

# Filter-And-Forward Relay Design for MIMO-OFDM Systems

Donggun Kim, *Student Member, IEEE*, Youngchul Sung, *Senior Member, IEEE*, and Jihoon Chung, *Student Member, IEEE*

**Abstract**—In this paper, the filter-and-forward (FF) relay design for multiple-input-multiple-output (MIMO) orthogonal frequency-division multiplexing (OFDM) systems is investigated. Due to the considered MIMO structure, the FF relay design is investigated in the framework of joint design together with the linear MIMO transceiver. As the design criterion, first, the minimization of weighted sum mean square error (MSE) is considered. The joint design in this case is approached based on alternating optimization that iterates between the optimal design of the FF relay for given MIMO precoding and decoding matrices and the optimal design of MIMO precoding and decoding matrices for a given FF relay filter. Second, a more advanced problem of joint design for rate maximization is considered. The second problem is approached based on the obtained result regarding the first problem of weighted sum MSE minimization and the existing result regarding the relationship between weighted MSE minimization and rate maximization. Numerical results show the effectiveness of the proposed FF relay design method and significant performance improvement by the proposed FF relay over widely considered simple AF relays for MIMO-OFDM systems.

**Index Terms**—Linear relays, filter-and-forward relays, weighted mean square error, rate maximization, MIMO-OFDM.

## I. INTRODUCTION

RECENTLY, the filter-and-forward (FF) relaying scheme has gained interest from the research communities as an alternative relaying strategy due to its capability of performance improvement over simple AF relays and low complexity compared to other advanced relaying schemes such as decode-and-forward (DF) and compress-and-forward (CF) schemes [1]–[7]. With some available information-theoretical results on this topic [1]–[3], many researchers considered the practical design of FF relays in various settings. For example, Chen *et al.* proposed a half-duplex FF relay design method to minimize the relay transmit power and to maximize the quality-of-service (QoS) in a setting with one transmitter, one receiver and multiple FF relays [4]. Under a similar setting, Liang *et al.* optimized the FF relay filter to maximize the signal-to-noise ratio (SNR) at the destination equipped with a simple linear or decision-feedback equalizer [6]. Kim *et al.* consid-

ered rate maximization for FF relays and proposed a joint source and FF relay filter design algorithm for rate maximization based on the adaptive projected subgradient method [5]. However, all these works regarding FF relay design were conducted for single-carrier transmission, which is not widely adopted in current wireless standards employing OFDM. In the research field of relaying techniques for OFDM systems, most researchers considered the case that the relay performs OFDM processing [8]–[13]. In OFDM-processing relays, the incoming signal to the relay is OFDM-demodulated, amplified or decoded, OFDM-remodulated, and transmitted to the destination. However, OFDM processing at the relay incurs heavy processing complexity and high power consumption, and requires a highly linear power amplifier at the relay. Thus, to eliminate such heavy complexity burden at the relay, Kim *et al.* introduced direct finite-impulse response (FIR) filtering *without OFDM processing* at the relay, proposed several FF relay design methods for OFDM systems based on worst subcarrier SNR maximization or direct rate maximization, and showed that such direct FIR filtering relays can yield comparable performance relative to OFDM-processing relays [7], [14]. However, the limitation of their work is that they considered single-input single-output (SISO)-OFDM only and their approach based on worst subcarrier SNR maximization or direct rate maximization is not easily extended to the MIMO-OFDM<sup>1</sup> case, since SNR is not clearly defined in the MIMO case.

In this paper, we consider the FF relay design for MIMO-OFDM systems. In the MIMO case, the FF relay should not be designed alone but it should be designed jointly with the MIMO precoder at the source and the MIMO decoder at the destination. Hence, in this paper we consider the joint design of the FF relay at the relay and the linear MIMO transceiver, i.e., MIMO precoding and decoding matrices, at the source and the destination for MIMO-OFDM systems. As mentioned already, in the MIMO case, it is not easy to use SNR as the design metric as in the SISO case [7]. Thus, we approach the design problem based on the tractable criterion of weighted sum MSE minimization first, and then consider the more advanced rate-maximizing design problem based on the equivalence relationship between rate maximization and weighted MSE minimization with a properly chosen weight matrix [15]–[19]. We tackle the complicated joint design problem by using alternating optimization, which enables us to exploit the existing results for the MIMO

Manuscript received October 10, 2013; revised February 27, 2014 and May 19, 2014; accepted June 3, 2014. Date of publication June 9, 2014; date of current version July 18, 2014. This work was supported by the Ministry of Science, ICT and Future Planning, Korea, under the ICT R&D Program 2013. The editor coordinating the review of this paper and approving it for publication was M. Juntti.

The authors are with the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea (e-mail: dg.kim@kaist.ac.kr; ysung@ee.kaist.ac.kr; j.chung@kaist.ac.kr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCOMM.2014.2329691

<sup>1</sup>There exists vast literature regarding the relay design for MIMO-OFDM systems, but most works are done for OFDM-processing relays not for direct FIR filtering relays.

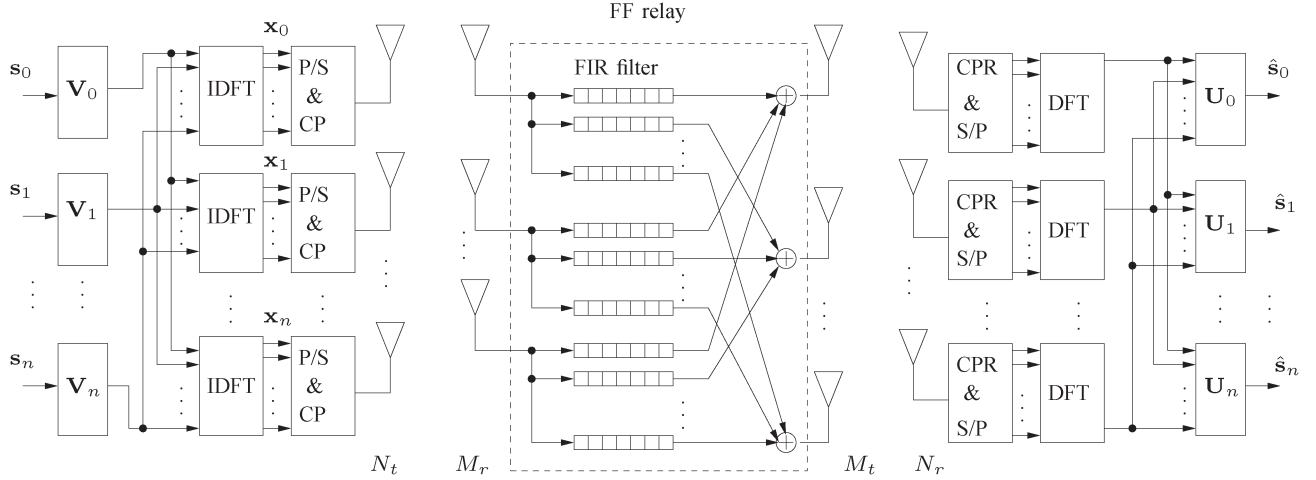


Fig. 1. System model.

precoder and decoder design when all channel information is given. The proposed alternating optimization is based on the iteration between optimal design of the FF relay for given MIMO precoding and decoding matrices and optimal design of MIMO precoding and decoding matrices for a given FF relay filter. Here, the linear MIMO transceiver design for a given FF relay filter can be cast into a conventional MIMO transceiver design problem and thus can be solved based on existing results e.g., [15]. However, the problem of optimal design of the FF relay for given MIMO precoding and decoding matrices is new. Hence, this problem is newly formulated mathematically based on the block circulant matrix theorem and reparameterization. It is shown that the FF relay design for given MIMO precoding and decoding matrices reduces to a convex quadratically constrained quadratic programming (QCQP) problem, which can be solved efficiently by a standard convex optimization solver. Numerical results show the effectiveness of the proposed FF relay design method; the FF relay designed by the proposed method shows significant performance improvement over widely-considered simple AF relays.

*Notation and Organization:* We will make use of standard notational conventions. Vectors and matrices are written in boldface with matrices in capitals. All vectors are column vectors. For a matrix  $\mathbf{X}$ ,  $\mathbf{X}^*$ ,  $\mathbf{X}^T$ ,  $\mathbf{X}^H$ ,  $\text{tr}(\mathbf{X})$ , and  $\mathbf{X}(i, j)$  indicate the complex conjugate, transpose, conjugate transpose, trace, and  $(i, j)$ -element of  $\mathbf{X}$ , respectively.  $\mathbf{X} \succeq 0$  and  $\mathbf{X} \succ 0$  mean that  $\mathbf{X}$  is positive semi-definite and that  $\mathbf{X}$  is strictly positive definite, respectively.  $\mathbf{I}_n$  stands for the identity matrix of size  $n$  (the subscript is omitted when unnecessary),  $\mathbf{I}_{m \times n}$  denotes the first  $m \times n$  submatrix of  $\mathbf{I}$ , and  $\mathbf{0}_{m \times n}$  denotes a  $m \times n$  matrix of all zero elements (the subscript is omitted when unnecessary). The notation  $\text{blkToeplitz}(\bar{\mathbf{F}}, N)$  indicates a  $NA \times (N + L_f - 1)B$  block Toeplitz matrix with  $N$  row blocks and  $[\bar{\mathbf{F}}, \mathbf{0}, \dots, \mathbf{0}]$  as its first row block, where  $\bar{\mathbf{F}} = [\mathbf{F}_0, \mathbf{F}_1, \dots, \mathbf{F}_{L_f-1}]$  is a row block composed of  $A \times B$  matrices  $\{\mathbf{F}_k\}$ .  $\text{diag}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$  means a (block) diagonal matrix with diagonal entries  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ . The notation  $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  means that  $\mathbf{x}$  is complex circularly-symmetric Gaussian distributed with mean vector  $\boldsymbol{\mu}$  and covariance matrix  $\boldsymbol{\Sigma}$ .  $\mathbb{E}\{\cdot\}$  denotes the expectation.  $\iota = \sqrt{-1}$ .

The remainder of this paper is organized as follows. The system model is described in Section II. In Section III, the joint linear MIMO transceiver and FF relay design problems for minimizing the weighted sum MSE and for maximizing the data rate are formulated and solved by using convex optimization theory and existing results. The performance of the proposed design methods is investigated in Section IV, followed by the conclusion in Section V.

## II. SYSTEM MODEL

We consider a point-to-point MIMO-OFDM system with a relay, as shown in Fig. 1, where the source has  $N_t$  transmit antennas, the relay has  $M_r$  receive antennas and  $M_t$  transmit antennas, and the destination has  $N_r$  receive antennas. The source and the destination employ MIMO-OFDM modulation and demodulation with  $N$  subcarriers, respectively, as in a conventional MIMO-OFDM system. However, we assume that the relay is a full-duplex FF relay equipped with a bank of  $M_t M_r$  FIR filters with order  $L_f$ , i.e., the relay performs FIR filtering on the incoming signals received at the  $M_r$  receive antennas at the chip rate<sup>2</sup> of the OFDM modulation and transmits the filtered signals instantaneously through the  $M_t$  transmit antennas to the destination without OFDM processing. Thus, the FF relay can be regarded as an extension of an amplify-and-forward (AF) relay and as an additional frequency-selective fading channel between the source and the destination. We assume that there is no direct link between the source and the destination and that the source-to-relay (SR) and relay-to-destination (RD) channels are multi-tap FIR filters and their state information is known to the system.

