H}}}^{\mathrm{H}}\} + \mathbb{E}\{\boldsymbol{\Delta}_{\breve{\mathcal{H}}} \boldsymbol{\Lambda}_{\breve{G}_r} \boldsymbol{\Lambda}_{\breve{G}_r}^{\mathrm{H}} \boldsymbol{\Delta}_{\breve{\mathcal{H}}}^{\mathrm{H}}\}) \\ & + P_{s_1} (\bar{\breve{\boldsymbol{\mathcal{H}}}} \mathbb{E}\{\boldsymbol{\Delta}_{\breve{H}} \boldsymbol{\Delta}_{\breve{H}}^{\mathrm{H}}\} \bar{\breve{\boldsymbol{\mathcal{H}}}}^{\mathrm{H}} + \mathbb{E}\{\boldsymbol{\Delta}_{\breve{\mathcal{H}}} \bar{\breve{\mathbf{H}}} \bar{\breve{\mathbf{H}}}^{\mathrm{H}} \boldsymbol{\Delta}_{\breve{\mathcal{H}}}^{\mathrm{H}}\} \\ & + \mathbb{E}\{\boldsymbol{\Delta}_{\breve{\mathcal{H}}} \boldsymbol{\Delta}_{\breve{H}} \boldsymbol{\Delta}_{\breve{H}}^{\mathrm{H}} \boldsymbol{\Delta}_{\breve{\mathcal{H}}}^{\mathrm{H}}\}), \end{aligned} \tag{22}$$

and

$$\begin{aligned} \boldsymbol{\Theta}_{22} \triangleq & P_{s_1} (\bar{\hat{\boldsymbol{\mathcal{G}}}} \bar{\hat{\mathbf{H}}}_r \bar{\hat{\mathbf{F}}}_r^{\mathrm{H}} \bar{\hat{\boldsymbol{\mathcal{G}}}}^{\mathrm{H}} + \hat{\boldsymbol{\mathcal{G}}} \mathbb{E}\{\boldsymbol{\Delta}_{\hat{H}_r} \boldsymbol{\Delta}_{\hat{H}_r}^{\mathrm{H}}\} \bar{\hat{\boldsymbol{\mathcal{G}}}}^{\mathrm{H}} \\ & + \mathbb{E}\{\boldsymbol{\Lambda}_{\hat{\mathcal{G}}} \bar{\hat{\mathbf{H}}}_r \bar{\hat{\mathbf{H}}}_r^{\mathrm{H}} \boldsymbol{\Lambda}_{\hat{\mathcal{G}}}^{\mathrm{H}}\} + \mathbb{E}\{\boldsymbol{\Lambda}_{\hat{\mathcal{G}}} \boldsymbol{\Delta}_{\hat{H}_r} \boldsymbol{\Delta}_{\hat{H}_r}^{\mathrm{H}} \boldsymbol{\Lambda}_{\hat{\mathcal{G}}}^{\mathrm{H}}\}) \\ & + P_{s_2} (\bar{\hat{\boldsymbol{\mathcal{G}}}} \mathbb{E}\{\boldsymbol{\Lambda}_{\hat{G}} \boldsymbol{\Lambda}_{\hat{G}}^{\mathrm{H}}\} \bar{\hat{\boldsymbol{\mathcal{G}}}}^{\mathrm{H}} + \mathbb{E}\{\boldsymbol{\Lambda}_{\hat{\mathcal{G}}} \bar{\hat{\mathbf{G}}} \bar{\hat{\mathbf{G}}}^{\mathrm{H}} \boldsymbol{\Lambda}_{\hat{\mathcal{G}}}^{\mathrm{H}}\} \\ & + \mathbb{E}\{\boldsymbol{\Lambda}_{\hat{\mathcal{G}}} \boldsymbol{\Lambda}_{\hat{G}} \boldsymbol{\Lambda}_{\hat{G}}^{\mathrm{H}} \boldsymbol{\Lambda}_{\hat{\mathcal{G}}}^{\mathrm{H}}\}), \\ \boldsymbol{\Theta}_{13} \triangleq & \bar{\breve{\boldsymbol{\mathcal{H}}}} \breve{\mathbf{I}} \mathbf{R}_{\breve{n}_r} \breve{\mathbf{I}}^{\mathrm{H}} \bar{\breve{\boldsymbol{\mathcal{H}}}}^{\mathrm{H}} + \mathbb{E}\{\boldsymbol{\Delta}_{\breve{\mathcal{H}}} \breve{\mathbf{I}} \mathbf{R}_{\breve{n}_r} \breve{\mathbf{I}}^{\mathrm{H}} \boldsymbol{\Delta}_{\breve{\mathcal{H}}}^{\mathrm{H}}\}, \\ \boldsymbol{\Theta}_{23} \triangleq & \bar{\breve{\boldsymbol{\mathcal{G}}}} \hat{\mathbf{I}} \mathbf{R}_{\hat{n}_r} \hat{\mathbf{I}}^{\mathrm{H}} \bar{\hat{\boldsymbol{\mathcal{G}}}}^{\mathrm{H}} + \mathbb{E}\{\boldsymbol{\Lambda}_{\hat{\mathcal{G}}} \hat{\mathbf{I}} \mathbf{R}_{\hat{n}_r} \hat{\mathbf{I}}^{\mathrm{H}} \boldsymbol{\Lambda}_{\hat{\mathcal{G}}}^{\mathrm{H}}\}. \end{aligned} \tag{23}$$

