MMSE Relaying Design for Multi-Antenna Two-Hop Downlinks with Finite-Rate Feedback

Wei Xu and Xiaodai Dong

Abstract—We consider a MIMO relay downlink system using precoding and limited feedback from two-hop channels. Since the conventional precoding matrices are directly obtained by treating the quantized channel state information (CSI) feedback as the real CSI, we propose an improved relay precoding strategy by taking the effect of channel quantization error into account. The proposed relaying technique is designed based on the minimum mean square error (MMSE) criterion and is robust to the channel uncertainties due to quantized CSI feedback. Numerical results verify the effectiveness of the proposed scheme.

I. INTRODUCTION

Multi-antenna relaying has become a promising technique for improving the system capacity and increasing the coverage at a low cost in next generation wireless communication systems including 3GPP LTE-Advanced and IEEE 802.16m [1], [2]. For multiple-input multiple-output (MIMO) systems, studies [3]–[5] have investigated the issue of using relay stations in multiuser MIMO networks to deal with remote mobile users, or equivalently to expand the coverage. In [3], sum capacity bounds are derived for the system. However, in order to achieve the full channel capacity, it is hard to obtain the optimal linear precoding scheme. As a result, a number of suboptimal methods are developed in [4], [5].

Although the recently developed linear precoding schemes are able to achieve a significant part of the full channel capacity, they generally require complete channel state information (CSI) of both two-hop links at the transmitter sides (CSIT). In practical wireless applications, however, the requirement of full CSIT could be infeasible to implement. Particularly in frequency division duplexed (FDD) systems, a large amount of CSI need to be sent back to the transmitters through feedback links, hence imposing a heavy feedback burden. To deal with the problem, one common way is to exploit limited feedback precoding which has been proposed in [6]–[8] for conventional multi-antenna systems without relay. Specifically for multiuser MIMO systems, [7] presents a pre-processing scheme with zero-forcing beamforming (ZFBF) and quantized CSI feedback, called finite-rate feedback (FRF). Further by considering the effects of both noise and CSI imperfection, [8] develops a robust beamforming design based on a minimum mean square error (MMSE) criterion.

Recently, an increasing attention has been paid to design limited feedback based precoding in relay systems. For a three-node two-hop system, an amplify-and-forward (AF) relaying scheme using Grassmannian beamforming is proposed in [9]. In [10], this scheme is extended to precoding design supporting multiple simultaneous data stream transmissions. For a multiuser relay downlink, the multiuser interference cancelation should be considered. Although the optimal precoding design for source and relay is still hard to obtain [3]–[5], recent work in [11] has presented an efficient limited feedback precoding for this system. In [11], SVD-based precoding is utilized for the source node and ZFBF-FRF is deployed at the relay. Generally, in these studies, the limited feedback model of FRF is considered for precoding design. These limited feedback preprocessing schemes directly employ precoding/beamforming techniques by assuming prefect CSIT and treating the quantized CSI feedback as the true channel information. However, to the best of our knowledge, few literature has considered to design robust relaying scheme for the multiuser system with imperfect CSI. A couple of modified beamforming approaches have been presented in [7], [8] for conventional multiuser systems without any relay, while the robust relay precoding design with respect to the limited feedback is still unknown.

In this study, we consider the multi-antenna downlink system where the base station (BS) broadcasts multiple data streams to a number of remote users via the help of a multi-antenna AF relay. Precoding matrices at both the BS and the relay are obtained through a finite-rate feedback channel from the corresponding receivers. We study the problem of robust relay precoding design via minimizing the expected minimum square error (MSE) with respect to the quantization error of both two-hop channels.

Notations: We use $A^T$, $A^H$, $\text{Tr}(A)$, and $\|A\|_F$ to denote the transpose, the conjugate transpose, the trace, and the Frobenius norm of a matrix $A$, respectively. $\|a\|$ denotes the Euclidean norm of a vector $a$. $\text{diag}\{\cdots\}$ represents a diagonal matrix with the given elements on the diagonal. $E[\cdot]$ stands for the statistical expectation of a random variable and $\Re\{\cdot\}$ returns the real part of the input.

II. SYSTEM MODEL

Consider a multiuser downlink channel in Fig. 1 with an $M$-antenna BS serving $K$ single-antenna users through an $N$-antenna relay station (RS). For simplicity, we assume that the number of antennas at the BS is no less than that at the RS, i.e., $M \geq N$. In order to focus our study on the effect of limited feedback scheme, we consider the system with $K = N$.