The considered baseband system model is described in detail as follows. At the source, a block of  $N$  input data vectors of size  $\Gamma \times 1$ , denoted as  $\{s_n = [s_n[1], s_n[2], \dots, s_n[\Gamma]]^T, n = 0, 1, \dots, N - 1\}$ , is processed for one OFDM symbol time. Here,  $s_n$  is the input data vector for the effective parallel flat

<sup>2</sup>In this paper, time samples after IDFT of OFDM processing are called chips. The FIR filtering is assumed to be performed at the baseband. Thus, up and down converters are necessary for FF operation and one common local oscillator (LO) at the relay is sufficient.

MIMO channel at the  $n$ -th subcarrier provided by MIMO-OFDM processing and  $\Gamma \leq \min(N_t, M_r, M_t, N_r)$  is the number of data streams for the effective flat MIMO channel at each subcarrier. We assume that each data symbol is a zero-mean independent complex Gaussian random variable with unit variance, i.e.,  $s_n[k] \sim \mathcal{CN}(0, 1)$  for  $k = 1, 2, \dots, \Gamma$  and  $n = 0, 1, \dots, N - 1$ . Let the concatenated data vector be denoted by  $\mathbf{s} = [\mathbf{s}_{N-1}^T, \mathbf{s}_{N-2}^T, \dots, \mathbf{s}_0^T]^T$ . Although MIMO precoding can be applied to the concatenated vector  $\mathbf{s}$ , such processing is inefficient and thus we assume that MIMO precoding is applied to the effective flat MIMO channel of each subcarrier separately, as in most practical MIMO-OFDM systems, with a precoding matrix  $\mathbf{V}_n$  for the  $n$ -th subcarrier MIMO channel. The MIMO precoded  $N$  symbols for each transmit antenna are collected and processed by inverse discrete Fourier transform (IDFT). By concatenating all IDFT symbols for all transmit antennas, we have the overall time-domain signal vector  $\mathbf{x}$ , given by

$$\mathbf{x} = (\mathbf{W}_N \otimes \mathbf{I}_{N_t}) \mathbf{V} \mathbf{s} \quad (1)$$

where

$$\mathbf{V} = \text{diag}(\mathbf{V}_{N-1}, \mathbf{V}_{N-2}, \dots, \mathbf{V}_0) \quad (2)$$

$$\mathbf{W}_N(k+1, l+1) = \frac{1}{\sqrt{N}} e^{j \frac{2\pi k l}{N}} \quad (3)$$

for  $k, l = 0, 1, \dots, N - 1$ , and  $\mathbf{x}$  is cyclic-prefix attached and transmitted. The cyclic prefix attached signal vector  $\mathbf{x}_{cp}$  can be expressed as

$$\mathbf{x}_{cp} = \left( \begin{bmatrix} \mathbf{I}_N \\ \mathbf{I}_{N_{cp}} \mathbf{0} \end{bmatrix} \otimes \mathbf{I}_{N_t} \right) \mathbf{x} \triangleq \mathbf{T}_{cp} \mathbf{x}, \quad (4)$$

where  $N_{cp}$  is the cyclic prefix length, and  $\mathbf{0}$  in (4) is an  $N_{cp} \times (N - N_{cp})$  all-zero matrix. We assume that the length of the overall FIR channel between the source and the destination is not larger than that of the OFDM cyclic prefix, i.e.,  $N_{cp} \geq L_f + L_r + L_g - 3$ , where  $L_f$ ,  $L_r$ , and  $L_g$  denote the SR channel length, the FIR filter order at the relay, and the RD channel length, respectively.

The transmitted signal  $\mathbf{x}_{cp}$  passes through the SR channel, the relay FIR filter, and the RD channel; is corrupted by white Gaussian noise; and is received at the destination. Then, the transmitted signal vector at the relay and the received signal vector at the destination are respectively given by

$$\mathbf{y}_t = \mathbf{R} \mathbf{F} \mathbf{x}_{cp} + \mathbf{R} \mathbf{n}_r \text{ and } \mathbf{y}_d = \mathbf{G} \mathbf{R} \mathbf{F} \mathbf{x}_{cp} + \mathbf{G} \mathbf{R} \mathbf{n}_r + \mathbf{n}_d, \quad (5)$$

where

$$\mathbf{y}_d = [\mathbf{y}_{d,N-1}^T, \mathbf{y}_{d,N-2}^T, \dots, \mathbf{y}_{d,0}^T]^T, \quad (6)$$

$$\mathbf{y}_t = [\mathbf{y}_{t,N-1}^T, \mathbf{y}_{t,N-2}^T, \dots, \mathbf{y}_{t,0}^T, \mathbf{y}_{t,-1}^T, \dots, \mathbf{y}_{t,-L_g+1}^T]^T, \quad (7)$$

$$\mathbf{x}_{cp} = [\mathbf{x}_{N-1}^T, \mathbf{x}_{N-2}^T, \dots, \mathbf{x}_0^T, \mathbf{x}_{-1}^T, \dots, \mathbf{x}_{-L_g-L_r-L_f+3}^T]^T, \quad (8)$$

$$\mathbf{n}_r = \left[ \mathbf{n}_{r,N-1}^T, \mathbf{n}_{r,N-2}^T, \dots, \mathbf{n}_{r,0}^T, \dots, \mathbf{n}_{r,-L_g-L_r+2}^T \right]^T, \quad (9)$$

$$\mathbf{n}_d = [\mathbf{n}_{d,N-1}^T, \mathbf{n}_{d,N-2}^T, \dots, \mathbf{n}_{d,0}^T]^T, \quad (10)$$

$$\mathbf{G} = \text{blkToeplitz}(\overline{\mathbf{G}}, N), \mathbf{R} = \text{blkToeplitz}(\overline{\mathbf{R}}, N + L_g - 1),$$

$$\mathbf{F} = \text{blkToeplitz}(\overline{\mathbf{F}}, N + L_g + L_r - 2), \quad (11)$$

$$\overline{\mathbf{G}} = [\mathbf{G}_0, \mathbf{G}_1, \dots, \mathbf{G}_{L_g-1}], \overline{\mathbf{R}} = [\mathbf{R}_0, \mathbf{R}_1, \dots, \mathbf{R}_{L_r-1}],$$

$$\overline{\mathbf{F}} = [\mathbf{F}_0, \mathbf{F}_1, \dots, \mathbf{F}_{L_f-1}]. \quad (12)$$

Here,  $\mathbf{y}_{d,k}$  and  $\mathbf{n}_{d,k}$  are  $N_r \times 1$  vectors;  $\mathbf{y}_{t,k}$  is a  $M_t \times 1$  vector;  $\mathbf{x}_k$  is a  $N_t \times 1$  vector;  $\mathbf{n}_{r,k}$  is a  $M_r \times 1$  vector;  $\mathbf{G}_k$  is a  $N_r \times M_t$  matrix;  $\mathbf{R}_k$  is a  $M_t \times M_r$  matrix; and  $\mathbf{F}_k$  is a  $M_r \times N_t$  matrix. The entries of the noise vectors,  $\mathbf{n}_{r,k}$  and  $\mathbf{n}_{d,k}$ , are independently and identically distributed (i.i.d.) Gaussian with  $\mathbf{n}_{r,k}[i] \stackrel{i.i.d.}{\sim} \mathcal{CN}(0, \sigma_r^2)$  and  $\mathbf{n}_{d,k}[i] \stackrel{i.i.d.}{\sim} \mathcal{CN}(0, \sigma_d^2)$ . Then, the (cyclic-prefix portion removed)  $N$ -point vector DFT of the received vector at the destination is given by

$$\mathbf{y} = (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{G} \mathbf{R} \mathbf{F} \mathbf{x}_{cp} + (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{G} \mathbf{R} \mathbf{n}_r$$

$$+ (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{n}_d,$$

$$= (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{G} \mathbf{R} \mathbf{F} \mathbf{T}_{cp} (\mathbf{W}_N \otimes \mathbf{I}_{N_t}) \mathbf{V} \mathbf{s}$$

$$+ (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{G} \mathbf{R} \mathbf{n}_r + (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{n}_d,$$

$$= (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{H}_c (\mathbf{W}_N \otimes \mathbf{I}_{N_t}) \mathbf{V} \mathbf{s}$$

$$+ (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{G} \mathbf{R} \mathbf{n}_r + (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{n}_d, \quad (13)$$

$$= \mathbf{D} \mathbf{V} \mathbf{s} + (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{G} \mathbf{R} \mathbf{n}_r + (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{n}_d, \quad (14)$$

where  $\mathbf{y} = [\mathbf{y}_{N-1}^T, \mathbf{y}_{N-2}^T, \dots, \mathbf{y}_0^T]^T$ ,  $\mathbf{y}_n$  is a  $N_r \times 1$  received signal vector at the  $n$ -th subcarrier,  $\mathbf{W}_N^H$  is the normalized DFT matrix of size  $N$ ,  $\mathbf{H}_c$  is a  $NN_r \times NN_t$  block circulant matrix generated from the block Toeplitz overall channel matrix  $\mathbf{G} \mathbf{R} \mathbf{F}$  from the source to the destination, and  $\mathbf{D} = (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{H}_c (\mathbf{W}_N \otimes \mathbf{I}_{N_t})$  is a block diagonal matrix generated by the block circulant matrix theorem described in the next section. The  $n$ -th subcarrier output of the  $N$ -point vector DFT is processed by a MIMO decoding matrix  $\mathbf{U}_n$  of size  $\Gamma \times N_r$  to yield an estimate  $\hat{\mathbf{s}}_n$  of  $\mathbf{s}_n$ . The overall receiver processing for all the subcarrier channels can be expressed as

$$\hat{\mathbf{s}} = \mathbf{U} \mathbf{D} \mathbf{V} \mathbf{s} + \mathbf{U} (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{G} \mathbf{R} \mathbf{n}_r + \mathbf{U} (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{n}_d$$

where  $\mathbf{U} = \text{diag}(\mathbf{U}_{N-1}, \mathbf{U}_{N-2}, \dots, \mathbf{U}_0)$ .

#### A. Derivation of the Subcarrier Channel and Mean Square Error

To facilitate the optimization problem formulation in the next section, we need to derive an explicit expression for the received signal vector  $\mathbf{y}_n$ ,  $n = 0, 1, \dots, N - 1$ , at the  $n$ -th subcarrier.