Similarly, the interferences and the noises are decomposed into the estimated terms and estimation error terms, which are calculated as the Appendix.

For the constraints in (16), the power of the m_{th} relay is written as

$$\begin{aligned} p_m =& \mathbb{E}\{\|\mathbf{t}_m\|^2\} \\ =& P_{s_1} \mathbf{e}_m^{\mathrm{T}} \mathbf{B}^{\mathrm{H}} \bar{\tilde{\boldsymbol{H}}} \bar{\tilde{\boldsymbol{H}}}^{\mathrm{H}} \mathbf{B} \mathbf{e}_m + P_{s_2} \mathbf{e}_m^{\mathrm{T}} \mathbf{B}^{\mathrm{H}} \bar{\tilde{\boldsymbol{G}}} \bar{\tilde{\boldsymbol{G}}}^{\mathrm{H}} \mathbf{B} \mathbf{e}_m \\ & + \mathbf{e}_m^{\mathrm{T}} \mathbf{B}^{\mathrm{H}} \mathbf{R}_\eta \mathbf{B} \mathbf{e}_m + P_{s_1} \mathbf{e}_m^{\mathrm{T}} \mathbf{B}^{\mathrm{H}} \boldsymbol{\Xi}_1 \mathbf{B} \mathbf{e}_m \\ & + P_{s_2} \mathbf{e}_m^{\mathrm{T}} \mathbf{B}^{\mathrm{H}} \boldsymbol{\Xi}_2 \mathbf{B} \mathbf{e}_m, \end{aligned} \tag{24}$$

where $\boldsymbol{\Xi}_1 \triangleq \mathbb{E}\{\boldsymbol{\Delta}_{\tilde{H}} \boldsymbol{\Delta}_{\tilde{H}}^{\mathrm{H}}\}$ and $\boldsymbol{\Xi}_2 \triangleq \mathbb{E}\{\boldsymbol{\Delta}_{\tilde{G}} \boldsymbol{\Delta}_{\tilde{G}}^{\mathrm{H}}\}$. $\mathbf{e}_m$ is the m_{th} column of the identity matrix. Using $\mathbf{E}_m \triangleq \mathrm{diag}\{\mathbf{e}_m\}$ and the properties of the Kronecker product, the m_{th} transmitted power can be rewritten as