Let $x \in \mathbb{C}^{N \times 1}$ be the transmit symbol vector at the BS. Note that the energy of the symbols is normalized, i.e.,
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2010 proceedings.

\[ \mathbb{E}[xx^H] = I_N. \]

Before transmitting this vector to the RS, the BS first pre-processes \( x \) by an \( M \times N \) precoding matrix \( W \). Then, at the RS, the received symbol vector is precoded by an \( N \times N \) matrix \( F \) before forwarded to the user terminal. Thus, the finally received symbol at the \( k \)th user can be written as

\[ y_k = g_k^H F HW x + g_k^H F n_k + n_k \]  

(1)

where \( g_k \in \mathbb{C}^{N \times 1} \) is the channel vector between the RS and the \( k \)th user, \( H \in \mathbb{C}^{N \times M} \) is the channel information between the BS and the RS, and \( n_k \) and \( n_k \) are the Gaussian noise with zero mean and unit variance at the relay and \( k \)th user, respectively. By stacking the received signals at all users into \( y = [y_1 \ldots y_N]^T \), we have

\[ y = GH x + n_U \]  

(2)

where \( G = [g_1 \cdots g_N] \) is the \( N \) RS-to-user channels and \( n_U \) is the noise at \( N \) user sides.

In this study, we assume that perfect channel knowledge available at the receiver sides. That is, RS perfectly knows \( H \) and each user \( k \) has the knowledge of \( g_k \). We exploit the structure of limited feedback precoding developed in [11]. By applying the singular value decomposition (SVD) to \( H \), it gives

\[ H = U \Sigma V^H \]  

(3)

where \( \Sigma \) is an \( N \times N \) diagonal matrix with non-zero singular values of \( H \) and \( V \) has \( N \) columns. As depicted in Fig. 1, the RS sends a quantized version of \( V \) via a finite number of bits back to the BS for precoding. The quantization process of \( V \) follows the conventional vector quantization method [6]-[8]. In detail, the RS uses \( B_1 \) bits to quantize each column of \( V \) with a codebook containing \( 2^{B_1} \) unit-norm \( M \times 1 \) complex vector codewords. Let \( v_i \) be the \( i \)th column of \( V \) and denote \( C_i = \{c_{i1}, \cdots, c_{i2^{B_1}}\} \) as the quantization codebook with unit-norm vectors. The quantized vector \( \tilde{v}_i \) of \( v_i \) is determined according to \( \tilde{v}_i = \arg \max_{c \in C_i} |v_i^H c| \). Then, the quantization of \( V \) is given by \( \tilde{V} = [\tilde{v}_1 \cdots \tilde{v}_N] \). The precoding matrix at the BS is designed by

\[ W = \sqrt{n} \tilde{V} \]  

(4)

where \( \rho_1 \) is a scalar to guarantee the transmit power constraint at the BS. Denote \( P_1 \) as the power constraint at the BS. Then, by considering \( \text{Tr}(WW^H) = P_1 \), it is easy to obtain \( \rho_1 = P_1/N \).

For the channels from RS to users, the precoding matrix \( F \) follows the structure given by

\[ F = F_1 U^H \]  

(5)

where \( U \) is perfectly known at the RS and \( F_1 \) is a properly designed precoding matrix according to the quantized CSI feedback from user terminals. By exploiting the model of FRF [7], in this study only the quantized channel direction information (CDI), i.e., the quantization of normalized channel information \( g_k^2 = g_k^2/\|g_k\| \), is available at the RS for precoding design. Denote \( \tilde{g}_k \) as the quantization of \( g_k \). Each user \( k \) quantizes its CDI with \( B_2 \) bits via the same vector quantization method and then sends the index of \( \tilde{g}_k \) back to the RS. Moreover, from (5), the average transmit power constraint at the RS can be described by

\[ \mathbb{E}_{H,W} [\text{Tr}(FHWW^HHF^HH + FF^H)] = P_2 \]  

(6)

where \( P_2 \) is the power constraint at the RS.