*Lemma 1:* If  $\mathbf{H}_c$  is a block circulant matrix with  $\mathbf{K} = [\mathbf{H}_0, \mathbf{H}_1, \dots, \mathbf{H}_{N-1}]$  as its first row block, then it is block-diagonalizable as

$$\mathbf{\Lambda}_b = (\mathbf{W}_N^H \otimes \mathbf{I}_{N_r}) \mathbf{H}_c (\mathbf{W}_N \otimes \mathbf{I}_{N_t})$$

where  $\Lambda_b$  is a block diagonal matrix defined as

$$\Lambda_b = \begin{bmatrix} \mathbf{K}(\sqrt{N}\mathbf{w}_{N-1}^H \otimes \mathbf{I}_{N_t})^T & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \mathbf{K}(\sqrt{N}\mathbf{w}_0^H \otimes \mathbf{I}_{N_t})^T \end{bmatrix}$$

with  $\sqrt{N}\mathbf{w}_k^H$  denoting the  $-(k-N)$ -th row of the DFT matrix  $\sqrt{N}\mathbf{W}_N^H$ , and

$$\mathbf{K}(\sqrt{N}\mathbf{w}_k^H \otimes \mathbf{I}_{N_t})^T = \sum_{n=0}^{N-1} \mathbf{H}_n e^{-j2\pi \frac{n(N-k-1)}{N}}.$$

*Proof:* In [20], it is shown that a circulant matrix can be diagonalized by a DFT matrix. This can easily be extended to the block circulant case [21].  $\square$

By lemma 1, to derive the diagonal blocks of  $\mathbf{D}$  in (14), we only need to know the first row block of  $\mathbf{H}_c$  in (13). Let the first row block of the RD channel matrix  $\mathbf{G}$  be denoted by a  $N_r \times M_t(N + L_g - 1)$  matrix  $\tilde{\mathbf{G}} = [\mathbf{G}_0, \mathbf{G}_1, \dots, \mathbf{G}_{L_g-1}, \mathbf{0}, \dots, \mathbf{0}]$ . Then, the first row block of the effective channel filtering matrix  $\mathbf{GRF}$  is given by  $\tilde{\mathbf{GRF}}$ . Note that the cyclic prefix adding and removing operations make  $\mathbf{GRF}$  into the block circulant matrix  $\mathbf{H}_c$  by truncating out the blocks of  $\mathbf{GRF}$  outside the first  $N \times N$  blocks and by moving the lower  $(L_g + L_r + L_f - 3) \times (L_g + L_r + L_f - 3)$  blocks of the truncated part to the lower left of the untruncated  $N \times N$  block matrix, where each block is a  $N_r \times N_t$  matrix. Therefore, the first row block  $\tilde{\mathbf{H}}_c$  of  $\mathbf{H}_c$  is simply the first  $N$  blocks of  $\tilde{\mathbf{GRF}}$ , given by

$$\tilde{\mathbf{H}}_c = \tilde{\mathbf{GRF}}\mathbf{T} \text{ and } \mathbf{T} = \begin{bmatrix} \mathbf{I}_{NN_t} \\ \mathbf{0}_{(L_f+L_r+L_g-3)N_t \times NN_t} \end{bmatrix} \quad (15)$$

where  $\mathbf{T}$  is a truncation matrix for truncating out the remaining blocks of  $\tilde{\mathbf{GRF}}$  except the first  $N$  column blocks. By using the first row block  $\tilde{\mathbf{H}}_c$  and Lemma 1, we obtain the diagonal blocks of  $\mathbf{D}$  as

$$\mathbf{D} = \text{diag} \left( \tilde{\mathbf{H}}_c(\sqrt{N}\mathbf{w}_{N-1}^H \otimes \mathbf{I}_{N_t})^T, \tilde{\mathbf{H}}_c(\sqrt{N}\mathbf{w}_{N-2}^H \otimes \mathbf{I}_{N_t})^T, \dots, \tilde{\mathbf{H}}_c(\sqrt{N}\mathbf{w}_0^H \otimes \mathbf{I}_{N_t})^T \right). \quad (16)$$

Based on (15) and (16), the received signal vector on the  $n$ -th subcarrier at the destination is expressed as

$$\begin{aligned} \mathbf{y}_n &= \sqrt{N}\tilde{\mathbf{GRF}}\mathbf{T}\mathbf{W}_{t,n}^T \mathbf{V}_n \mathbf{s}_n + \mathcal{W}_{r,n} \mathbf{GRn}_{r,n} + \mathcal{W}_{r,n} \mathbf{nd}_{n}, \\ &= \hat{\mathbf{y}}_n + \mathbf{z}_n, \end{aligned} \quad (17)$$

where  $\mathcal{W}_{t,n} = \mathbf{w}_n^H \otimes \mathbf{I}_{N_t}$ ,  $\mathcal{W}_{r,n} = \mathbf{w}_n^H \otimes \mathbf{I}_{N_r}$ ,  $\hat{\mathbf{y}}_n = \sqrt{N}\tilde{\mathbf{GRF}}\mathbf{T}\mathbf{W}_{t,n}^T \mathbf{V}_n \mathbf{s}_n$ , and  $\mathbf{z}_n = \mathcal{W}_{r,n} \mathbf{GRn}_{r,n} + \mathcal{W}_{r,n} \mathbf{nd}_{n}$ . This received signal vector  $\mathbf{y}_n$  is filtered by the MIMO decoding matrix  $\mathbf{U}_n$  and its output is given by

$$\begin{aligned} \hat{\mathbf{s}}_n &= \sqrt{N}\mathbf{U}_n \tilde{\mathbf{GRF}}\mathbf{T}\mathbf{W}_{t,n}^T \mathbf{V}_n \mathbf{s}_n + \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{GRn}_{r,n} \\ &\quad + \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{nd}_{n}. \end{aligned} \quad (18)$$

Finally, the weighted MSE between  $\mathbf{s}_n$  and  $\hat{\mathbf{s}}_n$  is given by

$$\begin{aligned} &\text{tr}(\Theta_n \mathcal{M}_n) \\ &= \text{tr}(\Theta_n \mathbb{E} \{ (\hat{\mathbf{s}}_n - \mathbf{s}_n)(\hat{\mathbf{s}}_n - \mathbf{s}_n)^H \}), \\ &= \text{tr}(\Theta_n \mathbb{E} \{ (\mathbf{U}_n \mathbf{y}_n - \mathbf{s}_n)(\mathbf{U}_n \mathbf{y}_n - \mathbf{s}_n)^H \}), \\ &= \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{U}_n \mathbf{y}_n \mathbf{y}_n^H \mathbf{U}_n^H \}) - \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{s}_n \mathbf{y}_n^H \mathbf{U}_n^H \}) \\ &\quad - \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{U}_n \mathbf{y}_n \mathbf{s}_n^H \}) + \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{s}_n \mathbf{s}_n^H \}), \\ &\stackrel{(a)}{=} \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{U}_n \hat{\mathbf{y}}_n \hat{\mathbf{y}}_n^H \mathbf{U}_n^H \}) + \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{U}_n \mathbf{z}_n \mathbf{z}_n^H \mathbf{U}_n^H \}) \\ &\quad - \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{s}_n \hat{\mathbf{y}}_n^H \mathbf{U}_n^H \}) - \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{U}_n \hat{\mathbf{y}}_n \mathbf{s}_n^H \}) \\ &\quad + \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{s}_n \mathbf{s}_n^H \}), \end{aligned} \quad (19)$$

where  $\mathcal{M}_n \triangleq \mathbb{E} \{ (\hat{\mathbf{s}}_n - \mathbf{s}_n)(\hat{\mathbf{s}}_n - \mathbf{s}_n)^H \}$  is the MSE matrix at the  $n$ -th subcarrier,  $\Theta_n$  is a  $\Gamma \times \Gamma$  diagonal positive definite weight matrix, and (a) follows from (17).

### III. PROBLEM FORMULATION AND THE PROPOSED DESIGN METHOD

In this section, we consider optimal design of the FIR MIMO relay filter  $\{\mathbf{R}_0, \mathbf{R}_1, \dots, \mathbf{R}_{L_r-1}\}$  and the linear MIMO precoders and decoders  $\{\mathbf{V}_n, \mathbf{U}_n, n = 0, 1, \dots, N-1\}$ . Among several optimality criteria, we first consider the minimization of the weighted sum mean-square-error (MSE) for given weight matrices, and then consider the rate maximization via the weighted sum MSE minimization based on the fact that the rate maximization for MIMO channels is equivalent to the weighted MSE minimization with properly chosen weight matrices  $\{\Theta_n\}$  [15]. (Here, the summation is across the subcarrier channels.) The first problem is formally stated as follows.

*Problem 1:* For given weight matrices  $\{\Theta_n\}_{n=0}^{N-1}$ , SR channel  $\mathbf{F}$ , RD channel  $\mathbf{G}$ , FF relay filter order  $L_r$ , maximum source transmit power  $P_{s,\max}$ , and maximum relay transmit power  $P_{r,\max}$ , optimize the MIMO precoding matrix  $\mathbf{V} = \text{diag}(\mathbf{V}_0, \dots, \mathbf{V}_{N-1})$ , the relay filter  $\bar{\mathbf{R}}$ , and the MIMO decoding matrix  $\mathbf{U} = \text{diag}(\mathbf{U}_0, \dots, \mathbf{U}_{N-1})$  in order to minimize the weighted sum MSE:

$$\begin{aligned} &\min_{\mathbf{V}, \bar{\mathbf{R}}, \mathbf{U}} \sum_{n=0}^{N-1} \text{tr}(\Theta_n \mathcal{M}_n) \\ &\text{s.t. } \text{tr}(\mathbf{V}\mathbf{V}^H) \leq P_{s,\max} \text{ and } \text{tr}(\mathbf{y}_t \mathbf{y}_t^H) \leq P_{r,\max}. \end{aligned} \quad (20)$$

Note that Problem 1 is a complicated non-convex optimization problem, which does not yield an easy solution. To circumvent the difficulty in the joint optimization, we approach the problem based on alternating optimization. That is, we first optimize the relay filter for given MIMO precoding and decoding matrices under the power constraints. Then, with the obtained relay filter we optimize the MIMO precoding and decoding matrices. Problem 1 is solved in this alternating fashion until the iteration converges. A solution to each step is provided in the following subsections.

#### A. Relay Filter Optimization

Whereas the linear MIMO precoder  $\mathbf{V}_n$  and MIMO decoder  $\mathbf{U}_n$  are applied to each subcarrier channel separately, the relay

filter affects all the subcarrier channels simultaneously since the FF relay does not perform OFDM processing. Here, we consider the relay filter optimization for given MIMO precoding and decoding matrices, and the problem is formulated as follows.

*Problem 1-1:* For given weight matrices  $\{\Theta_n\}$ , SR channel  $\mathbf{F}$ , RD channel  $\mathbf{G}$ , FF relay filter order  $L_r$ , MIMO precoder  $\mathbf{V}$ , MIMO decoder  $\mathbf{U}$ , and maximum relay transmit power  $P_{r,\max}$ , optimize the relay filter  $\bar{\mathbf{R}}$  in order to minimize the weighted sum MSE:

$$\min_{\bar{\mathbf{R}}} \sum_{n=0}^{N-1} \text{tr}(\Theta_n \mathcal{M}_n) \text{ s.t. } \text{tr}(\mathbf{y}_t \mathbf{y}_t^H) \leq P_{r,\max}. \quad (21)$$