$$p_m = \mathbf{b}^{\mathrm{H}} \boldsymbol{\Theta}_{3m} \mathbf{b}, \tag{25}$$

where the matrix $\boldsymbol{\Theta}_{3m}$ is equivalent to

$$\begin{aligned} \boldsymbol{\Theta}_{3m} \triangleq & p_{s_1} (\mathbf{I}_{L_b} \otimes \mathbf{E}_m) \tilde{\boldsymbol{H}} \tilde{\boldsymbol{H}}^{\mathrm{H}} (\mathbf{I}_{L_b} \otimes \mathbf{E}_m)^{\mathrm{H}} \\ & + p_{s_2} (\mathbf{I}_{L_b} \otimes \mathbf{E}_m) \tilde{\boldsymbol{G}} \tilde{\boldsymbol{G}}^{\mathrm{H}} (\mathbf{I}_{L_b} \otimes \mathbf{E}_m)^{\mathrm{H}} \\ & + p_{s_1} (\mathbf{I}_{L_b} \otimes \mathbf{E}_m) \boldsymbol{\Xi}_1 (\mathbf{I}_{L_b} \otimes \mathbf{E}_m)^{\mathrm{H}} \\ & + p_{s_2} (\mathbf{I}_{L_b} \otimes \mathbf{E}_m) \boldsymbol{\Xi}_2 (\mathbf{I}_{L_b} \otimes \mathbf{E}_m)^{\mathrm{H}} \\ & + \sigma_\eta^2 (\mathbf{I}_{L_b} \otimes \mathbf{E}_m)(\mathbf{I}_{L_b} \otimes \mathbf{E}_m)^{\mathrm{H}}. \end{aligned} \tag{26}$$

Therefore, the optimization problem can be reformulated as

$$\begin{aligned} \max_{\mathbf{b}} \; & \rho \, \frac{P_{s_2} \mathbf{b}^{\mathrm{H}} \boldsymbol{\Theta}_{11} \mathbf{b}}{\mathbf{b}^{\mathrm{H}} (\boldsymbol{\Theta}_{12} + \boldsymbol{\Theta}_{13}) \mathbf{b} + \sigma_{n_1}^2} \\ & + (1-\rho) \, \frac{P_{s_1} \mathbf{b}^{\mathrm{H}} \boldsymbol{\Theta}_{21} \mathbf{b}}{\mathbf{b}^{\mathrm{H}} (\boldsymbol{\Theta}_{22} + \boldsymbol{\Theta}_{23}) \mathbf{b} + \sigma_{n_1}^2} \\ & s.t. \; \mathbf{b}^{\mathrm{H}} \boldsymbol{\Theta}_{3m} \mathbf{b} \le P_{r_m max}, \; m = 1, \ldots, N_R. \end{aligned} \tag{27}$$

According to the equalities $\mathrm{Tr}\{\mathbb{ABC}\} = \mathrm{Tr}\{\mathbb{CAB}\}$ and $\boldsymbol{\alpha}^{\mathrm{H}} \boldsymbol{\beta} = \mathrm{Tr}\{\boldsymbol{\beta} \boldsymbol{\alpha}^{\mathrm{H}}\}$, we introduce the semi-definite matrix $\boldsymbol{\Phi} \triangleq \mathbf{b}\mathbf{b}^{\mathrm{H}}$, then the above problem is equivalent to

$$\begin{aligned} \max_{\boldsymbol{\Phi}} \; & \rho \, \frac{\mathrm{Tr}\{\boldsymbol{\Phi} \boldsymbol{\Theta}_{11}\}}{\mathrm{Tr}\{\boldsymbol{\Phi}(\boldsymbol{\Theta}_{12} + \boldsymbol{\Theta}_{13})\} + \sigma_{n_1}^2} \\ & + (1-\rho) \, \frac{\mathrm{Tr}\{\boldsymbol{\Phi} \boldsymbol{\Theta}_{21}\}}{\mathrm{Tr}\{\boldsymbol{\Phi}(\boldsymbol{\Theta}_{22} + \boldsymbol{\Theta}_{23})\} + \sigma_{n_2}^2} \\ s.t. \; & \mathrm{Tr}\{\boldsymbol{\Phi} \boldsymbol{\Theta}_{3m}\} \le P_{r_m max}, \; m = 1, \ldots, N_R, \\ & \mathrm{rank}(\Phi) = 1. \end{aligned} \tag{28}$$