III. ROBUST MMSE RELAYING DESIGN

This section focuses on the design of relating matrix \( F_1 \) with the knowledge of only quantized CDI from users. In [11], the ZFBF-FRF technique has been utilized for designing \( F_1 \) in a straightforward manner by simply treating \( \tilde{g}_k \) as the real channel information. However, the effects of both the CDI quantization error and the lack of channel magnitude information have not been considered in the precoding design. As a result, in order to include these effects and improve the relay precoding design in FRF systems, we propose the robust relating matrix \( F_1 \) via minimizing the MSE averaged over the two-hop channel realizations conditioned on the known information of quantized CDI.

Start with the MSE cost function of the received symbol in the system. According to the definition of MSE in [12] and from (2), the MSE is defined by

\[ e(F_1, \beta) = \mathbb{E}_{x,n_R,n_U} [||\beta^{-1}y - x||^2] \]

\[ = \mathbb{E}_{x,n_R,n_U} [||\beta^{-1}GHFWx + \beta^{-1}GFWU^{-1} + \beta^{-1}F_1n_R - x||^2] \]

\[ = ||\beta^{-1}GHFW1 + \beta^{-2}GFW1||^2_F + ||\beta^{-1}n_R||^2_F \]

\[ + \beta^{-2}N \]  

(7)

where \( \beta \) is a positive scalar at the receiver. Equality (7) is obtained by using (3) and \( n_R = U^H n_U \) is the equivalent noise with zero-mean and unit variance. Since only partial channel information in terms of quantized CDI is available at the transmitter, it is natural to modify the MSE objective as an expected MSE with respect to the random CSI conditioned on known channel quantization information. Furthermore, by considering the transmit power constraint at the BS as described in (6), the optimization problem of determining \( F_1 \) can be formulated as follows:

\[ \min_{F_1, \beta} \mathbb{E}_{V,G} [e(F_1, \beta)] \]

s.t. \( \mathbb{E}_{H,W} [\text{Tr}(FHWW^HHF^HH + FF^H)] = P_2 \).  

(9)

In order to obtain a closed-form solution to the above problem, we need to rewrite (9) into a simpler and concise form as the function of \( F_1 \) and \( \beta \) before solving this problem.
We first consider the objective function in (9). From [7], any channel vector \( g_k \) can be decomposed by
\[
g_k = \|g_k\| (\sqrt{1 - \epsilon_k} g_k + \sqrt{\epsilon_k} s_k) \tag{10}
\]
where \( \epsilon_k = 1 - \|g_k\|^2 \) is the quantization error and \( s_k \) is a unit-norm vector orthogonal to \( g_k \). Accordingly, by defining \( G = [g_1 \cdots g_N] \), the stacked channel matrix \( G \) follows
\[
G = (G(I - Z_g)^{1/2} + S_g Z_g^{1/2}) A_g \tag{11}
\]
where \( Z_g = \text{diag}(\epsilon_1 \cdots \epsilon_N) \), \( S_g = [s_1 \cdots s_N] \), and \( A_g = \text{diag}(\|g_1\| \cdots \|g_N\|) \). Subsequently, by substituting (8) into the objective function in (9) and applying (11), the objective function is rewritten by

\[
E_{V,G,V,G} [e(F_1, \beta)] = E\left[\sqrt{\rho_1} \beta^{-1} A_g (I - Z_g)^{1/2} \hat{G} H + Z_g^{1/2} S_g \right] F_1 \Sigma V H \hat{V} \bigg| I_{p_g} \bigg] + \beta^{-2} E\left[\|A_g (I - Z_g)^{1/2} \hat{G} H + Z_g^{1/2} S_g \| F_1 \bigg| I_{p_g} \bigg] \tag{12}
\]

\[
+ \beta^{-2} \left[\|A_g (I - Z_g)^{1/2} \hat{G} H F_1 \| I_{p_g} \bigg] + (\beta^{-2} + 1) N \tag{12}
\]

where the last equality uses both \( E\left[ S_g G \right] = 0 \) and \( E\left[ S_g \right] = 0 \) because any vector \( s_k \) is isotropically distributed in the unit-norm sphere [7], and follows by the fact that \( S_g \) is independent to other random matrices. Subsequently, the expectations in (12) will be calculated term by term as follows.

