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 considered. Due to the considered MIMO structure, the problem of joint design of the linear MIMO transceiver at the source and the destination and the FF relay at the relay is considered. As the design criterion, the minimization of weighted sum mean-square-error (MSE) is considered first, and the joint design in this case is approached based on alternating optimization that iterates between optimal design of the FF relay for a given set of MIMO precoder and decoder and optimal design of the MIMO precoder and decoder for a given FF relay filter. Next, the joint design problem for rate maximization is considered based on the obtained result regarding weighted sum MSE 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 and significant performance improvement by FF relays over widely-considered simple AF relays for MIMO-OFDM systems.

Index Terms

†Corresponding author

The authors are with the Dept. of Electrical Engineering, KAIST, Daejeon 305-701, South Korea. E-mail:{dg.kim@, ysung@ee., and j.chung@}kaist.ac.kr. This research was funded by the MSIP(Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013.
Recently, the filter-and-forward (FF) relaying scheme has gained an interest from the research communities as an alternative relaying strategy due to its capability of performance improvement over simple AF relays and still low complexity compared with other relaying strategies such as decode-and-forward (DF) and compress-and-forward (CF) schemes [1]–[7]. It is shown that the FF scheme can outperform the AF scheme considerably. However, most of the works regarding the FF relay scheme were conducted for single-carrier systems [1], [3], [5], [6]. Recently, Kim et al. considered the FF relay design for single-input and single-output (SISO) OFDM systems [7], [8], but their result based on worst subcarrier signal-to-noise ratio (SNR) maximization or direct rate maximization is not easily extended to the MIMO case since SNR is not clearly defined for MIMO channels and furthermore in the MIMO case the design of the MIMO precoder at the source and the MIMO decoder at the destination should be considered jointly with the FF relay design. Thus, although there exists vast literature regarding the relay design for MIMO-OFDM systems in the case that the relay performs OFDM processing [9]–[14], not many results are available for the FF relay design for MIMO-OFDM transmission, which is the current industry standard for the physical layer of many commercial wireless communication systems.

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 without considering the MIMO precoder and decoder at the source and the destination. Thus, we consider the problem of joint design of the linear MIMO transceiver at the source and the destination and the FF relay at the relay. As mentioned, in the MIMO case, it is not easy to use SNR as the design metric as in the SISO case [7]. Thus,

*In this case, each subcarrier channel is independent and we only need to consider a single flat MIMO channel.*
we approach the design problem based on the tractable criterion of minimization of weighted sum MSE first and then consider the 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 problems by using alternating optimization, which enables us to exploit the existing results for the MIMO 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 a given set of MIMO precoder and decoder and optimal design of the MIMO precoder and decoder for a given FF relay filter. While the linear MIMO transceiver design for a given FF relay filter can be addressed by existing results e.g. [15], the problem of optimal design of the FF relay for a given MIMO transceiver is newly formulated based on the block circulant matrix theorem and reparameterization. It is shown that the FF relay design problem for a given MIMO transceiver reduces to a quadratically constrained quadratic program (QCQP) problem and a solution to this QCQP problem is proposed based on conversion to a semi-definite program (SDP). Numerical results show the effectiveness of the proposed FF relay design and significant performance improvement by FF relays over widely-considered simple AF relays, and suggests that it is worth considering the FF relaying scheme for MIMO-OFDM systems over the AF scheme with a certain amount of complexity increase.

A. Notation and Organization

In this paper, 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 $X$, $X^*$, $X^T$, $X^H$, $\text{tr}(X)$, and $X(i,j)$ indicate the complex conjugate, transpose, conjugate transpose, trace, and $(i,j)$-element of $X$, respectively. $X \succeq 0$ and $X \succ 0$ mean that $X$ is positive semi-definite and that $X$ is strictly positive definite, respectively. $I_n$ stands for the identity matrix of size $n$ (the subscript is omitted when unnecessary), $I_{m \times n}$ denotes the first $m \times n$ submatrix of $I$, and $0_{m \times n}$ denotes a $m \times n$ matrix of all zero elements (the subscript is omitted when unnecessary).
notation \( \text{blkToeplitz}(\mathbf{F}, N) \) indicates a \( NA \times (N + L_f - 1)B \) block Toeplitz matrix with \( N \) row blocks and \([\mathbf{F}, 0, \cdots, 0]\) as its first row block, where \( \mathbf{F} = [\mathbf{F}_0, \mathbf{F}_1, \cdots, \mathbf{F}_{L_f-1}] \) is a row block composed of \( A \times B \) matrices \( \{\mathbf{F}_k\} \); \( \text{diag}(X_1, X_2, \cdots, X_n) \) means a (block) diagonal matrix with diagonal entries \( X_1, X_2, \cdots, X_n \). The notation \( x \sim \mathcal{CN}(\mu, \Sigma) \) means that \( x \) is complex circularly-symmetric Gaussian distributed with mean vector \( \mu \) and covariance matrix \( \Sigma \). \( \mathbb{E}\{\cdot\} \) 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 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. I, 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 \( M_t M_r \) finite impulse response (FIR) filters with order \( L_g \), i.e., the relay performs FIR filtering on the incoming signals received at the \( M_r \) receive antennas at the chip rate‡ 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

†In the case of half-duplex, the problem can be formulated similarly.
‡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.
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 filters with finite impulse responses and their state information is known to the system.

\[
s_0, s_1, \ldots, s_n
\]

\[
V_0, V_1, \ldots, V_n
\]

\[
N_t, M_r, M_t, N_r
\]

\[
\hat{s}_0, \hat{s}_1, \ldots, \hat{s}_n
\]

\[
V_0, V_1, \ldots, V_n
\]

\[
P/S, & \text{CP}
\]

\[
\text{IDFT}
\]

\[
\text{FIR filter}
\]

\[
M_r, M_t
\]

\[
A_0, A_1, \ldots, A_n
\]

\[
\text{CPR, S/P}
\]

\[
\text{DFT}
\]

\[
\text{FF relay}
\]