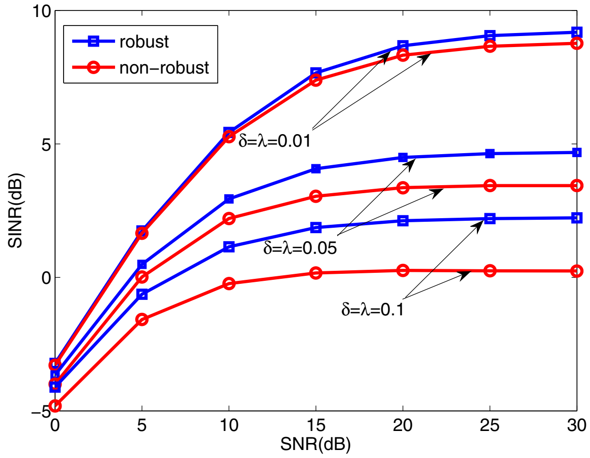

Fig. 2. The total SINR performance of the destinations subjected to the individual relay transmitted power for the different uncertainty size.

Clearly, the above problem is non-convex since the objective function is not a concave function. After semi-definite relaxation (SDR) [11] and introducing auxiliary variables, the optimization with individual power constraints can be reformulated as

$$\begin{aligned} \max_{\mathbf{\Phi},\Gamma} \quad & \Gamma = \gamma_1 + \gamma_2 \\ \text{s.t. } & \mathrm{Tr}\{\mathbf{\Phi\Theta}_{11}\} - \frac{\gamma_1}{\rho}\mathrm{Tr}\{\mathbf{\Phi}(\mathbf{\Theta}_{12}+\mathbf{\Theta}_{13})\} \geq \frac{\gamma_1}{\rho}\sigma_{n_1}^2, \\ & \mathrm{Tr}\{\mathbf{\Phi\Theta}_{21}\} - \frac{\gamma_2}{1-\rho}\mathrm{Tr}\{\mathbf{\Phi}(\mathbf{\Theta}_{22}+\mathbf{\Theta}_{23})\} \geq \frac{\gamma_2}{1-\rho}\sigma_{n_2}^2, \\ & \mathrm{Tr}\{\mathbf{\Phi\Theta}_{3m}\} \leq P_{r_m max}, \ m = 1, \ldots, N_R, \\ & \mathbf{\Phi} \succeq 0, \end{aligned} \tag{29}$$

where Γ is the sum signal-to-noise ratio (SNR) given a SNR profile factors $(k,\bar{k})$ with $0 \leqslant k \leqslant 1$ and $\bar{k} = 1-k$. Positive variables γ_1 and γ_2 equal to $\gamma_1 = k\Gamma$ and $\gamma_2 = \bar{k}\Gamma$. Since there are N_R individual power constraints, we aim at solving a series of the following problem via bi-section search over the SNR region $\tau \in [0, \tau_{max}]$. For each given value of τ, the above problem is convex over $\mathbf{\Phi}$, whose solution details are given as follows [12]:

- Initialize $\tau_{low} = 0$, $\tau_{up} = \tau_{max}$.
- Repeat
  1) Set $\tau = \frac{1}{2}(\tau_{low} + \tau_{up})$, then $\gamma_1 = k\tau$ and $\gamma_2 = \bar{k}\tau$.
  2) Solve problem (29) with given τ.
  3) Update τ with the bi-section method: If $\mathrm{Tr}\{\mathbf{\Phi\Theta}_{3m}\} \leq P_{r_m max}$, set $\tau_{low} = \tau$; otherwise, $\tau_{up} = \tau$.
- Until $\tau_{up} - \tau_{low} < \epsilon$, where ϵ is a small positive accuracy parameter.

IV. Simulation Results

In this section, we consider a two-way relaying network with $N_R = 3$ relays and quasi-static frequency selective channels with the length $L_h = L_g = 3$. The transmitted signal power for both terminals and relay nodes are assumed as $P_{s_1} = P_{s_2} = 1$ and $P_{r_m max} = 2$. The constant ρ is set equal to 0.5 in this simulation. The true channel impulse response coefficients are modeled as zero-mean complex Gaussian random variables with an exponential power delay profile. The SINR performances of the destinations subjected to the individual relay transmitted power for the different uncertainty level, δ=λ=0.01, 0.05 and 0.1, are shown in Fig.2. As can be seen, the robust design results in a better total SINR performance than the non-robust case because it takes the estimation error into account. As the increase of the uncertainty level, this advantage gets more obvious.