To solve Problem 1-1, we first need to express each term in (21) as a function of the design variable  $\bar{\mathbf{R}}$ . Note that the relay block-Toeplitz filtering matrix  $\mathbf{R}$  is redundant since the true design variable  $\bar{\mathbf{R}}$  is embedded in the block Toeplitz structure of  $\mathbf{R}$ . (See (11).) Hence, taking  $\mathbf{R}$  as the design variable directly is inefficient and we need reparameterization of the weighted MSE in terms of  $\bar{\mathbf{R}}$ . This is possible through successive manipulation of the terms constructing the weight MSE shown in (19). First, using similar techniques to those used in [7], we can express the first term of (19) in terms of  $\bar{\mathbf{R}}$  as follows:

$$\begin{aligned} & \text{tr}(\Theta_n \mathbb{E}[\mathbf{U}_n \hat{\mathbf{y}}_n \hat{\mathbf{y}}_n^H \mathbf{U}_n^H]) \\ &= N \text{tr} \left( \Theta_n \mathbb{E} \left[ \mathbf{U}_n \tilde{\mathbf{G}} \mathbf{R} \mathbf{F}^T \mathbf{W}_{t,n}^T \mathbf{V}_n \mathbf{s}_n \right. \right. \\ & \quad \left. \left. \mathbf{s}_n^H \mathbf{V}_n^H \mathbf{W}_{t,n}^* \mathbf{T}^H \mathbf{F}^H \mathbf{R}^H \tilde{\mathbf{G}}^H \mathbf{U}_n^H \right] \right), \\ & \stackrel{(a)}{=} N \text{tr} \left( \mathbf{V}_n^H \mathbf{W}_{t,n}^* \mathbf{T}^H \mathbf{F}^H \mathbf{R}^H \tilde{\mathbf{G}}^H \mathbf{U}_n^H \Theta_n \right. \\ & \quad \left. \mathbf{U}_n \tilde{\mathbf{G}} \mathbf{R} \mathbf{F}^T \mathbf{W}_{t,n}^T \mathbf{V}_n \mathbb{E}[\mathbf{s}_n \mathbf{s}_n^H] \right), \\ &= N \text{tr} \left( \mathbf{V}_n^H \mathbf{W}_{t,n}^* \mathbf{T}^H \mathbf{F}^H \mathbf{R}^H \tilde{\mathbf{G}}^H \mathbf{U}_n^H \Theta_n \right. \\ & \quad \left. \mathbf{U}_n \tilde{\mathbf{G}} \mathbf{R} \mathbf{F}^T \mathbf{W}_{t,n}^T \mathbf{V}_n \right), \\ &= N \text{tr} \left( \Theta_n^{\frac{1}{2}} \mathbf{U}_n \tilde{\mathbf{G}} \mathbf{R} \mathbf{K}_n \mathbf{R}^H \tilde{\mathbf{G}}^H \mathbf{U}_n^H \Theta_n^{\frac{1}{2}} \right), \\ & \stackrel{(b)}{=} N \left[ \text{vec}(\mathbf{R}^T \tilde{\mathbf{G}}^T \mathbf{U}_n^T \Theta_n^{\frac{1}{2}}) \right]^T \bar{\mathbf{K}}_n \left[ \text{vec}(\mathbf{R}^T \tilde{\mathbf{G}}^T \mathbf{U}_n^T \Theta_n^{\frac{1}{2}}) \right]^*, \\ & \stackrel{(c)}{=} N \left[ \text{vec}(\mathbf{R}^T) \right]^T \left( \Theta_n^{\frac{1}{2}} \mathbf{U}_n \tilde{\mathbf{G}} \otimes \mathbf{I}_Q \right)^T \bar{\mathbf{K}}_n \left( \Theta_n^{\frac{1}{2}} \mathbf{U}_n \tilde{\mathbf{G}} \otimes \mathbf{I}_Q \right)^* \\ & \quad \times \left[ \text{vec}(\mathbf{R}^T) \right]^*, \\ & \stackrel{(d)}{=} N \mathbf{r}^T \mathbf{E}_1 \left( \Theta_n^{\frac{1}{2}} \mathbf{U}_n \tilde{\mathbf{G}} \otimes \mathbf{I}_Q \right)^T \bar{\mathbf{K}}_n \left( \Theta_n^{\frac{1}{2}} \mathbf{U}_n \tilde{\mathbf{G}} \otimes \mathbf{I}_Q \right)^* \mathbf{E}_1^H \mathbf{r}^*, \\ &= \mathbf{r}^H \mathbf{Q}_{1,n} \mathbf{r}, \end{aligned} \quad (22)$$

where

$$\begin{aligned} \mathbf{K}_n &= \mathbf{F}^T \mathbf{W}_{t,n}^T \mathbf{V}_n \mathbf{V}_n^H \mathbf{W}_{t,n}^* \mathbf{T}^H \mathbf{F}^H; \\ \bar{\mathbf{K}}_n &= \mathbf{I}_\Gamma \otimes \mathbf{K}_n; \quad \mathbf{I}_Q = \mathbf{I}_{(N+L_r+L_g-2)M_r}; \quad \mathbf{r} = \text{vec}(\bar{\mathbf{R}}^T); \\ \mathbf{Q}_{1,n} &= N \mathbf{E}_1^* \left( \Theta_n^{\frac{1}{2}} \mathbf{U}_n \tilde{\mathbf{G}} \otimes \mathbf{I}_Q \right)^H \bar{\mathbf{K}}_n \left( \Theta_n^{\frac{1}{2}} \mathbf{U}_n \tilde{\mathbf{G}} \otimes \mathbf{I}_Q \right) \mathbf{E}_1^T; \end{aligned}$$

and  $\mathbf{E}_1$  is defined in Appendix A. Here, (a) holds due to  $\text{tr}(\mathbf{UBC}) = \text{tr}(\mathbf{CUB})$ ; (b) holds due to  $\text{tr}(\mathbf{XK}_n \mathbf{X}^H) = \text{vec}(\mathbf{X}^T)^T \bar{\mathbf{K}}_n \text{vec}(\mathbf{X}^T)^*$ ; (c) holds due to the Kronecker product identity,  $\text{vec}(\mathbf{IBC}) = (\mathbf{C}^T \otimes \mathbf{I}) \text{vec}(\mathbf{B})$ ; and (d) is obtained

because  $\mathbf{R} = \text{blkToeplitz}(\bar{\mathbf{R}}, N + L_g - 1)$  and  $\text{vec}(\mathbf{R}^T) = \mathbf{E}_1^T \mathbf{r}$ . In a similar way, the remaining terms of (19) and the relay power constraint can also be represented as functions of the design variable  $\mathbf{r}$ . That is, the second term of (19) can be rewritten as

$$\begin{aligned} & \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{U}_n \mathbf{z}_n \mathbf{z}_n^H \mathbf{U}_n^H \}) \\ &= \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{G} \mathbf{R} \mathbf{n}_{r,n} \mathbf{n}_{r,n}^H \mathbf{R}^H \mathbf{G}^H \mathcal{W}_{r,n}^H \mathbf{U}_n^H \}) \\ & \quad + \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{n}_{d,n} \mathbf{n}_{d,n}^H \mathcal{W}_{r,n}^H \mathbf{U}_n^H \}), \\ &= \text{tr}(\mathbf{R}^H \mathbf{G}^H \mathcal{W}_{r,n}^H \mathbf{U}_n^H \Theta_n \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{G} \mathbf{R} \mathbb{E} \{ \mathbf{n}_{r,n} \mathbf{n}_{r,n}^H \}) \\ & \quad + \text{tr}(\Theta_n \mathbf{U}_n \mathcal{W}_{r,n} \mathbb{E} \{ \mathbf{n}_{d,n}^H \mathbf{n}_{d,n} \} \mathcal{W}_{r,n}^H \mathbf{U}_n^H), \\ &= \sigma_r^2 \text{tr}(\mathbf{R}^H \mathbf{M}_n \mathbf{R}) + \sigma_d^2 \text{tr}(\Theta_n \mathbf{U}_n \mathcal{W}_{r,n} \mathcal{W}_{r,n}^H \mathbf{U}_n^H), \\ &= \sigma_r^2 \text{tr}(\mathbf{R}^H \mathbf{M}_n \mathbf{R}) \\ & \quad + \sigma_d^2 \text{tr}(\Theta_n \mathbf{U}_n (\mathbf{w}_n^H \otimes \mathbf{I}_{N_r}) (\mathbf{w}_n \otimes \mathbf{I}_{N_r}) \mathbf{U}_n^H), \\ & \stackrel{(a)}{=} \sigma_r^2 \text{vec}(\mathbf{R})^H \bar{\mathbf{M}}_n \text{vec}(\mathbf{R}) \\ & \quad + \sigma_d^2 \text{tr}(\Theta_n \mathbf{U}_n (\mathbf{w}_n^H \mathbf{w}_n \otimes \mathbf{I}_{N_r}) \mathbf{U}_n^H), \\ & \stackrel{(b)}{=} \sigma_r^2 \mathbf{r}^H \mathbf{E}_2 \bar{\mathbf{M}}_n \mathbf{E}_2^H \mathbf{r} + \sigma_d^2 \text{tr}(\Theta_n \mathbf{U}_n \mathbf{U}_n^H), \\ &= \mathbf{r}^H \mathbf{Q}_{2,n} \mathbf{r} + c_n, \end{aligned} \quad (23)$$

where

$$\begin{aligned} \mathbf{M}_n &= \mathbf{G}^H \mathcal{W}_{r,n}^H \mathbf{U}_n^H \Theta_n \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{G}, \\ \bar{\mathbf{M}}_n &= \mathbf{I}_{(N+L_g+L_r-2)M_r} \otimes \mathbf{M}_n, \quad \mathbf{Q}_{2,n} = \sigma_r^2 \mathbf{E}_2 \bar{\mathbf{M}}_n \mathbf{E}_2^H, \\ c_n &= \sigma_d^2 \text{tr}(\Theta_n \mathbf{U}_n \mathbf{U}_n^H) \end{aligned}$$

and  $\mathbf{E}_2$  is defined in Appendix A. Here, (a) follows from the Kronecker product identity  $(\mathbf{UB} \otimes \mathbf{CD}) = (\mathbf{U} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D})$ , and (b) is obtained due to  $\text{vec}(\mathbf{R})^H = \mathbf{r}^H \mathbf{E}_2$ . The third term of (19) can be rewritten as

$$\begin{aligned} & \text{tr}(\Theta_n \mathbb{E} \{ \mathbf{s}_n \hat{\mathbf{y}}_n^H \mathbf{U}_n^H \}) \\ &= \sqrt{N} \text{tr} \left( \Theta_n \mathbb{E} \{ \mathbf{s}_n \mathbf{s}_n^H \} \mathbf{V}_n^H \mathcal{W}_{t,n}^* \mathbf{T}^H \mathbf{F}^H \mathbf{R}^H \tilde{\mathbf{G}}^H \mathbf{U}_n^H \right), \\ &= \sqrt{N} \text{tr} \left( \Theta_n \mathbf{V}_n^H \mathcal{W}_{t,n}^* \mathbf{T}^H \mathbf{F}^H \mathbf{R}^H \tilde{\mathbf{G}}^H \mathbf{U}_n^H \right), \\ &= \sqrt{N} \text{tr} \left( \mathbf{R}^H \tilde{\mathbf{G}}^H \mathbf{U}_n^H \Theta_n \mathbf{V}_n^H \mathcal{W}_{t,n}^* \mathbf{T}^H \mathbf{F}^H \right), \\ &= \sqrt{N} \text{vec}(\mathbf{R})^H \text{vec} \left( \tilde{\mathbf{G}}^H \mathbf{U}_n^H \Theta_n \mathbf{V}_n^H \mathcal{W}_{t,n}^* \mathbf{T}^H \mathbf{F}^H \right), \\ &= \sqrt{N} \mathbf{r}^H \mathbf{E}_2 \text{vec} \left( \tilde{\mathbf{G}}^H \mathbf{U}_n^H \Theta_n \mathbf{V}_n^H \mathcal{W}_{t,n}^* \mathbf{T}^H \mathbf{F}^H \right), \\ &= \mathbf{r}^H \mathbf{q}_n, \end{aligned} \quad (24)$$

where  $\mathbf{q}_n = \sqrt{N} \mathbf{E}_2 \text{vec}(\tilde{\mathbf{G}}^H \mathbf{U}_n^H \Theta_n \mathbf{V}_n^H \mathcal{W}_{t,n}^* \mathbf{T}^H \mathbf{F}^H)$ . Finally, the relay transmit power can be rewritten as