Fig. 1: System model

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 \( \{s_n = [s_n[1], s_n[2], \ldots, s_n[\Gamma]]^T, n = 0, 1, \ldots, N - 1\} \), is processed for one OFDM symbol time. Here, \( s_n \) is the input data vector for the effective parallel flat 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 \( k = 1, 2, \ldots, \Gamma \) and \( n = 0, 1, \ldots, N - 1 \). Let the concatenated data vector be denoted by \( s = [s_N^T, s_{N-1}^T, \ldots, s_0^T]^T \). Although MIMO precoding can be applied to the concatenated vector \( s \), such processing is complexity-wise 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 \( V_n \) 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 $x$, given by

$$x = (W_N \otimes I_{N_t})Vs$$  \hspace{1cm} (1)$$

where

$$V = \text{diag}(V_{N-1}, V_{N-2}, \cdots, V_0)$$ \hspace{1cm} (2)$$

$$W_N(k+1, l+1) = \frac{1}{\sqrt{N}} e^{\frac{2\pi i kl}{N}}, \quad k, l = 0, 1, \cdots, N-1,$$ \hspace{1cm} (3)$$

and $x$ is cyclic-prefix attached and transmitted. The cyclic prefix attached signal vector $x_{cp}$ can be expressed as

$$x_{cp} = \left( \begin{array}{c} I_N \\ I_{N_{cp}} \\ 0 \end{array} \right) \otimes I_{N_t} x, \quad \Delta = \mathbf{t}_{cp}$$ \hspace{1cm} (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 $x_{cp}$ 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

$$y_t = RFx_{cp} + Rn_r \quad \text{and} \quad y_d = GRFx_{cp} + GRn_r + n_d,$$ \hspace{1cm} (5)$$
where

\[
y_d = [y_{d,N-1}^T, y_{d,N-2}^T, \ldots, y_{d,0}^T]^T,
\]

\[
y_t = [y_{t,N-1}^T, y_{t,N-2}^T, \ldots, y_{t,0}^T, y_{t,-1}^T, \ldots, y_{t,-L_g+1}^T]^T,
\]

\[
x_{cp} = [x_{t,N-1}^T, x_{t,N-2}^T, \ldots, x_{0}^T, x_{-1}^T, \ldots, x_{-L_g-L_c-L_f+3}^T]^T,
\]

\[
n_r = [n_{r,N-1}^T, n_{r,N-2}^T, \ldots, n_{r,0}^T, n_{r,-1}^T, \ldots, n_{r,-L_g-L_c-2}^T]^T,
\]

\[
n_d = [n_{d,N-1}^T, n_{d,N-2}^T, \ldots, n_{d,0}^T]^T,
\]

\[
G = \text{blkToeplitz}(\mathcal{G}, N), \quad R = \text{blkToeplitz}(\mathcal{R}, N + L_g - 1), \quad F = \text{blkToeplitz}(\mathcal{F}, N + L_g + L_r - 2),
\]

\[
\mathcal{G} = [G_0, G_1, \ldots, G_{L_g-1}], \quad \mathcal{R} = [R_0, R_1, \ldots, R_{L_r-1}], \quad \mathcal{F} = [F_0, F_1, \ldots, F_{L_f-1}].
\]

Here, \( y_{d,k} \) and \( n_{d,k} \) are \( N_r \times 1 \) vectors; \( y_{t,k} \) is a \( M_t \times 1 \) vector; \( x_k \) is a \( N_t \times 1 \) vector; \( n_{r,k} \) is a \( M_r \times 1 \) vector; \( G_k \) is a \( N_r \times M_t \) matrix; \( R_k \) is a \( M_t \times M_r \) matrix; and \( F_k \) is a \( M_r \times N_t \) matrix.

The entries of the noise vectors, \( n_{r,k} \) and \( n_{d,k} \), are independently and identically distributed (i.i.d) Gaussian with \( n_{r,k}[i] \overset{i.i.d.}{\sim} \mathcal{CN}(0, \sigma_r^2) \) and \( n_{d,k}[i] \overset{i.i.d.}{\sim} \mathcal{CN}(0, \sigma_d^2) \). Then, the (cyclic-prefix portion removed) \( N \)-point vector DFT of the received vector at the destination is given by

\[
y = (W_N^H \otimes I_{N_r}) \text{GRFx}_{cp} + (W_N^H \otimes I_{N_r}) \text{GRn}_r + (W_N^H \otimes I_{N_r}) n_d,
\]

\[
= (W_N^H \otimes I_{N_r}) \text{GRFT}_{cp}(W_N \otimes I_{N_t}) V \mathbf{s} + (W_N^H \otimes I_{N_r}) \text{GRn}_r + (W_N^H \otimes I_{N_r}) n_d,
\]

\[
= (W_N^H \otimes I_{N_r}) H_c(W_N \otimes I_{N_t}) V \mathbf{s} + (W_N^H \otimes I_{N_r}) \text{GRn}_r + (W_N^H \otimes I_{N_r}) n_d,
\]

\[
= D V \mathbf{s} + (W_N^H \otimes I_{N_r}) \text{GRn}_r + (W_N^H \otimes I_{N_r}) n_d,
\]

where \( y = [y_{N-1}^T, y_{N-2}^T, \ldots, y_0^T]^T \), \( y_n \) is a \( N_r \times 1 \) received signal vector at the \( n \)-th subcarrier, \( W_N^H \) is the normalized DFT matrix of size \( N \), \( H_c \) is a \( NN_r \times NN_t \) block circulant matrix generated from the block Toeplitz overall channel matrix \( \text{GRF} \) from the source to the destination, and \( D = (W_N^H \otimes I_{N_r}) H_c(W_N \otimes I_{N_t}) \) 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 linear receiver filter \( U_n \) of size \( \Gamma \times N_r \) to yield an estimate of \( s_n \). The
overall receiver processing for all the subcarrier channels can be expressed as

\[ \hat{s} = UDVs + U(W_N^H \otimes I_{N_r})GRn + U(W_N^H \otimes I_{N_r})n_d, \]  

(15)

where \( U = \text{diag}(U_{N-1}, U_{N-2}, \cdots, U_0) \).

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 \( y_n, n = 0, 1, \cdots, N-1 \), at the \( n \)-th subcarrier.