V. Conclusion

In this paper, the distributed filter-and-forward beamforming for two-way relaying networks over the frequency selective fading channels has been investigated. The inevitable channel errors following the stochastic error model has been considered and taken into account in the beamforming design. The FF beamforers was designed based on the optimization criteria of maximizing the total system SINR under the individual relay power constraints. Simulation results have validated the effectiveness of the proposed beamformer.

VI. Appendix

The terms with respect to the estimation errors in (19), (22), (23) and (26) can be calculated based on the following expressions, i.e. $\mathbb{E}\{\mathbf{XX}^{\mathrm{H}}\}$, $\mathbb{E}\{\mathbf{YAY}^{\mathrm{H}}\}$ and $\mathbb{E}\{\mathbf{YXX}^{\mathrm{H}}\mathbf{Y}^{\mathrm{H}}\}$, in which $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{A}$ can be written as

$$\begin{aligned} &\mathbf{X}(p,:) = [\mathbf{0}_{1\times(k+i)}, \ x_{j,0}, \ \ldots \ , x_{j,L_1-1}, \mathbf{0}_{1\times(L_b+L_1-2-k-i)}], \\ &\mathbf{Y}(q,:) = [\mathbf{0}_{1\times(i\Omega+j-1)}, \ x_{j,0}, \ \mathbf{0}_{1\times(\Omega L_b-1)}, \ x_{j,1}, \ \mathbf{0}_{1\times(\Omega L_b-1)}, \\ &\qquad\qquad \ldots, x_{j,L_1-1}, \mathbf{0}_{1\times((L_b-i)\Omega-j)}], \\ &\mathbf{A}(p,p') \ \ = \{a_{p,p'}\}, \\ &p = k\Omega L_b + i\Omega + j, \ q = i\Omega + j, \ p' = k'\Omega L_b + i'\Omega + j', \\ &k,k' = 0,\ldots,L_1-1, \ i,i' = 0,\ldots,L_b-1, \ j,j' = 1,\ldots,\Omega, \end{aligned} \tag{30}$$

where $\mathbf{X}(p,:)$ denotes the p_{th} row of the matrix $\mathbf{X}$ and $\mathbf{A}(p,p')$ denotes the p_{th} row and the p'_{th} column element of the matrix $\mathbf{A}$. In this Appendix, we assume that $L_1 = L_h(L_g)$ and $\Omega = N_R$.

A. $\mathbb{E}\{\mathbf{XX}^{\mathrm{H}}\}$

Clearly, we have

$$\mathbb{E}\{\mathbf{XX}^{\mathrm{H}}\}(p,p') = \mathbb{E}\{\mathbf{X}(p,:)\mathbf{X}^{\mathrm{H}}(p',:)\} \tag{31}$$

Due to the statistic independence of the channel estimation error and the definitions in (2), (31) can be derived as

$$\mathbb{E}\{\mathbf{XX}^{\mathrm{H}}\}(p,p') = L_1, \ when \ \mathbf{X}(p,:) = \mathbf{X}(p',:). \tag{32}$$