$$\begin{aligned} & \mathbb{E} \{ \text{tr}(\mathbf{y}_t \mathbf{y}_t^H) \} \\ &= \text{tr}(\mathbf{R} \mathbf{F}^T \mathbf{C}_p (\mathbf{W}_N \otimes \mathbf{I}_{N_t}) \mathbf{V} \mathbb{E} \{ \mathbf{s} \mathbf{s}^H \} \mathbf{V}^H \\ & \quad (\mathbf{W}_N^H \otimes \mathbf{I}_{N_t}) \mathbf{T}_p^H \mathbf{F}^H \mathbf{R}^H) + \text{tr}(\mathbf{R} \mathbb{E} \{ \mathbf{n}_r \mathbf{n}_r^H \} \mathbf{R}^H), \\ &= \text{tr}(\mathbf{R} \mathbf{F}^T \mathbf{C}_p (\mathbf{W}_N \otimes \mathbf{I}_{N_t}) \mathbf{V} \mathbf{V}^H (\mathbf{W}_N^H \otimes \mathbf{I}_{N_t}) \mathbf{T}_p^H \mathbf{F}^H \mathbf{R}^H) \\ & \quad + \sigma_r^2 \text{tr}(\mathbf{R} \mathbf{R}^H), \\ &= \text{tr}(\mathbf{R} \mathbf{I} \mathbf{R}^H), \\ &= \text{vec}(\mathbf{R}^T)^T \bar{\mathbf{I}} \text{vec}(\mathbf{R}^T)^*, \\ &= \mathbf{r}^H \bar{\mathbf{I}} \mathbf{r}, \end{aligned} \quad (25)$$

where

$$\begin{aligned} \mathbf{\Pi} &= \mathbf{F}\mathbf{T}_{cp}(\mathbf{W}_N \otimes \mathbf{I}_{N_t})\mathbf{V}\mathbf{V}^H(\mathbf{W}_N^H \otimes \mathbf{I}_{N_t})\mathbf{T}_{cp}^H\mathbf{F}^H + \sigma_r^2\mathbf{I}, \\ \bar{\mathbf{\Pi}} &= \mathbf{I}_{(N+L_g-1)M_t} \otimes \mathbf{\Pi}, \text{ and } \tilde{\mathbf{\Pi}} = \mathbf{E}_1^*\bar{\mathbf{\Pi}}^*\mathbf{E}_1^T. \end{aligned}$$

Based on (22)–(25), the weighted MSE for the  $n$ -th subcarrier channel is expressed as

$$\text{tr}(\Theta_n \mathcal{M}_n) = \mathbf{r}^H \mathbf{Q}_n \mathbf{r} - \mathbf{r}^H \mathbf{q}_n - \mathbf{q}_n^H \mathbf{r} + z_n \quad (26)$$

where  $\mathbf{Q}_n = \mathbf{Q}_{1,n} + \mathbf{Q}_{2,n}$  and  $z_n = c_n + \text{tr}(\Theta_n)$ , and Problem 1-1 is reformulated as

$$\begin{aligned} \min_{\mathbf{r}} \quad & \mathbf{r}^H \mathbf{Q} \mathbf{r} - \mathbf{r}^H \mathbf{q} - \mathbf{q}^H \mathbf{r} + z \\ \text{s.t.} \quad & \mathbf{r}^H \tilde{\mathbf{\Pi}} \mathbf{r} \leq P_{r,\max} \end{aligned} \quad (27)$$

where  $\mathbf{Q} = \sum_{n=1}^N \mathbf{Q}_n$ ,  $\mathbf{q} = \sum_{n=1}^N \mathbf{q}_n$ , and  $z = \sum_{n=1}^N z_n$ .

The key point in the derivation of (27) is that Problem 1-1 reduces to a convex optimization problem. One can easily see that  $\mathbf{Q}_{1,n}$  and  $\mathbf{Q}_{2,n}$  are positive semi-definite from (22) and (23). Hence,  $\mathbf{Q}$  is positive semi-definite since it is the sum of positive semi-definite matrices. One can also see from (25) that  $\tilde{\mathbf{\Pi}}$  is positive definite. Thus, the problem (27) is a convex problem due to the positive semi-definiteness of  $\mathbf{Q}$  and  $\tilde{\mathbf{\Pi}}$  [22]. Problem (27) is a *quadratically constrained quadratic programming (QCQP) problem* and its optimal solution can be obtained efficiently by using the standard interior point method for convex optimization [22]–[25].

## B. MIMO Precoder and Decoder Optimization

Now consider the design of the MIMO precoding and decoding matrices  $\{(\mathbf{V}_n, \mathbf{U}_n), n = 0, 1, \dots, N-1\}$  for a given relay FIR filter. Note that when the transmit power  $P_{n,\max} (\geq \text{tr}(\mathbf{V}_n \mathbf{V}_n^H))$  for every  $n$  and the relay filter are given, the problem simply reduces to  $N$  independent problems of designing the MIMO precoder  $\mathbf{V}_n$  and the MIMO decoder  $\mathbf{U}_n$  for the  $n$ -th subcarrier (flat) MIMO channel for  $n = 0, \dots, N-1$ , as in typical MIMO-OFDM systems. This is because we have an independent MIMO channel per subcarrier owing to MIMO-OFDM processing. However, here we have an additional freedom to distribute the total source transmit power  $P_{s,\max}$  to  $N$  subcarriers such that  $P_{s,\max} = \sum_{n=0}^{N-1} P_{n,\max}$ , and should take this overall power allocation into consideration. We solve the problem by separating the power allocation problem out and applying the existing result [15] to the problem. First, consider the MIMO precoder and decoder design problem when the transmit power  $P_{n,\max}$  for every  $n$  and the relay filter are given:

*Problem 1-2:* For given weight matrices  $\{\Theta_n\}$ , maximum per-subcarrier transmit power  $P_{n,\max}$  for  $n = 0, 1, \dots, N-1$ , SR channel  $\mathbf{F}$ , RD channel  $\mathbf{G}$ , relay filtering matrix  $\mathbf{R}$ , jointly optimize  $(\mathbf{V}_n, \mathbf{U}_n)$  in order to minimize the weighted MSE at the  $n$ -th subcarrier MIMO channel:

$$\begin{aligned} \min_{\mathbf{V}_n, \mathbf{U}_n} \quad & \text{tr}(\Theta_n \mathcal{M}_n) \\ \text{s.t.} \quad & \text{tr}(\mathbf{V}_n \mathbf{V}_n^H) \leq P_{n,\max}, \text{ for } n = 0, 1, \dots, N-1. \end{aligned} \quad (28)$$

Problem 1-2 has already been solved and the optimal transceiver structure for Problem 1-2 is available in [15] and [26]. It is shown in [15] that the optimal transmit precoder and receive decoder diagonalize the MIMO channel into eigen-subchannels for any weight matrix. Lemma 1 and Theorem 1 of [15] provide the optimal transmit precoder  $\mathbf{V}_n$  and receive decoder  $\mathbf{U}_n$ , and the solution can be expressed as  $\mathbf{V}_n = \tilde{\mathbf{V}}_n \tilde{\mathbf{P}}_n$ , where  $\tilde{\mathbf{V}}_n^H \tilde{\mathbf{V}}_n = \mathbf{I}_\Gamma$  and  $\tilde{\mathbf{P}}_n$  is a diagonal matrix with non-negative entries s.t.  $\text{tr}(\tilde{\mathbf{P}}_n^2) = P_{n,\max}$  determining the transmit power of each of  $\Gamma$  data streams of the  $n$ -th subcarrier MIMO channel. (Please see [15] for details.)

Note that the solution to Problem 1-2 only optimizes the power allocation within the multiple data streams for each subcarrier when the transmit power is allocated to each subcarrier. Now, consider the problem of total source power allocation  $P_{s,\max}$  to the subcarrier channels. Here, we exploit the *diagonalizing* property [15] of the solution to Problem 1-2, take the direction information only for the transmit precoder from the solution to Problem 1-2, and apply alternating optimization. That is, when the relay filtering matrix  $\mathbf{R}$  from Problem 1-1 and the normalized transmit MIMO precoders  $\{\tilde{\mathbf{V}}_n\}$  and the receive MIMO decoders  $\{\mathbf{U}_n\}$  from Problem 1-2 are given, each subcarrier MIMO channel is diagonalized into eigen-subchannels. Thus, the effective parallel MIMO channel (18) for the  $n$ -th subcarrier is rewritten as

$$\begin{aligned} \hat{\mathbf{s}}_n &= \sqrt{N} \mathbf{U}_n \tilde{\mathbf{G}} \mathbf{R} \mathbf{F}^T \mathbf{W}_{t,n}^T \mathbf{V}_n \mathbf{s}_n + \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{G} \mathbf{R} \mathbf{n}_{r,n} \\ &\quad + \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{n}_{d,n}, \\ &= \sqrt{N} \mathbf{U}_n \tilde{\mathbf{G}} \mathbf{R} \mathbf{F}^T \mathbf{W}_{t,n}^T \tilde{\mathbf{V}}_n \tilde{\mathbf{P}}_n \mathbf{s}_n + \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{G} \mathbf{R} \mathbf{n}_{r,n} \\ &\quad + \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{n}_{d,n}, \\ &= \mathbf{D}_n \tilde{\mathbf{P}}_n \mathbf{s}_n + \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{G} \mathbf{R} \mathbf{n}_{r,n} + \mathbf{U}_n \mathcal{W}_{r,n} \mathbf{n}_{d,n} \end{aligned} \quad (29)$$

where  $\mathbf{D}_n = \text{diag}(d_n[1], d_n[2], \dots, d_n[\Gamma])$  is obtained from the optimal transceiver  $(\tilde{\mathbf{V}}_n, \mathbf{U}_n)$  of Problem 1-2 with each  $d_n[k]$  being a non-negative value [15], and  $\tilde{\mathbf{P}}_n = \text{diag}(p_n[1], p_n[2], \dots, p_n[\Gamma])$ . Therefore, we obtain  $N\Gamma$  parallel eigen-subchannels for the overall MIMO-OFDM system as

$$\hat{\mathbf{s}}_n[k] = d_n[k] p_n[k] s_n[k] + n_n[k] \quad (31)$$

for  $n = 0, 1, \dots, N-1$  and  $k = 1, 2, \dots, \Gamma$ , where  $n_n[k] = \mathbf{U}_{n,k}^H \mathcal{W}_{r,n} (\mathbf{G} \mathbf{R} \mathbf{n}_{r,n} + \mathbf{n}_{d,n})$  and  $\mathbf{U}_{n,k}^H$  is the  $k$ -th row of  $\mathbf{U}_n$ . The total power  $P_{s,\max}$  should now be optimally allocated to these  $N\Gamma$  parallel channels to minimize the weighted sum MSE, where the weighted sum MSE of  $N\Gamma$  parallel eigen-subchannels is derived as

$$\begin{aligned} & \sum_{n=0}^{N-1} \sum_{k=1}^{\Gamma} \theta_{nk} \mathbb{E} \left\{ |\hat{\mathbf{s}}_n[k] - s_n[k]|^2 \right\} \\ &= \sum_{n=0}^{N-1} \sum_{k=1}^{\Gamma} \theta_{nk} (d_n[k]^2 p_n[k]^2 - 2d_n[k] p_n[k] + c_n[k]) \end{aligned} \quad (32)$$

where  $c_n[k] = \sigma_r^2 \mathbf{U}_{n,k}^H \mathcal{W}_{r,n} \mathbf{G} \mathbf{R} \mathbf{R}^H \mathbf{G}^H \mathcal{W}_{r,n}^H \mathbf{U}_{n,k} + \sigma_d^2 \mathbf{U}_{n,k}^H \mathbf{U}_{n,k} + 1$ , and  $\theta_{nk}$  is properly derived from  $\Theta_n$ . Thus, the problem of overall source power allocation to minimize the

weight sum MSE subject to the source power constraint is stated as follows.