Lemma 1: If \( H_c \) is a block circulant matrix with \( K = [H_0, H_1, \cdots, H_{N-1}] \) as its first row block, then it is block-diagonalizable as

\[ \Lambda_b = (W_N^H \otimes I_{N_r}) H_c (W_N \otimes I_{N_t}) \]

where \( \Lambda_b \) is a block diagonal matrix defined as

\[ \Lambda_b = \begin{bmatrix} K(\sqrt{N}w_{N-1}^H \otimes I_{N_t})^T & 0 \\ \vdots \\ 0 & K(\sqrt{N}w_0^H \otimes I_{N_t})^T \end{bmatrix} \]

with \( \sqrt{N}w_k^H \) denoting the \(-(k-N)\)-th row of the DFT matrix \( \sqrt{N}W_N^H \), and

\[ K(\sqrt{N}w_k^H \otimes I_{N_t})^T = \sum_{n=0}^{N-1} H_n e^{-j2\pi \frac{n(N-k-1)}{N}}. \]

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. \( \square \)

By lemma 1, to derive the diagonal blocks of \( D \) in (14), we only need to know the first row block of \( H_c \) in (13). Let the first row block of the RD channel matrix \( G \) be denoted by a \( N_r \times M_t(N + L_g - 1) \) matrix \( \tilde{G} = [G_0, G_1, \cdots, G_{L_g-1}, 0, \cdots, 0] \). Then, the first row block of the effective channel filtering matrix \( GRF \) is given by \( \tilde{GRF} \). Note that the cyclic prefix adding and removing operations make \( GRF \) into the block circulant matrix \( H_c \) by truncating out the
blocks of 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 $N_c \times N_t$ matrix. Therefore, the first row block $\tilde{H}_c$ of $H_c$ is simply the first $N$ blocks of $\tilde{GRF}$, given by

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

(16)

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

$$D = \text{diag}(\tilde{H}_c(\sqrt{N}w_{N-1}^H \otimes I_{N_t})^T, \tilde{H}_c(\sqrt{N}w_{N-2}^H \otimes I_{N_t})^T, \cdots, \tilde{H}_c(\sqrt{N}w_0^H \otimes I_{N_t})^T).$$

(17)

Based on (16) and (17), the received signal vector on the $n$-th subcarrier at the destination is expressed as

$$y_n = \sqrt{N}\tilde{GRFT}W_{t,n}^TV_n s_n + W_{r,n}GR_{r,n} + W_{r,n}n_{d,n},$$

(18)

$$= \hat{y}_n + z_n,$$

(19)

where $W_{t,n} = w_n^H \otimes I_{N_t}$, $W_{r,n} = w_n^H \otimes I_{N_r}$, $\hat{y}_n = \sqrt{N}\tilde{GRFT}W_{t,n}^TV_n s_n$, and $z_n = W_{r,n}GR_{r,n} + W_{r,n}n_{d,n}$. This received signal vector $y_n$ is filtered by the receive filter $U_n$ and its output is given by

$$\hat{s}_n = \sqrt{N}U_n\tilde{GRFT}W_{t,n}^TV_n s_n + U_nW_{r,n}GR_{r,n} + U_nW_{r,n}n_{d,n}.$$  

(20)
Finally, the weighted MSE between $s_n$ and $\hat{s}_n$ is given by

$$\text{tr}(\Theta_n \mathcal{M}_n) = \text{tr} \left( \Theta_n \mathbb{E} \left\{ (\hat{s}_n - s_n)(\hat{s}_n - s_n)^H \right\} \right),$$

$$= \text{tr} \left( \Theta_n \mathbb{E} \left\{ (U_n y_n - s_n)(U_n y_n - s_n)^H \right\} \right),$$

$$= \text{tr} \left( \Theta_n \left( \mathbb{E} \left\{ U_n y_n y_n^H U_n^H \right\} - \mathbb{E} \left\{ s_n y_n^H U_n^H \right\} - \mathbb{E} \left\{ U_n y_n s_n^H \right\} + \mathbb{E} \left\{ s_n s_n^H \right\} \right) \right),$$

$$= \text{tr} \left( \Theta_n \mathbb{E} \left\{ U_n \hat{y}_n \hat{y}_n^H U_n^H \right\} \right) + \text{tr} \left( \Theta_n \mathbb{E} \left\{ U_n z_n z_n^H U_n^H \right\} \right) - \text{tr} \left( \Theta_n \mathbb{E} \left\{ s_n \hat{y}^H U_n^H \right\} \right)$$

$$- \text{tr} \left( \Theta_n \mathbb{E} \left\{ U_n \hat{y}_n s_n^H \right\} \right) + \text{tr} \left( \Theta_n \mathbb{E} \left\{ s_n s_n^H \right\} \right),$$

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

**III. PROBLEM FORMULATION AND PROPOSED DESIGN METHOD**

In this section, we consider optimal design of the FIR MIMO relay filter $\{R_0, R_1, \cdots, R_{L_r-1}\}$ and the linear precoders and decoders $\{V_n, U_n, n = 0, 1, \cdots, N-1\}$. 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 $\{\Theta_n\}$. (Here, the summation is across the subcarrier channels.) The first problem is formally stated as follows.

**Problem 1:** For given weight matrices $\{\Theta_n\}$, SR channel $F$, RD channel $G$, FF relay filter order $L_r$, maximum source transmit power $P_{s,\text{max}}$, and maximum relay transmit power $P_{r,\text{max}}$, optimize the transmit filter $V = \text{diag}(V_0, \cdots, V_{N-1})$, the relay filter $\bar{R}$, and the receive filter $U = \text{diag}(U_0, \cdots, U_{N-1})$ in order to minimize the weighted sum MSE:

$$\min_{V, \bar{R}, U} \sum_{n=0}^{N-1} \text{tr}(\Theta_n \mathcal{M}_n) \quad \text{s.t.} \quad \text{tr}(VV^H) \leq P_{s,\text{max}} \quad \text{and} \quad \text{tr}(y_t y_t^H) \leq P_{r,\text{max}}. \quad (22)$$

Note that Problem 1 is a complicated non-convex optimization problem, which does not yield an easy solution. To circumvent the difficulty in joint optimization, we approach the problem...
based on alternating optimization. That is, we first optimize the relay filter for given transmit and receive filters under the power constraints. Then, with the obtained relay filter we optimize the transmit and receive filters. 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 precoder \( V_n \) and decoder \( U_n \) 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 transmit and receive filters, and the problem is formulated as follows.