$$
\mathbb{E}\{\mathbf{Y}\mathbf{X}\mathbf{X}^{\mathrm{H}}\mathbf{Y}^{\mathrm{H}}\}
= \left[\begin{array}{cccccccccc}
\boldsymbol{\Sigma} & \mathbf{0}_{\Omega\times\Omega} & (L_1-1)_\Omega & \mathbf{0}_{\Omega\times\Omega} & (L_1-2)_\Omega & \cdots & (L_1-\Pi+1)_\Omega & \mathbf{0}_{\Omega\times\Omega} & \cdots & \mathbf{0}_{\Omega\times\Omega} \\
\mathbf{0}_{\Omega\times\Omega} & \boldsymbol{\Sigma} & \mathbf{0}_{\Omega\times\Omega} & (L_1-1)_\Omega & \mathbf{0}_{\Omega\times\Omega} & & & & & \vdots \\
(L_1-1)_\Omega & \mathbf{0}_{\Omega\times\Omega} & \boldsymbol{\Sigma} & & & & & & & \mathbf{0}_{\Omega\times\Omega} \\
\mathbf{0}_{\Omega\times\Omega} & (L_1-1)_\Omega & \mathbf{0}_{\Omega\times\Omega} & \ddots & & & & & \mathbf{0}_{\Omega\times\Omega} & (L_1-\Pi+1)_\Omega \\
(L_1-2)_\Omega & \mathbf{0}_{\Omega\times\Omega} & (L_1-1)_\Omega & & \boldsymbol{\Sigma} & & (L_1-1)_\Omega & & & \vdots \\
\vdots & \vdots & \ddots & & & \ddots & & \ddots & & (L_1-2)_\Omega \\
(L_1-\Pi+1)_\Omega & \mathbf{0}_{\Omega\times\Omega} & & \ddots & & & \boldsymbol{\Sigma} & \mathbf{0}_{\Omega\times\Omega} & (L_1-1)_\Omega & \mathbf{0}_{\Omega\times\Omega} \\
\mathbf{0}_{\Omega\times\Omega} & (L_1-\Pi+1)_\Omega & & & & \ddots & & \ddots & \mathbf{0}_{\Omega\times\Omega} & (L_1-1)_\Omega \\
\vdots & \vdots & & \ddots & & \ddots & & \ddots & \boldsymbol{\Sigma} & \mathbf{0}_{\Omega\times\Omega} \\
\mathbf{0}_{\Omega\times\Omega} & \cdots & \mathbf{0}_{\Omega\times\Omega} & (L_1-\Pi+1)_\Omega & \cdots & (L_1-2)_\Omega & \mathbf{0}_{\Omega\times\Omega} & (L_1-1)_\Omega & \mathbf{0}_{\Omega\times\Omega} & \boldsymbol{\Sigma}
\end{array}\right], \tag{38}
$$

B. $\mathbb{E}\{\mathbf{YAY}^{\mathrm{H}}\}$

The matrix $\mathbf{YA}$ equals to

$$
\begin{aligned}
\{\mathbf{YA}\}(\alpha,d) &= \sum_{k=0}^{L_1-1} x_{j,k} a_{k\Omega L_b+i\Omega+j,d}, \\
&\alpha = i\Omega + j, \quad d = 1,\ldots,\Omega L_b L_1.
\end{aligned} \tag{33}
$$

Then, we have the product $\mathbf{YAY}^{\mathrm{H}}$ as follows:

$$
\begin{aligned}
\{\mathbf{YAY}^{\mathrm{H}}\}(\alpha,\beta) &= \sum_{k=0}^{L_1-1}\sum_{k'=0}^{L_1-1} x_{j,k}\ a_{k\Omega L_b+i\Omega+j,d}\ x_{j',k'} \\
&= \sum_{k=0}^{L_1-1}\sum_{k'=0}^{L_1-1} a_{k\Omega L_b+i\Omega+j,d}\ x_{j,k}\ x_{j',k'}, \\
&\alpha = i\Omega + j,\ \beta = i'\Omega + j', \\
&i,i' = 0,\ldots,L_b-1,\ j,j' = 1,\ldots,\Omega.
\end{aligned} \tag{34}
$$

When $j = j'$ $k = k'$, $\mathbb{E}\{x_{j,k}\ x_{j',k'}\} = 1$, otherwise $\mathbb{E}\{x_{j,k}\ x_{j',k'}\} = 0$. Then, we have

$$
\begin{aligned}
\mathbb{E}\{\mathbf{YAY}^{\mathrm{H}}\}(\alpha,\beta) &= \mathbb{E}\{\mathbf{Y}(\alpha,:)\mathbf{AY}^{\mathrm{H}}(\beta,:)\} \\
&= \sum_{k=0}^{L_1-1}\sum_{k'=0}^{L_1-1} a_{k\Omega L_b+\alpha,\ k'\Omega L_b+\beta} \\
&\alpha = 1,\ldots,\Omega L_b,\ \beta = 1,\ldots,\Omega L_b.
\end{aligned} \tag{35}
$$