*Problem 1-3:* For given any weight matrices  $\{\Theta_n\}$ , SR channel  $\mathbf{F}$ , RD channel  $\mathbf{G}$ , relay filtering matrix  $\mathbf{R}$ , maximum source power  $P_{s,\max} = \sum_{n=0}^{N-1} P_{n,\max}$ , normalized transmit precoders  $\{\tilde{\mathbf{V}}_n\}$ , and receive decoders  $\{\mathbf{U}_n\}$ ,

$$\begin{aligned} \min_{p_n[k]} \quad & \sum_{n=0}^{N-1} \sum_{k=1}^{\Gamma} \theta_{nk} (d_n[k]^2 p_n[k]^2 - 2d_n[k]p_n[k] + c_n[k]) \\ \text{s.t.} \quad & \sum_{n=0}^{N-1} \sum_{k=1}^{\Gamma} p_n[k]^2 = P_{s,\max}. \end{aligned} \quad (33)$$

Note that Problem 1-3 is a convex optimization problem with respect to  $p_n[k]$ . The optimal solution to Problem 1-3 is given in the following proposition:

*Proposition 1:* The optimal solution to Problem 1-3 is given by

$$\begin{aligned} p_n[k] &= \left( \frac{\theta_{nk} d_n[k]}{\theta_{nk} d_n[k]^2 + \mu} \right)_+ \\ \text{s.t.} \quad & \sum_{n=0}^{N-1} \sum_{k=1}^{\Gamma} \left( \frac{\theta_{nk} d_n[k]}{\theta_{nk} d_n[k]^2 + \mu} \right)^2 = P_{s,\max}. \end{aligned} \quad (34)$$

*Proof:* See Appendix B.

The solution in Proposition 1 allocates power inverse-proportionally to the power of the effective channel  $d_n[k]$  in most cases similarly to the method in [26].

Now summarizing the results, we propose our method to design the linear transceiver at the source and the destination and the FF relay filter jointly to minimize the weighted sum MSE, based on alternating optimization solving Problem 1-1, Problem 1-2, and Problem 1-3 iteratively.

*Algorithm 1:* Given parameters:  $\{\Theta_n\}$ ,  $\mathbf{F}$ ,  $\mathbf{G}$ ,  $L_r$ ,  $P_{s,\max}$ , and  $P_{r,\max}$

- Step 1: Initialize  $\{\tilde{\mathbf{P}}_n\}$ ,  $\{\tilde{\mathbf{V}}_n\}$ , and  $\{\mathbf{U}_n\}$  for each  $n$ . For example,  $p_n[k] = P_{s,\max}/N\Gamma$ ,  $\tilde{\mathbf{V}}_n = \mathbf{I}_{N_t \times \Gamma}$ , and  $\mathbf{U}_n = \mathbf{I}_{\Gamma \times N_r}$ .
- Step 2: Solve Problem 1-1 and obtain  $\mathbf{R}$ .
- Step 3: Solve Problem 1-2 and obtain  $\{\tilde{\mathbf{V}}_n, \mathbf{U}_n\}$ .
- Step 4: Solve Problem 1-3 and obtain  $\{\tilde{\mathbf{P}}_n\}$ .
- Step 5: Go to Step 2 and repeat until the change in the weighted sum MSE falls within a given tolerance.

The weighted sum MSE is a function of  $\mathbf{R}$  and  $\{\tilde{\mathbf{V}}_n, \mathbf{U}_n, \tilde{\mathbf{P}}_n\}$  denoted by  $\mathcal{M}(\mathbf{R}, \tilde{\mathbf{V}}_n, \mathbf{U}_n, \tilde{\mathbf{P}}_n)$ . Let  $\mathbf{X}^{(i)}$  denotes the solution at the  $(i)$ -th step. Then, it is easy to see that  $\mathcal{M}(\mathbf{R}^{(0)}, \tilde{\mathbf{V}}_n^{(0)}, \mathbf{U}_n^{(0)}, \tilde{\mathbf{P}}_n^{(0)}) \geq \mathcal{M}(\mathbf{R}^{(1)}, \tilde{\mathbf{V}}_n^{(0)}, \mathbf{U}_n^{(0)}, \tilde{\mathbf{P}}_n^{(0)}) \geq \mathcal{M}(\mathbf{R}^{(1)}, \tilde{\mathbf{V}}_n^{(2)}, \mathbf{U}_n^{(2)}, \tilde{\mathbf{P}}_n^{(0)}) \geq \mathcal{M}(\mathbf{R}^{(1)}, \tilde{\mathbf{V}}_n^{(2)}, \mathbf{U}_n^{(2)}, \tilde{\mathbf{P}}_n^{(3)}) \geq \dots \geq 0$ , because the optimal solution is obtained at each step and the possible solution set of the current step includes the solution of the previous step. By the monotone convergence theorem, the proposed algorithm converges although it yields a suboptimal solution and the initialization of the algorithm affects its performance.

### C. Rate Maximization

Now we consider the problem of rate maximization. In general, the rate maximization problem is not equivalent to the MSE minimization problem. However, they are closely related to each other. The relationship has been studied in [15]–[17]. By using the relationship, the rate maximization problem for MIMO broadcast channels and MIMO interference-broadcast channels has recently been considered in [18] and [19]. In the case of the joint design of the FF relay at the relay and the linear MIMO transceiver at the source and the destination, the result regarding the weighted sum MSE minimization in the previous subsection can be modified and used to maximize the sum rate based on the existing relationship between the weighed MSE and the rate. It was shown in [15] that the rate maximization for the  $n$ -th subcarrier MIMO channel (29) is equivalent to the weighted MSE minimization when the weight matrix  $\Theta_n$  is set as a diagonal matrix composed of the eigenvalues of  $\mathbf{H}^H \Sigma_n^{-1} \mathbf{H}$ , where  $\mathbf{H} = \sqrt{N} \tilde{\mathbf{G}} \mathbf{R} \mathbf{F} \mathbf{W}_{t,n}^T$  is the effective MIMO channel matrix and  $\Sigma_n$  is the effective noise covariance matrix of the  $n$ -th subcarrier MIMO channel (29). (See Lemma 3 of [15].) Exploiting this result, we propose our algorithm to design the linear MIMO transceiver and the relay filter to maximize the sum rate below.

*Algorithm 2:* Given parameters:  $\mathbf{F}$ ,  $\mathbf{G}$ ,  $L_r$ ,  $P_{s,\max}$ , and  $P_{r,\max}$

- Step 1: Initialize  $\{\Theta_n\}$ ,  $\{\tilde{\mathbf{P}}_n\}$ ,  $\{\tilde{\mathbf{V}}_n\}$ , and  $\{\mathbf{U}_n\}$  for each  $n$ . For example,  $\Theta_n = \mathbf{I}$ ,  $p_n[k] = P_{s,\max}/N\Gamma$ ,  $\tilde{\mathbf{V}}_n = \mathbf{I}_{N_t \times \Gamma}$ , and  $\mathbf{U}_n = \mathbf{I}_{\Gamma \times N_r}$ .
- Step 2: Solve Problem 1-1 and obtain  $\mathbf{R}$ .
- Step 3: Solve Problem 1-2 and obtain  $\{\tilde{\mathbf{V}}_n, \mathbf{U}_n, \Theta_n\}$ .<sup>3</sup>
- Step 4: Compute  $\{\tilde{\mathbf{P}}_n\}$  for the  $N\Gamma$  parallel scalar channels obtained from Step 3 by water-filling.
- Step 5: Go to Step 2 and repeat until the change in the weighted sum MSE falls within a given tolerance.

Note that the weight matrices  $\{\Theta_n\}$  in Algorithm 2 are updated in each iteration so that the weighted MSE minimization is equivalent to the rate maximization for an updated relay filter, whereas the weight matrices are fixed over iterations in Algorithm 1.

Now consider the complexity of the proposed algorithms. Note that solving Problem 1-2 involves  $N$  separate small MIMO systems of size  $N_r \times N_t$ , and the solution to Problem 1-3 (Algorithm 1) and the water-filling power allocation solution (Algorithm 2) are explicitly given. Thus, the main complexity of the proposed algorithms lies in solving Problem 1-1 that requires solving a convex QCQP problem of size  $M_t L_r M_r \times 1$ . Due to the existence of fast approximate algorithms for solving convex problems [27], the proposed algorithm is implementable if the number of iterations for convergence is not so large, as seen in Fig. 4 in Section IV. For other practical issues such as channel estimation and self-interference caused by full-duplex operation, please see [7].

<sup>3</sup>When  $\mathbf{R}$  is given, all the parallel subcarrier MIMO channels are determined and a solution  $\{\tilde{\mathbf{V}}_n, \mathbf{U}_n, \Theta_n\}$  is given by Lemma 1 and Theorem 1 of [15].

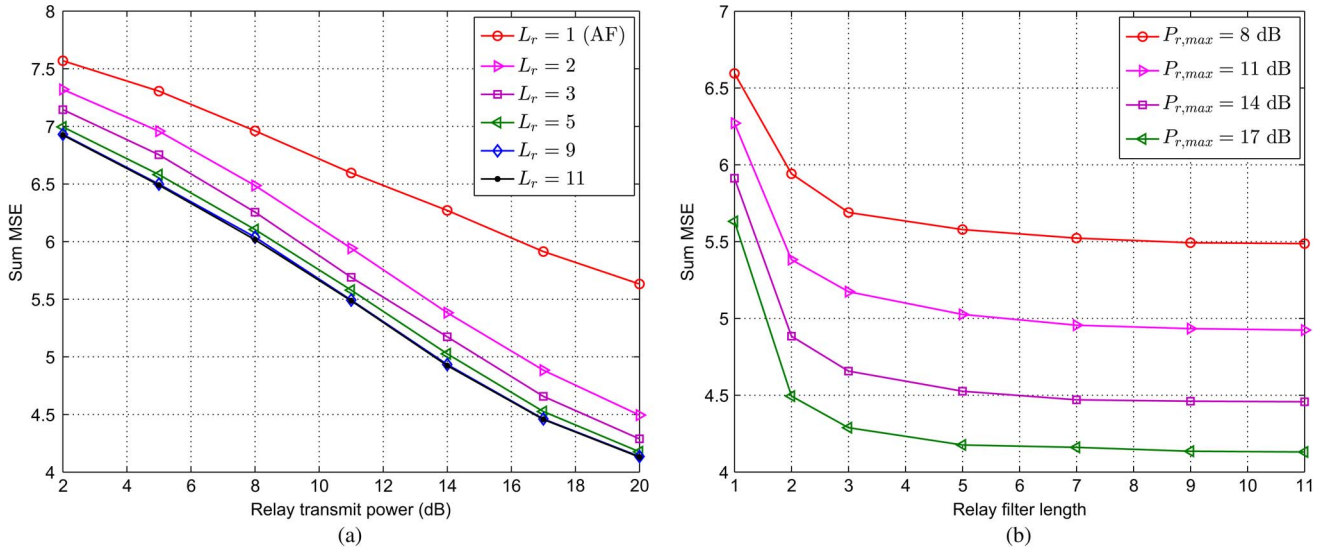


Fig. 2. (a) Sum MSE versus FF relay transmit power and (b) sum MSE versus relay filter length.