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

\[
\min_{R} \sum_{n=0}^{N-1} \text{tr}(\Theta_n M_n) \quad \text{s.t.} \quad \text{tr}(y_t y_t^H) \leq P_{r,max}.
\]  

(23)

To solve Problem 1-1, we first need to express each term in (23) as a function of the design variable \( \overline{R} \). Note that the relay block-Toeplitz filtering matrix \( R \) is redundant since the true design variable \( \overline{R} \) is embedded in the block Toeplitz structure of \( R \). (See (11).) Hence, taking \( R \) as the design variable directly is inefficient and we need reparameterization of the weighted MSE in terms of \( \overline{R} \). This is possible through successive manipulation of the terms constructing the weight MSE shown in (21). First, using similar techniques to those used in [7], we can
express the first term of (21) in terms of \( \overline{R} \) as follows:

\[
\begin{align*}
\text{tr}(\Theta_n \mathbb{E}\{ U_n \hat{y}_n \hat{y}_n^H \}) &= N \text{tr}(\Theta_n \mathbb{E}\{ U_n \bar{G} \bar{F} \bar{R} \mathcal{T}_{t,n}^T V_n \mathbf{s}_n \mathbf{S}_n^H \mathbf{V}_n^H \mathbf{W}_n^* T^H F^H R^H \bar{G}^H U_n^H \}), \\
&= N \text{tr}(V_n^H \mathbf{W}_n^* T^H F^H R^H \bar{G}^H U_n^H \Theta_n U_n \bar{G} \bar{F} \bar{R} \mathcal{T}_{t,n}^T V_n \mathbb{E}\{ \mathbf{s}_n \mathbf{S}_n^H \}), \\
&= N \text{tr}(V_n^H \mathbf{W}_n^* T^H F^H R^H \bar{G}^H U_n^H \Theta_n U_n \bar{G} \bar{F} \bar{R} \mathcal{T}_{t,n}^T V_n), \\
&= N \text{tr}(\Theta_n^{1/2} U_n \bar{G} \bar{F} \bar{R} \mathcal{T}_{t,n}^T V_n V_n^H \mathbf{W}_n^* T^H F^H R^H \bar{G}^H U_n^H \Theta_n^{1/2}), \\
&= N \left[ \text{vec}(R^T \bar{G}^T U_n^T \Theta_n^{1/2}) \right]^T \overline{K}_n \left[ \text{vec}(R^T \bar{G}^T U_n^T \Theta_n^{1/2}) \right]^*, \\
&= N \left[ \text{vec}(R^T) \right]^T \left( \Theta_n^{1/2} U_n \bar{G} \otimes I_Q \right)^T \overline{K}_n \left( \Theta_n^{1/2} U_n \bar{G} \otimes I_Q \right)^* \left[ \text{vec}(R^T) \right]^*, \\
&= N r^T E_1 (\Theta_n^{1/2} U_n \bar{G} \otimes I_Q)^T \overline{K}_n (\Theta_n^{1/2} U_n \bar{G} \otimes I_Q)^* E_1^T r^*, \\
&= r^H Q_{1,n} r, \quad \text{(24)}
\end{align*}
\]

where

\[
\overline{K}_n = I_r \otimes K_n; \quad I_Q = I_{(N+L_r+L_g-2)M_r}; \quad r = \text{vec}(\overline{R}^T); \\
Q_{1,n} = NE_1 (\Theta_n^{1/2} U_n \bar{G} \otimes I_Q)^H \overline{K}_n (\Theta_n^{1/2} U_n \bar{G} \otimes I_Q) E_1^T.
\]

and \( E_1 \) is defined in Appendix A. Here, (a) holds due to \( \text{tr}(\overline{U} \overline{C} \overline{B}) = \text{tr}(\overline{C} \overline{U} \overline{B}) \); (b) holds due to \( \text{tr}(\overline{X} \overline{K}_n \overline{X}^H) = \text{vec}(\overline{X}^T)^T \overline{K}_n \text{vec}(\overline{X}^T)^* \); (c) holds due to the kronecker product identity, \( \text{vec}(\overline{I} \overline{B}) = (C^T \otimes I) \text{vec}(\overline{B}) \); and (d) is obtained because \( \overline{R} = \text{blkToeplitz}(\overline{R}, N + L_g - 1) \) and \( \text{vec}(\overline{R}^T) = E_1^T r \). In a similar way, the remaining terms of (21) and the relay power constraint can also be represented as functions of the design variable \( r \). That is, the second term of (21)
can be rewritten as

\[
\text{tr}(\Theta_n E \{ U_n z_n z_n^H U_n^H \}) = \text{tr} \left( \Theta_n E \left\{ U_n W_{r,n} G R_{r,n} n_{r,n}^H R^H G^H W_{r,n}^H U_n^H \right\} + \Theta_n E \left\{ U_n W_{r,n} n_{d,n} n_{d,n}^H W_{r,n}^H U_n^H \right\} \right),
\]

\[
= \text{tr} \left( R^H G^H W_{r,n}^H U_n^H \Theta_n U_n W_{r,n} G R \{ n_{r,n}^H \} \right) + \text{tr} \left( \Theta_n U_n W_{r,n} E \{ n_{d,n}^H \} W_{r,n}^H U_n^H \right),
\]

\[
= \sigma_r^2 \text{tr}(R^H G^H W_{r,n}^H U_n^H \Theta_n U_n W_{r,n} G R) + \sigma_d^2 \text{tr}(\Theta_n U_n W_{r,n}^H U_n^H),
\]

\[
= \sigma_r^2 \text{tr}(R^H M_n R) + \sigma_d^2 \text{tr}(\Theta_n U_n (w_n^H \otimes I_{N_r}) (w_n \otimes I_{N_r}) U_n^H),
\]

\[
\overset{(a)}{=} \sigma_r^2 \text{vec}(R)^H \overline{M}_n \text{vec}(R) + \sigma_d^2 \text{tr}(\Theta_n U_n (w_n^H w_n \otimes I_{N_r}) U_n^H),
\]

\[
\overset{(b)}{=} \sigma_r^2 \text{tr}^2 E_2 \overline{M}_n E_2^H r + \sigma_d^2 \text{tr}(\Theta_n U_n U_n^H),
\]

\[
= r^H Q_{2,n} r + c_n,
\]

where

\[
\overline{M}_n = I_{(N+L_0+L_\epsilon-2)M_r} \otimes M_n, \quad Q_{2,n} = \sigma_r^2 E_2 \overline{M}_n E_2^H, \quad c_n = \sigma_d^2 \text{tr}(\Theta_n U_n U_n^H),
\]

and \( E_2 \) is defined in Appendix A. Here, (a) follows from the kronecker product identity \((UB \otimes CD) = (U \otimes C)(B \otimes D)\), and (b) is obtained due to \text{vec}(R)^H = r^H E_2. The third term of (21) can be rewritten as