C. $\mathbb{E}\{\mathbf{YXX}^{\mathrm{H}}\mathbf{Y}^{\mathrm{H}}\}$

The α_{th} row of the matrix $\mathbf{YX}$ can be expressed as

$$
\mathbf{YX}(\alpha,:) = \left[\mathbf{0}_{1\times i}\ \ \mathbf{b}_{1\times(2L_1-1)}\ \mathbf{0}_{1\times L_b-1-i}\right], \tag{36}
$$

where

$$
\begin{aligned}
\mathbf{b}(\varphi) &= \begin{cases} \sum\limits_{\varepsilon=0}^{\varphi-1} x_{j,\varepsilon}x_{j,\varphi-\varepsilon-1}, & \varphi \le L_1 \\ \sum\limits_{\varepsilon=0}^{2L_1-\varphi-1} x_{j,L_1-\varepsilon-1}x_{j,\varphi-L_1+\varepsilon}, & \varphi > L_1 \end{cases} \\
&\varphi = 1,\ldots,2L_1-1
\end{aligned} \tag{37}
$$

Thus, we can obtain $\mathbb{E}\{\mathbf{YXX}^{\mathrm{H}}\mathbf{Y}^{\mathrm{H}}\}$ shown as (38), where

$$
\boldsymbol{\Sigma} = \begin{bmatrix} 2L_1^2-L_1 & L_1 & \cdots & L_1 \\ L_1 & 2L_1^2-L_1 & \cdots & L_1 \\ \vdots & \vdots & \ddots & \vdots \\ L_1 & \cdots & L_1 & 2L_1^2-L_1 \end{bmatrix}_{\Omega\times\Omega},
$$

$$
\Pi = \begin{cases} (L_b+1)/2, & L_b\ is\ odd; \\ L_b/2, & L_b\ is\ even. \end{cases}
$$

References

[1] F. Gao, R. Zhang and Y. Liang, "Channel Estimation for OFDM Modulated Two-Way Relay Networks," *IEEE Trans. Signal Process.*, vol.57, no.11, pp.4443-4455, Nov. 2009

[2] M. Zeng, R. Zhang and S. Cui, "On Design of Collaborative Beamforming for Two-Way Relay Networks," *IEEE Trans. Signal Process.*, vol. 59, pp. 2284-2295, 2011.

[3] B. Nazer, and M. Gastpar, "Reliable Physical Layer Network Coding," *Proceedings of the IEEE*, vol. 99, no.3, pp.438-460, 2011.

[4] H. Q. Ngo, T. Q. S. Quek, and H. Shin, "Amplify-and-Forward Two-Way Relay Networks: Error Exponents and Resource Allocation," *IEEE Trans. Commun.*, vol. 58, no.9, pp.2653-2666, Sep. 2010.

[5] Y. Zhou, Z. G. Pan, L. Tian, G. Sun, and J. L. Shi, "Cell Throughput of Multicast Services in OFDM-Based Distributed Antenna Systems," in *IEEE Proc. VTC'2011-Spring*, May 2011, pp. 1-5.

[6] H. Chen, A. B. Gershman and S. Shahbazpanahi, "Filter-and-Forward Distributed Beamforming in Relay Networks With Frequency Selective Fading," in *IEEE Trans. Signal Process.*, vol. 58, pp. 1251-1262, March, 2010.

[7] Y. Liang, A. Ikhleti, W. Gerstackertt, and R. Schobert, "Cooperative Two-Way Filter-and-Forward Beamforming for Frequency-Selective Channels," in *IEEE Proc. ICC'2011*, July 2011, pp. 1-6.

[8] H. Chen, A. B. Gershman and S. Shahbazpanahi, "Filter-and-Forward Distributed Beamforming for Two-Way Relay Networks with Frequency Selective Channels," in *IEEE Proc. GlOCOM'2010*, Jan. 2010, pp. 1-5.

[9] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z.-Q. Luo, "Distributed beamforming for relay networks based on second-order statistics of the channel state information," *IEEE Trans. Signal Process.*, vol. 56, pp. 4306C4316, Sep. 2008.

[10] T. Q. S. Quek, M. Z. Win, and M. Chiani, "Robust Power Allocation Algorithms for Wireless Relay Networks," *IEEE Trans. Commun.*, vol. 58, no.7, pp. 1931-1938, Jul. 2010.

[11] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[12] M. Zeng, R. Zhang, and S. Cui, "On Design of Collaborative Beamforming for Two-Way Relay Networks ", *IEEE Trans. Signal Process.*, vol. 56, pp. 2284 - 2295, Jan. 2011.