IV. NUMERICAL RESULTS

In this section, we provide some numerical results to evaluate the performance of the proposed FF relay design in Section III. Throughout the simulation, we fixed the number of OFDM subcarriers as  $N = 16$  with a minimal cyclic prefix covering the overall FIR channel length in each simulation case. In all cases, each channel tap coefficient of the SR and RD channel matrices,  $\mathbf{F}_k$  and  $\mathbf{G}_k$ , was generated independently according to a Rayleigh distribution, i.e.,  $\mathbf{F}_k(i, j) \stackrel{i.i.d.}{\sim} \mathcal{CN}(0, \sigma_{f,k}^2)$  and  $\mathbf{G}_k(i, j) \stackrel{i.i.d.}{\sim} \mathcal{CN}(0, \sigma_{g,k}^2)$ . The MIMO setting was  $N_t = M_r = M_t = N_r = 2$ . The relay and the destination had the same noise power  $\sigma_r^2 = \sigma_d^2 = 1$ , and the source transmit power was 20 dB higher than the noise power, i.e.,  $P_{s,max} = 100$ . (From here on, all dB power values are relative to  $\sigma_r^2 = \sigma_d^2 = 1$ .)

First, we fixed the SR channel length and the RD channel length as  $L_f = L_g = 3$  with  $\sigma_{f,k} = \sigma_{g,k} = 1$  for  $\forall k$ , and evaluated the performance. We first evaluated the MSE performance of the proposed FF relay design method, Algorithm 1, to minimize the sum MSE subject to a source power constraint and a relay power constraint. Fig. 2 shows the resulting sum MSE over all subcarriers. For the curves in Fig. 2, 200 channels were randomly realized and each plotted value is the average over the 200 channel realizations. As expected, it is seen in Fig. 2 that the performance of the FF relay improves as the FF relay filter length increases, and the FF relay significantly outperforms the simple AF relay ( $L_r = 1$ ). It is also seen that most of the gain is achieved by only a few filter taps for the FF relay.

Second, we investigated the BER performance corresponding to Fig. 2(a). Here, we assumed uncoded QPSK modulation for each subcarrier channel. From the result of Fig. 2(a), we obtained the SNR of each subcarrier channel of the total  $N = 16$  subcarrier channels for the designed FF relay filter, transmit precoder, receive decoder and source power allocation. Based on this, we computed the subcarrier BER based on the SNR of each subcarrier and averaged all the subcarrier channel BERs to obtain the overall BER, and the result is shown in Fig. 3. It is seen in Fig. 3 that the FF relay significantly improves the BER

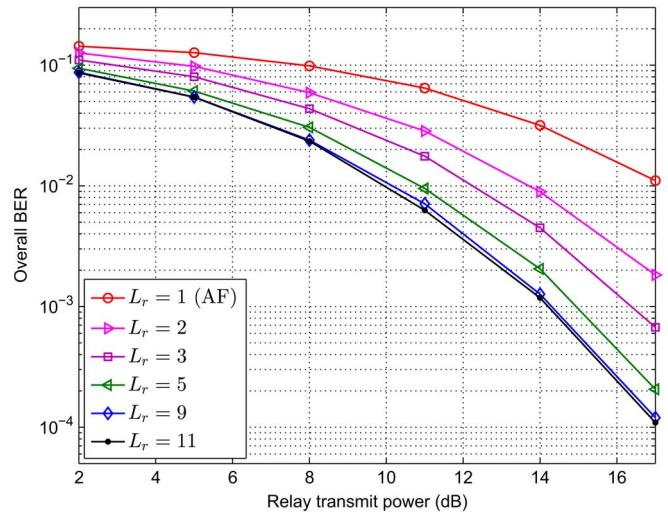


Fig. 3. Overall BER versus FF relay transmit power.

performance over the AF relay. Next, we tested the convergence property of the proposed algorithm, and Fig. 4 shows the result. It is seen that the proposed algorithm converges with a few iterations.

Next, we examined the rate performance of the proposed rate-targeting design method, Algorithm 2. (Rate maximization may be the ultimate goal of design in many cases.) Fig. 5 shows the result. Again, for the figure, 200 channels were randomly realized with  $L_f = L_g = 3$  and  $\sigma_{f,k} = \sigma_{g,k} = 1$ . Each plotted value is the average over the 200 channel realizations and the sum rate is the sum over the total subcarrier channels. It is shown in Fig. 5 that the FF relay improves the rate performance as the FF relay filter length increases, and the improvement gap shows that it is worth considering FF relays over simple AF relays even though FF relays require more processing than AF relays.

Finally, we investigated the impact of the channel length and power delay profile on the MSE performance and the rate performance. The different channel length and power delay profile of SR and RD channels are summarized as  $\mathbf{f} = [\sigma_{f,1}^2, \sigma_{f,2}^2,$



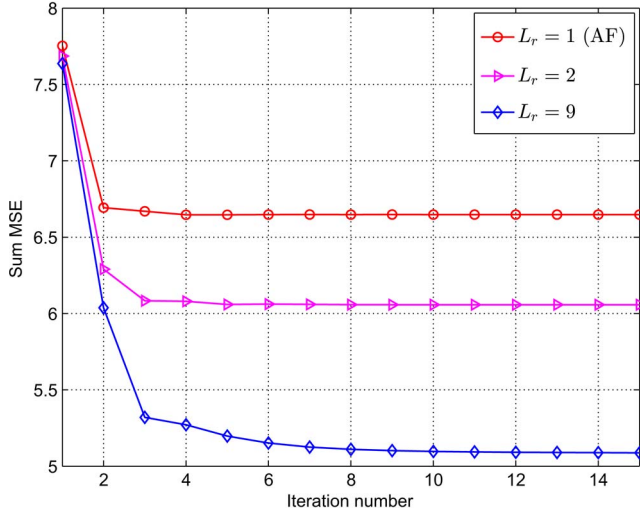


Fig. 4. Sum MSE versus the number of iteration.

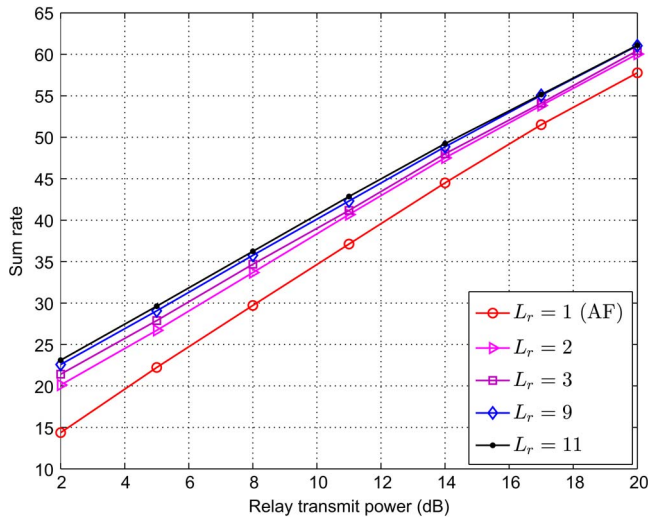


Fig. 5. Sum rate versus FF relay transmit power.

$\dots, \sigma_{f,L_f}^2$ ] and  $\mathbf{g} = [\sigma_{g,1}^2, \sigma_{g,2}^2, \dots, \sigma_{g,L_g}^2]$ , where  $L_f$  and  $L_g$  are the SR channel length and the RD channel length, respectively. Fig. 6(a) shows the resulting sum MSE over all subcarriers in the case that the sum of the averaged channel tap power is identical but the channel length is different. For the curves in the figure, 200 channels were randomly realized according to each channel length and power delay profile and each plotted value is the average over the 200 channel realizations. It is seen in the figure that when both the SR channel length and the RD channel length are one, there is no performance difference between AF ( $L_r = 1$ ) and FF ( $L_r \geq 2$ ). However, it is also seen in the figure that for the AF, there is performance degradation as the SR channel length and the RD channel length increase. Note that the FF improves the degraded performance of the AF. Fig. 6(b) shows the rate performance for the same setting as that in Fig. 6(a). Similar behavior is observed in this case too.

## V. CONCLUSION

In this paper, we have considered the FF relaying scheme for MIMO-OFDM systems and the joint design of the FF relay and

the linear MIMO transceiver based on two criteria, weighted sum MSE minimization and sum rate maximization, and have proposed algorithms for this purpose based on alternating optimization that iterates between optimal design of the FF relay for a given MIMO transceiver and optimal design of the MIMO transceiver for a given FF relay filter. We have shown that the FF relay design problem for a given MIMO transceiver reduces to a convex quadratically constrained quadratic program (QCQP) which can easily be solved by a standard convex optimization solver. We have provided some numerical results to evaluate the performance gain of the FF relaying scheme over the simple AF scheme. Numerical results show the effectiveness of the proposed FF relay design and suggest that it is worth considering the FF relaying scheme over the widely-considered simple AF scheme for MIMO-OFDM systems.

## APPENDIX

### A. $\mathbf{E}_1$ and $\mathbf{E}_2$ Matrices

$\mathbf{E}_1$  and  $\mathbf{E}_2$  are  $M_t L_r M_r \times M_t(N + L_r + L_g - 2)(N + L_g - 1)M_r$  matrices and defined as follows:

$$\mathbf{E}_1 = \begin{bmatrix} \text{blkToeplitz} \left( \begin{bmatrix} \mathbf{I}_{L_r} \\ \mathbf{0}_{(N+L_g-2) \times L_r} \end{bmatrix}, M_t \right) \\ \text{blkToeplitz} \left( \begin{bmatrix} \mathbf{0}_{1 \times L_r} \\ \mathbf{I}_{L_r} \\ \mathbf{0}_{(N+L_g-3) \times L_r} \end{bmatrix}, M_t \right) \\ \text{blkToeplitz} \left( \begin{bmatrix} \mathbf{0}_{2 \times L_r} \\ \mathbf{I}_{L_r} \\ \mathbf{0}_{(N+L_g-4) \times L_r} \end{bmatrix}, M_t \right) \\ \vdots \\ \text{blkToeplitz} \left( \begin{bmatrix} \mathbf{0}_{(N+L_g-2) \times L_r} \\ \mathbf{I}_{L_r} \end{bmatrix}, M_t \right) \end{bmatrix}^T \otimes \mathbf{I}_{M_r} \quad (35)$$

$$\mathbf{E}_2 = \begin{bmatrix} \text{blkToeplitz} \left( [\mathcal{E}_1 \cdots \mathcal{E}_{(N+L_g-2)M_t+1}], L_r \right) \\ \text{blkToeplitz} \left( [\mathcal{E}_2 \cdots \mathcal{E}_{(N+L_g-2)M_t+2}], L_r \right) \\ \vdots \\ \text{blkToeplitz} \left( [\mathcal{E}_{M_t} \cdots \mathcal{E}_{(N+L_g-2)M_t+M_t}], L_r \right) \end{bmatrix} \quad (36)$$

where

$$\mathcal{E}_k = \begin{bmatrix} \mathbf{e}_k^T \\ \mathbf{e}_{(N+L_g-1)M_t+k}^T \\ \mathbf{e}_{2(N+L_g-1)M_t+k}^T \\ \vdots \\ \mathbf{e}_{(M_r-1)(N+L_g-1)M_t+k}^T \end{bmatrix}$$

and  $\mathbf{e}_i^T$  is the  $i$ -th row of  $\mathbf{I}_{(N+L_g-1)M_t M_r}$ .