\[
\text{tr} \left( \Theta_n E \{ s_n \tilde{y}_n^H U_n^H \} \right) = \sqrt{N} \text{tr} \left( \Theta_n E \left\{ s_n s_n^H \right\} V_n^H W_{t,n}^* T^F R^H \tilde{G}^H U_n^H \right),
\]

\[
= \sqrt{N} \text{tr} \left( \Theta_n V_n^H W_{t,n}^* T^F R^H \tilde{G}^H U_n^H \right),
\]

\[
= \sqrt{N} \text{tr} \left( R^H \tilde{G}^H U_n^H \Theta_n V_n^H W_{t,n}^* T^F \right),
\]

\[
= \sqrt{N} \text{vec}(R)^H \text{vec}(\tilde{G}^H U_n^H \Theta_n V_n^H W_{t,n}^*) \text{vec}(T^F),
\]

\[
= \sqrt{N} r^H E_2 \text{vec}(\tilde{G}^H U_n^H \Theta_n V_n^H W_{t,n}^*) \text{vec}(T^F),
\]

\[
= r^H q_n,
\]
where \( q_n = \sqrt{N} \mathbb{E}_2 \text{vec}(\tilde{G}^H U_n^H \Theta_n V_n^H W_{t,n}^* T^H F^H) \). Finally, the relay transmit power can be rewritten as

\[
\mathbb{E}\{\text{tr}(y_t y_t^H)\} = \text{tr}\left( R F T_{cp} (W_N \otimes I_{N_t}) V \mathbb{E}\{s s^H\} V^H (W_N^H \otimes I_{N_t}) T^H F^H R^H \right) + \text{tr}\left( \mathbb{R} \mathbb{E}\{n_n n_n^H\} R^H \right) ,
\]

\[
= \text{tr}\left( R F T_{cp} (W_N \otimes I_{N_t}) V V^H (W_N^H \otimes I_{N_t}) T^H F^H R^H + \sigma_r^2 \text{tr}(R R^H) \right) ,
\]

\[
= \text{vec}(R^T)^T \Pi \text{vec}(R^T) ,
\]

\[
= r^H \tilde{\Pi} r ,
\]

(27)

where \( \Pi = I_{(N+L_g-1)M_t} \otimes \Pi \) and \( \tilde{\Pi} = E_T^1 \tilde{\Pi} E_T^T . \)

Based on (24), (25), (26), and (27), the weighted MSE for the \( n \)-th subcarrier channel is expressed as

\[
\text{tr}(\Theta_n M_n) = r^H Q_n r - r^H q_n - q_n^H r + z_n
\]

(28)

where \( Q_n = Q_{1,n} + Q_{2,n} \) and \( z_n = c_n + \text{tr}(\Theta_n) \), and Problem 1-1 is reformulated as

\[
\min_r r^H Q r - r^H q - q^H r + z
\]

s.t. \( r^H \tilde{\Pi} r \leq P_{r,max} \).

(29)

where \( Q = \sum_{n=1}^{N} Q_n , q = \sum_{n=1}^{N} q_n , \) and \( z = \sum_{n=1}^{N} z_n . \)

The key point of the derivation of (29) is that Problem 1-1 reduces to a quadratically constrained quadratic programming (QCQP) problem with a constraint. It is known that QCQP is NP-hard in general. However, QCQP has been well studied in the case that the number of constraints is small. Using the results of [21] and [22], we obtain an optimal solution to Problem 1-1 as follows. Let \( \frac{r}{t} = r , \) where \( t \in \mathcal{C} , \) and \( \tilde{r} = [r^T, t]^T \in \mathcal{C}^{(M_1 L_r M_r + 1) \times 1} . \) Then, we rewrite
equivalently as
\[
\min_{\mathbf{r}} \quad \mathbf{r}^H \mathbf{B}_1 \mathbf{r}
\]
\[
\text{s.t.} \quad \mathbf{r}^H \mathbf{B}_2 \mathbf{r} \leq 0
\]
where
\[
\mathbf{B}_1 = \begin{bmatrix}
    \mathbf{Q} & -\mathbf{q} \\
    -\mathbf{q}^H & \mathbf{z}
\end{bmatrix}
\quad \text{and} \quad
\mathbf{B}_2 = \begin{bmatrix}
    \mathbf{T} & 0 \\
    0 & -P_{r,\text{max}}
\end{bmatrix}.
\]

By defining \( \mathcal{R} := \mathbf{r}\mathbf{r}^H \) and removing the rank-one constraint \( \text{rank}(\mathcal{R}) = 1 \), we obtain the following convex optimization problem:
\[
\min_{\mathcal{R}} \quad \text{tr}(\mathbf{B}_1 \mathcal{R})
\]
\[
\text{s.t.} \quad \text{tr}(\mathbf{B}_2 \mathcal{R}) \leq 0
\]
which is a semi-definite program (SDP) and can be solved efficiently by using the standard interior point method for convex optimization [23]–[26]. With an additional constraint \( \text{rank}(\mathcal{R}) = 1 \), the problem (31) is equivalent to Problem 1-1. That is, if the optimal solution of (31) has rank one, then it is also the optimal solution of Problem 1-1. However, there is no guarantee that an algorithm for solving the problem (31) yields a rank-one solution. In such a case, a rank-one solution from \( \mathcal{R} \) can always be obtained by using the rank-one decomposition procedure [22].

B. Transmit and receive filter optimization

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

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

$$
\min_{V_n, U_n} \text{tr}(\Theta_n M_n) \quad \text{s.t.} \quad \text{tr}(V_n V_n^H) \leq P_{n,\text{max}}, \quad \text{for} \quad n = 0, 1, \cdots, N-1.
$$

(32)

Problem 1-2 has already been solved and the optimal transceiver structure for Problem 1-2 is available in [15] and [27]. It is shown in [15] that the optimal transmit filter and receive filter diagonalize the MIMO channel into eigen-subchannels for any weight matrix. Lemma 1 and Theorem 1 of [15] provide the optimal transmit filter $V_n$ and receive filter $U_n$, and the solution can be expressed as $V_n = \tilde{V}_n \tilde{P}_n$, where $\tilde{V}_n^H \tilde{V}_n = I_{\Gamma}$ and $\tilde{P}_n$ is a diagonal matrix with nonnegative entries s.t. $\text{tr}(\tilde{P}_n^2) = P_{n,\text{max}}$ determining the transmit power of each of $\Gamma$ data streams of the $n$-th subcarrier MIMO channel. (Please refer to [15].)

Note that the solution to Problem 1-2 only optimizes the power allocation within multiple data streams for each subcarrier when the transmit power is allocated to each subcarrier. Now, consider the problem of total source power allocation $P_{s,\text{max}}$ to subcarrier channels. Here, we exploit the **diagonalizing** property [15] of the solution to Problem 1-2, take the direction information only for the transmit filter from the solution to Problem 1-2, and apply alternating optimization. That is, when the relay filtering matrix $R$ from Problem 1-1 and the normalized transmit filters $\{\tilde{V}_n\}$ and the receive filters $\{U_n\}$ from Problem 1-2 are given, each subcarrier MIMO channel is diagonalized into eigen-subchannels. Thus, the effective parallel MIMO channel (20) for the
\( n \)-th subcarrier is rewritten as

\[
\hat{s}_n = \sqrt{N} \mathbf{U}_n \mathbf{GRFT}_n \mathbf{V}_n s_n + \mathbf{U}_n \mathbf{W}_r \mathbf{n}_d + \mathbf{U}_n \mathbf{W}_r \mathbf{n}_d,
\]

\[
= \sqrt{N} \mathbf{U}_n \mathbf{GRFT}_n \mathbf{V}_n \tilde{\mathbf{P}} \mathbf{s}_n + \mathbf{U}_n \mathbf{W}_r \mathbf{n}_d + \mathbf{U}_n \mathbf{W}_r \mathbf{n}_d,
\]

\[
= \mathbf{D}_n \tilde{\mathbf{P}} \mathbf{s}_n + \mathbf{U}_n \mathbf{W}_r \mathbf{n}_d + \mathbf{U}_n \mathbf{W}_r \mathbf{n}_d,
\]

(33)

(34)

where \( \mathbf{D}_n = \text{diag}(d_n[1], d_n[2], \ldots, d_n[\Gamma]) \) 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 \( \tilde{\mathbf{P}}_n = \text{diag}(p_n[1], p_n[2], \ldots, p_n[\Gamma]) \).

Therefore, we obtain \( N \Gamma \) parallel eigen-subchannels for the overall MIMO-OFDM system as

\[
\hat{s}_n[k] = d_n[k] p_n[k] s_n[k] + n_n[k], \quad \text{for } n = 0, 1, \ldots, N - 1 \text{ and } k = 1, 2, \ldots, \Gamma,
\]

(35)

where \( n_n[k] = \mathbf{U}^H_{n,k} \mathbf{W}_r \mathbf{n}_d + \mathbf{U}^H_{n,k} \mathbf{n}_d \) and \( \mathbf{U}^H_{n,k} \) 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

\[
\sum_{n=0}^{N-1} \sum_{k=1}^{B} \theta_{nk} \mathbb{E}\{ |\hat{s}_n[k] - s_n[k]|^2 \} = \sum_{n=0}^{N-1} \sum_{k=1}^{\Gamma} \theta_{nk} (d_n[k]^2 p_n[k]^2 - 2 d_n[k] p_n[k] + c_n[k])
\]

(36)

where \( c_n[k] = \sigma^2 U^H_{n,k} \mathbf{W}_r \mathbf{GRFT}_n \mathbf{GRFT}_n^H \mathbf{W}_r \mathbf{U}_n k + \sigma^2 d^2 U^H_{n,k} \mathbf{U}_n k + 1 \), 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 filters \( \{\tilde{\mathbf{V}}_n\} \), and receive filters \( \{\mathbf{U}_n\} \),

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

(37)

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

\[ p_n[k] = \left( \frac{\theta_{nk} \mathcal{d}_n[k]}{\theta_{nk} \mathcal{d}_n[k]^2 + \mu} \right)_+ \quad \text{s.t.} \quad \sum_{n=0}^{N-1} \sum_{k=1}^{\Gamma} \left( \frac{\theta_{nk} \mathcal{d}_n[k]}{\theta_{nk} \mathcal{d}_n[k]^2 + \mu} \right)^2 = P_{s,max}. \] (38)

Proof: See Appendix B

The solution in Proposition 1 allocates power inverse-proportionally to the power of the effective channel \( \mathcal{d}_n[k] \) in most cases similarly to the method in [27].

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 \( \{\tilde{P}_n\}, \{\tilde{V}_n\}, \) and \( \{U_n\} \) for \( n = 0, 1, \ldots, N-1 \). For example, \( p_n[k] = \frac{P_{s,max}}{N \Gamma} \), \( \tilde{V}_n = \mathbf{I}_{N_t \times \Gamma} \), and \( U_n = \mathbf{I}_{\Gamma \times N_r} \).

Step 2: Solve Problem 1-1 and obtain \( \mathbf{R} \).

Step 3: Solve Problem 1-2 and obtain \( \{\tilde{V}_n, U_n\} \).

Step 4: Solve Problem 1-3 and obtain \( \{\tilde{P}_n\} \).

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 \( \{\tilde{V}_n, U_n, \tilde{P}_n\} \) denoted by \( \mathcal{M} (\mathbf{R}, \tilde{V}_n, U_n, \tilde{P}_n) \).

Let \( \mathbf{X}^{(i)} \) denotes the solution at the \( (i) \)-th step. Then, it is easy to see that \( \mathcal{M} (\mathbf{R}^{(0)}, \tilde{V}_n^{(0)}, U_n^{(0)}, \tilde{P}_n^{(0)}) \geq \mathcal{M} (\mathbf{R}^{(1)}, \tilde{V}_n^{(0)}, U_n^{(0)}, \tilde{P}_n^{(0)}) \geq \mathcal{M} (\mathbf{R}^{(2)}, \tilde{V}_n^{(2)}, U_n^{(2)}, \tilde{P}_n^{(0)}) \geq \mathcal{M} (\mathbf{R}^{(3)}, \tilde{V}_n^{(2)}, U_n^{(2)}, \tilde{P}_n^{(0)}) \geq \cdots \geq 0 \) 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. In this way, the proposed algorithm converges by the monotone convergence theorem 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 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 (33) is equivalent to the weighted MSE minimization when the weight matrix \( \Theta_n \) is set as a diagonal matrix composed of the eigenvalues of \( H^H \Sigma_n^{-1} H \), where \( H = \sqrt{N} \tilde{G} \tilde{R} \tilde{F} \tilde{W}_{t,n}^T \) is the effective MIMO channel matrix and \( \Sigma_n \) is the effective noise covariance matrix of the \( n \)-th subcarrier MIMO channel (33). (See Lemma 3 of [15].) Exploiting this result, we propose our algorithm to design the linear transceiver and the relay filter to maximize the sum rate below.