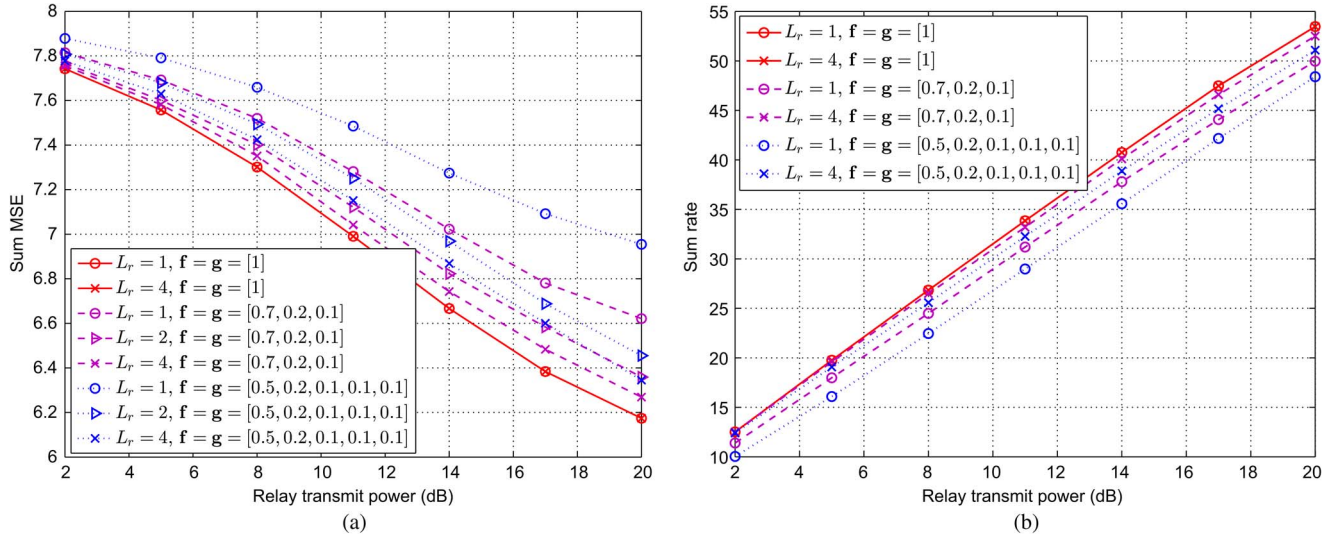


Fig. 6. Impact of  $\mathbf{f}$  and  $\mathbf{g}$ : (a) sum MSE versus FF relay transmit power and (b) sum rate versus FF relay transmit power.

### B. Proof of Proposition 1

The Lagrangian of (33) is given by

$$\begin{aligned} \mathcal{L}(p_n[k], \mu) &= \sum_{n=0}^{N-1} \sum_{k=1}^B \theta_{nk} (d_n[k]^2 p_n[k]^2 - 2d_n[k]p_n[k] + c_n[k]) \\ &+ \mu \left( \sum_{n=0}^{N-1} \sum_{k=1}^B p_n[k]^2 - P_{s,\max} \right) - \sum_{n=0}^{N-1} \sum_{k=1}^B \lambda_{n,k} p_n[k] \end{aligned} \quad (37)$$

where  $\mu \in \mathcal{R}$  and  $\lambda_{n,k} \geq 0$  are dual variables associated with the source power constraint and the positiveness of power, respectively. Then, the following KKT conditions are necessary and sufficient for optimality because the problem (33) is a convex optimization problem:

$$p_n[k] \geq 0, \quad \sum_{n=0}^{N-1} \sum_{k=1}^B p_n[k]^2 - P_{s,\max} = 0, \quad (38)$$

$$\mu \in \mathcal{R}, \quad \lambda_{n,k} \geq 0, \quad (39)$$

$$\lambda_{n,k} p_n[k] = 0 \quad (40)$$

$$\begin{aligned} \nabla_{p_n[k]} \mathcal{L} &= 2\theta_{nk} d_n[k]^2 p_n[k] - 2\theta_{nk} d_n[k] \\ &+ 2\mu p_n[k] - \lambda_{n,k} = 0 \end{aligned} \quad (41)$$

for  $n = 0, 1, \dots, N-1$  and  $k = 1, \dots, B$ .

The gradient (41) can be rewritten as  $\lambda_{n,k} = 2(\theta_{nk} d_n[k]^2 + \mu)p_n[k] - 2\theta_{nk} d_n[k]$ . Plugging this into (39) and (40), we get

$$\mu p_n[k] \geq \theta_{nk} d_n[k] - \theta_{nk} d_n[k]^2 p_n[k] \quad (42)$$

$$((\theta_{nk} d_n[k]^2 + \mu) p_n[k] - \theta_{nk} d_n[k]) p_n[k] = 0. \quad (43)$$

Let us consider the case that  $p_n[k] = 0$ . Then, (42) is satisfied only if  $d_n[k] = 0$  because  $d_n[k] \geq 0$ . If  $p_n[k] > 0$ ,  $p_n[k] = (\theta_{nk} d_n[k]) / (\theta_{nk} d_n[k]^2 + \mu)$  by the complementary slackness (43). This also satisfies (42). Therefore, we get the desired result satisfying the primal constraints (38).  $\square$

### REFERENCES

- [1] A. El Gamal, M. Mohseni, and S. Zahedi, "Bounds on capacity and minimum energy-per-bit for AWGN relay channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1545–1561, Apr. 2006.
- [2] A. del Coso and C. Ibars, "Linear relaying for the Gaussian multiple-access and broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 2024–2035, Apr. 2009.
- [3] Y. Sung and C. Kim, "The Capacity for the Linear Time-Invariant Gaussian Relay Channel," Sep. 2011. [Online]. Available: ArXiv: <http://arxiv.org/abs/1109.5426> and <http://wisrl.kaist.ac.kr/papers/Sungetal12ISITsubmission.pdf>
- [4] H. Chen, A. Gershman, and S. Shahbazpanahi, "Filter-and-forward distributed beamforming in relay networks with frequency selective fading," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1251–1262, Mar. 2010.
- [5] C. Kim, Y. Sung, and Y. H. Lee, "A joint time-invariant filtering approach to the linear Gaussian relay problem," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4360–4375, Aug. 2012.
- [6] Y. Liang, A. Ikhlef, W. Gerstacker, and R. Schober, "Cooperative filter-and-forward beamforming for frequency-selective channels with equalization," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 228–239, Jan. 2011.
- [7] D. Kim, J. Seo, and Y. Sung, "Filter-and-forward transparent relay design for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4392–4407, Nov. 2013.
- [8] I. Hammerström and A. Wittneben, "On the optimal power allocation for nongenerative OFDM relay links," in *Proc. ICC*, Jun. 2006, pp. 4463–4468.
- [9] T. Ng and W. Yu, "Joint optimization of relay strategies and resource allocations in cooperative cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 328–339, Feb. 2007.
- [10] M. Dong and S. Shahbazpanahi, "Optimal spectrum sharing and power allocation for OFDM-based two-way relaying," in *Proc. ICASSP*, Mar. 2010, pp. 3310–3313.
- [11] W. Dang, M. Tao, H. Mu, and J. Huang, "Subcarrier-pair based resource allocation for cooperative multi-relay OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1640–1649, May 2010.
- [12] Z. Fang, Y. Hua, and J. C. Koshy, "Joint source and relay optimization for a non-regenerative MIMO relay," in *Proc. IEEE SAM*, Jul. 2006, pp. 239–243.
- [13] S. Simoens, O. Munoz-Medina, J. Vidal, and A. del Coso, "On the Gaussian MIMO relay channel with full channel state information," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3588–3599, Sep. 2009.
- [14] D. Kim, J. Seo, and Y. Sung, "Filter-and-forward relay design for OFDM systems for quality-of-service enhancement," in *Proc. APSIPA ASC*, Hollywood, CA, USA, Dec. 2012, pp. 1–4.
- [15] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.
- [16] D. Guo, S. Shitz, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.

- [17] D. Palomar and S. Verdu, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 141–154, Jan. 2006.
- [18] S. S. Christensen, R. Argawal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.
- [19] Q. Shi, M. Razaviyayn, Z. Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.
- [20] R. M. Gray, *Toeplitz and Circulant Matrices: A Review*. Delft, The Netherlands: Now Publishers, 2006.
- [21] P. J. Davis, *Circulant Matrices*, 2nd ed. New York, NY, USA: Chelsea, 1994.
- [22] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.
- [23] C. Helmberg, "Semidefinite programming," *Eur. J. Oper. Res.*, vol. 137, no. 3, pp. 461–482, 2002.
- [24] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optim. Methods Softw. Spec. Issue Interior Point Methods*, vol. 11/12, no. 1–4, pp. 625–653, 1999.
- [25] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming, ver. Version 1.21 2011.
- [26] H. Sampath and A. Paulraj, "Joint transmit and receive optimization for high data rate wireless communications using multiple antennas," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Oct. 1999, vol. 1, pp. 215–219.
- [27] S. Arora, E. Hazan, and S. Kale, "The multiplicative weights update method: A meta-algorithm and applications," *Theory Comput.*, vol. 8, pp. 121–164, 2012.



**Donggun Kim** (S'10) received the B.S. degree in electronics and communications engineering from Hanyang University, Seoul, Korea, in 2010. He is currently working toward the Ph.D. degree with the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Korea.

His research interests include convex optimization, matrix theory, and signal processing for wireless communications.



**Youngchul Sung** (S'92–M'93–SM'09) received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Seoul, Korea, in 1993 and 1995, respectively, and the Ph.D. degree in electrical and computer engineering from Cornell University, Ithaca, NY, USA, in 2005. From 1995 to 2000, he was with LG Electronics, Ltd., Seoul. From 2005 to 2007, he was a Senior Engineer with the Corporate R&D Center, Qualcomm, Inc., San Diego, CA, USA, and participated in the design of WCDMA base station modem. Since 2007, he has been with

the faculty of the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Korea. His research interests include statistical signal processing, asymptotic statistics, and information geometry with applications to next-generation wireless communications and related areas.

Dr. Sung is a member of Signal and Information Processing Theory and Methods Technical Committee of the Asia-Pacific Signal and Information Processing Association (APSIPA), the Vice-Chair of IEEE ComSoc Asia-Pacific Board ISC, and a Technical Program Committee member of conferences, including Globecom 2009–2014, ICC 2011/2014, VTC 2014 Track Chair, APSIPA 2009–2013, DCOSS 2010/2013, MILCOM 2010/2014, WiOpt 2009 and its sponsorship chair, etc. He served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS from 2012 to 2014 and as a Guest Editor of the JSAC Special Issue on "Theories and Methods for Advanced Wireless Relays" in 2012.



**Jihoon Chung** (S'13) received the B.S. degree in electronics and computer engineering in 2012 from Hanyang University, Seoul, Korea, and the M.S. degree in electrical engineering in 2014 from Korea Advanced Institute of Science and Technology, Daejeon, Korea, where he is currently working toward the Ph.D. degree with the Department of Electrical Engineering.

His current research interests include signal processing for wireless communications and convex optimization.