Algorithm 2: Given parameters: \( F, G, L_r, P_{s,max}, \) and \( P_{r,max} \)

Step 1: Initialize \( \{\Theta_n\}, \{\tilde{P}_n\}, \{\tilde{V}_n\}, \) and \( \{U_n\} \) for \( n = 0, 1, \cdots, N - 1 \). For example, \( \Theta_n = I, p_n[k] = \frac{P_{s,max}}{N \Gamma}, \tilde{V}_n = I_{N_t \times \Gamma}, \) and \( U_n = I_{\Gamma \times N_r} \).

Step 2: Solve Problem 1-1 and obtain \( R \).

Step 3: Solve Problem 1-2 and obtain \( \{\tilde{V}_n, U_n, \Theta_n\} \).

Step 4: Compute \( \{\tilde{P}_n\} \) 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.

\(^3\)When \( R \) is given, all the parallel subcarrier MIMO channels are determined and a solution \( \{\tilde{V}_n, U_n, \Theta_n\} \) is given by Lemma 1 and Theorem 1 of [15].
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 an SDP problem of size \( M_t M_r L_g \). Due to the existence of fast approximate algorithms for solving SDP problems [28], [29], the proposed algorithm is implementable if the number of iterations for convergence is not so large, which will be seen in Fig. 5. For other practical issues such as channel estimation and self-interference caused by full-duplex operation, please see [7].

### 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, \( F_k \) and \( G_k \), was generated i.i.d according to a Rayleigh distribution, i.e., \( F_k(i, j) \overset{\text{i.i.d.}}{\sim} \mathcal{CN}(0, \sigma_f^2) \) and \( G_k(i, j) \overset{\text{i.i.d.}}{\sim} \mathcal{CN}(0, \sigma_g^2) \), where \( \sigma_f = \sigma_g = 1 \). The SR channel length and the RD channel length were set as \( L_f = L_g = 3 \), and \( N_t = M_r = M_t = N_r = 2 \). 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., \( P_{s,\text{max}} = 100 \). (From here on, all dB power values are relative to \( \sigma_r^2 = \sigma_d^2 = 1 \).

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. Figs. 2 and 3 show the resulting sum MSE over all subcarriers. For the curves in the figures,
200 channels were randomly realized with $L_f = L_g = 3$ and each plotted value is the average over the 200 channel realizations. As expected, it is seen in Figs. 2 and 3 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.

Next, we investigated the BER performance corresponding to Fig. 2. Here, we assumed uncoded QPSK modulation for each subcarrier channel. From the result of Fig. 2, we obtained the SNR of each subcarrier channel of the total $N = 16$ subcarrier channels for the designed FF relay filter, transmit filer, receive filter 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. 4. It is seen in Fig. 4 that the FF relay significantly improves the BER performance over the AF relay. Next, we tested the convergence property of the proposed algorithm, and Fig. 5 shows the result. It is seen that the proposed algorithm converges with a few iterations.

Fig. 2: Sum MSE versus FF relay transmit power.
Fig. 3: Sum MSE versus relay filter length.

Fig. 4: Overall BER versus FF relay transmit power.
Finally, 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. 6 shows the result. Again, for the figure 200 channels were randomly realized with $L_f = L_g = 3$ and 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. 6 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.

![Sum MSE versus the number of iteration.](image)

**Fig. 5: Sum MSE versus the number of iteration.**

V. CONCLUSION

In this paper, we have considered the joint design of the linear transceiver and the FF relay for MIMO-OFDM systems for 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 MIMO transceiver at the source and the destination 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 quadratically constrained quadratic program (QCQP) and have proposed a solution to this QCQP problem based on conversion to a semi-definite program (SDP). We have provided some numerical results to evaluate the performance gain of the FF relaying scheme over the simple AF scheme for MIMO-OFDM systems. 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.

Fig. 6: Sum rate versus FF relay transmit power.
APPENDIX A

**E₁ AND E₂ MATRICES**

E₁ and E₂ are \( M_L L_r M_r \times M_L (N + L_r + L_g - 2)(N + L_g - 1) M_r \) matrices and defined as follows:

\[
E_1 =
\begin{bmatrix}
I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

\[
\otimes I_{M_r}
\]

\[\text{(39)}\]

where \( I = I_{L_r} \).
\[ \mathbf{E}_2 = \begin{bmatrix} \mathbf{E}_1 & \mathbf{E}_{M_t+1} & \mathbf{E}_{(L_r-1)M_t+1} & \mathbf{E}_{(L_r)M_t+1} & \mathbf{E}_{(N+L_g-2)M_t+1} & \mathbf{0} \\ \vdots & \mathbf{E}_1 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \mathbf{E}_1 & \mathbf{E}_{M_t+1} & \mathbf{E}_{(N+L_g-L_r-1)M_t+1} & \mathbf{E}_{(N+L_g-L_r)M_t+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{E}_{M_t} & \mathbf{E}_{M_t+M_t} & \mathbf{E}_{(L_r-1)M_t+M_t} & \mathbf{E}_{(L_r)M_t+M_t} & \mathbf{E}_{(N+L_g-2)M_t+M_t} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \]

The \( (N + L_g - 1) \)-th block and \( (N + L_g) \)-th block are

\[ \begin{align*} 
\mathbf{E}_{(N+L_g-2)M_t+1} & \\
\vdots & \\
\mathbf{E}_{(N+L_g-2)M_t+M_t} & \\
\end{align*} \]

where \( \mathbf{e}_i^T \) is the \( i \)-th row of \( \mathbf{I}_{(N+L_g-1)M_tM_t^*} \).
APPENDIX B

Proof of Proposition 1

The Lagrangian of (37) is given by

\[ L(p_n[k], \mu) = \sum_{n=0}^{N-1} \sum_{k=1}^{B} \theta_{nk}(d_n[k]^2p_n[k]^2 - 2d_n[k]p_n[k] + c_n[k]) + \mu\left(\sum_{n=0}^{N-1} \sum_{k=1}^{B} p_n[k]^2 - P_{s,max}\right) \]

\[- \sum_{n=0}^{N-1} \sum_{k=1}^{B} \lambda_{n,k}p_n[k] \]

(41)

where \( \mu \in \mathbb{R} \) 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 (37) is a convex optimization problem:

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

(42)

\[ \mu \in \mathbb{R}, \quad \lambda_{n,k} \geq 0, \]

(43)

\[ \lambda_{n,k}p_n[k] = 0 \]

(44)

\[ \nabla_{p_n[k]} L = 2\theta_{nk}d_n[k]^2p_n[k] - 2\theta_{nk}d_n[k] + 2\mu p_n[k] - \lambda_{n,k} = 0 \]

(45)

for \( n = 0, 1, \ldots, N - 1 \) and \( k = 1, \ldots, B \).

The gradient (45) 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 (43) and (44), we get

\[ \mu p_n[k] \geq \theta_{nk}d_n[k] - \theta_{nk}d_n[k]^2p_n[k] \]

(46)

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

(47)

Let us consider the case that \( p_n[k] = 0 \). Then, (46) is satisfied only if \( d_n[k] = 0 \) because \( d_n[k] \geq 0 \). If \( p_n[k] > 0 \), \( p_n[k] = \left(\frac{\theta_{nk}d_n[k]}{\theta_{nk}d_n[k]^2 + \mu}\right) \) by the complementary slackness (47). This also satisfies (46). Therefore, we get the desired result satisfying the primal constraints (42). \( \square \)
REFERENCES

[1] A. El Gamal, M. Mohseni, and S. Zahedi, “Bounds on capacity and minimum energy-per-bit for AWGN relay channels,” IEEE Trans. Inform. Theory, vol. 52, 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, pp. 2024 – 2035, Apr. 2009.

[3] 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, pp. 1251 – 1262, Mar. 2010.

[4] 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, pp. 228 – 239, Jan. 2011.

[5] Y. Sung and C. Kim, “The capacity for the linear time-invariant Gaussian relay channel,” ArXiv. http://arxiv.org/abs/1109.5426, Sep. 2011.

[6] 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, pp. 4360 – 4375, Aug. 2012.

[7] D. Kim and J. Seo and Y. Sung, “Filter-and-Forward Transparent Relay Design for OFDM systems,” IEEE Trans. Vech. Technol., vol. 62, pp. 1 – 16, Nov. 2013.

[8] D. Kim and J. Seo and Y. Sung, “Filter-and-forward relay design for OFDM systems for quality-of-service enhancement,” in Proc. of APSIPA ASC, Hollywood, CA, Dec. 2012.

[9] I. Hammerström and A. Wittneben, “On the optimal power allocation for nongenerative OFDM relay links,” in Proc. of ICC, vol. 10, pp. 4463–4468, June. 2006.

[10] T. Ng and W. Yu, “Joint optimization of relay strategies and resource allocations in cooperative cellular networks,” IEEE J. Sel. Areas Communi., vol. 25, pp. 328 – 339, Feb. 2007.

[11] M. Dong and S. Shahbazpanahi, “Optimal spectrum sharing and power allocation for OFDM-based two-way relaying,” in Proc. of ICASSP, pp. 3310–3313, Mar. 2010.

[12] W. Dang, M. Tao, H. Mu and J. Huang, “Subcarrier-pair based resource allocation for cooperative multi-relay OFDM systems,” IEEE Trans. Wireless Commun., pp. 1640 – 1649, May. 2010.

[13] Z. Fang and Y. Hua and J. C. Koshy, “Joint source and relay optimization for a non-regenerative MIMO relay,” in Proc. of 2006 IEEE SAM, 2006.

[14] S. Simoens and O. Munoz-Medina and J. Vidal and A. del Coso, “On the Gaussian MIMO relay channel with full channel state information,” IEEE Trans. Signal Process., vol. 57, Sep. 2009.

[15] H. Sampath and P. Stoica and A. Paulraj, “Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion,” IEEE Trans. Commun., vol. 49, pp. 2198 – 2206, Dec. 2001.

[16] D. Guo and S. Shitz and S. Verdu, “Mutual information and minimum mean-square error in Gaussian channels,” IEEE Trans. Inform. Theory, vol. 51, pp. 1261 – 1282, Apr. 2005.
[17] D. Palomar and S. Verdu, “Gradient of mutual information in linear vector gaussian channels,” IEEE Trans. Inform. Theory, vol. 52, pp. 141 – 154, Jan. 2006.

[18] S. S. Christensen and R. Argawal and 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, pp. 1 – 7, Dec. 2008.

[19] Q. Shi and M. Razaviyayn and 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, pp. 4331 – 4340, Sep. 2011.

[20] R. M. Gray, Toeplitz and Circulant Matrices : A Review. Now publishers, 2006

[21] Y. Huang and S. Zhang, “Complex matrix decomposition and quadratic programming,” Math. Oper. Res., vol. 32, pp. 758 – 768, Aug. 2007.

[22] W. Ai and Y. Huang and S. Zhang, “New results on Hermitian matrix rank-one decomposition,” Math. Programm., vol. 128, pp. 253 – 283, Aug. 2009.

[23] S. Boyd and L. Vandenberghe, Convex Optimization. New York: Cambridge University Press, 2004.

[24] C. Helmberg, “Semidefinite programming,” European Journal of Operation Research , vol. 137, pp. 461 – 482, 2002.

[25] J. F. Sturm, “Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones,” Optimization Methods ans Software, Special Issue on Interior Point Methods, vol. 11/12, pp. 625 – 563, 1999.

[26] M. Grant and S. Boyd, (CVX): Matlab Software for Disciplined Convex Programming. ver. Version 1.21, 2011.

[27] H. Sampath and A. Paulraj, “Joint transmit and receiver optimization for high data rate wireless communications using multiple antennas,” in Proc. of Asilomar Conf. Signals, Systems and Computers, vol. 1, 1999.

[28] S. Arora and E. Hazan and S. Kale, “Fast algorithms for approximate semidefinite programming using the multiplicative weights update method,” in Proc. of 46th FOCS, pp. 339–348, 2005.

[29] S. Arora and E. Hazan and S. Kale, “The Multiplicative Weights Update Method: a Meta-Algorithm and Applications,” Theory of Computing (TOC), vol. 8, pp. 121 – 164, 2012.