The first term in (12) can be further calculated as
\[
\rho_1 \beta^{-2} E\left[\|A_g (I - Z_g)^{1/2} \hat{G} H F_1 \| \right] = \rho_1 \beta^{-2} \left[\|A_g (I - Z_g)^{1/2} \hat{G} H F_1 \| \right] \tag{13}
\]

where we use \( E\left[ A_g^2 \right] = N I \) and \( E\left[ \Sigma^2 \right] = M I \) because these random matrices are diagonal and composed of singular values of Wishart distributed matrices [14]. Moreover, equality (13) also uses \( E\left[ I - Z_g \right] = (1 - \frac{N - 1}{N}) I \) [7] and \( E\left[ \hat{G} H \hat{V} \hat{V} H V \Sigma F_1 H \right] = (1 - \frac{N - M - N}{N}) I \) [11] where \( \tau_g = 2^{-b_2/(N - 1)} \) and \( \tau_v = 2^{-b_2/(M - 1)} \).

Then, by using the above existing results and further by applying [8, Eq. (13)]
\[
E\left[S_g S_g^H \right] = \frac{N}{N - 1} I - \frac{1}{N - 1} G G^H, \tag{14}
\]

the second expectation in (12) is equivalent to
\[
\rho_1 \beta^{-2} E\left[\|\Lambda_g Z_g^{1/2} S_g H F_1 \| \right] = \rho_1 \beta^{-2} \tau_g (N - (M - N) \tau_v) \left\{ Tr \left(F_1^H F_1^H \right) \right\} \tag{15}
\]

Considering the third term in (12), we have
\[
\sqrt{\rho_1 \beta^{-1}} \left\{ Tr \left( A_g (I - Z_g)^{1/2} \hat{G} H F_1 \right) \right\} = \sqrt{\rho_1 \beta^{-1}} \left\{ Tr \left( \Sigma V H \hat{V} A_g (I - Z_g)^{1/2} \right) \right\}. \tag{16}
\]

According to Lemma 1 in Appendix A, it is not hard to conclude that
\[
E\left[\Sigma V H \hat{V} A_g (I - Z_g)^{1/2} \right] = a I \tag{17}
\]
where \( a \) is a proper real value because all the component matrices in the expectation are independent and their expectations are all scaled identity matrices. Therefore, from (16) and (17), the third term in (12) is equivalent to
\[
-2 \sqrt{\rho_1 \beta^{-1}} \left\{ Tr \left( A_g (I - Z_g)^{1/2} \hat{G} H F_1 \right) \right\} = -2 a \beta^{-1} \sqrt{\rho_1} \left\{ Tr \left( \hat{G} H F_1 \right) \right\}. \tag{18}
\]

With similar techniques, the last 2 expectation terms in (12) can also be calculated as
\[
\beta^{-2} E\left[\|A_g Z_g^{1/2} S_g H F_1 \| \right] + \beta^{-2} E\left[\|A_g Z_g^{1/2} S_g H F_1 \| \right] = \beta^{-2} (N - (N - 1) \tau_g) \left\{ Tr \left( \hat{G} H F_1 \right) \right\} + (\beta^{-2} + 1) N \tag{19}
\]

By combining (13)–(19), the objective function in (9) is now equivalent to
\[
E_{V,G,V,G} [e(F_1, \beta)] = \rho_1 \beta^{-2} (N - (N - 1) \tau_g) (M - (M - N) \tau_v) \left\{ Tr \left( \hat{G} H F_1 \right) \right\} + \rho_1 \beta^{-2} (M - (M - N) \tau_v) \left\{ Tr \left( \hat{G} H F_1 \right) \right\} + (\beta^{-2} + 1) N \tag{20}
\]

where \( c_1 = N (1 - \tau_g) (1 + \rho_1 (M - (M - N) \tau_v)) \) and \( c_2 = N N (1 + \tau_v) \). Now, we consider the power constraint in (9). By substituting (3)–(5) into the power constraint expression, it can be rewritten by
\[
P_2 = \rho_1 E\left[ Tr \left( F_1 \Sigma V H \hat{V} H V \Sigma F_1 H \right) \right] + Tr \left( F_1 F_1^H \right) = (\rho_1 (M - (M - N) \tau_v) + 1) Tr \left( F_1 F_1^H \right) \tag{21}
\]
Hence, the power constraint in (9) can be simplified to
\[ \text{Tr} \left( \mathbf{F}_1 \mathbf{F}_1^H \right) = P_2 / [\rho_1 (M - (M - N) r_i) + 1] \triangleq P'_2. \] (22)

Thus far, by further substituting (22) into (20), the primal optimization problem in (9) is equivalent to
\[
\begin{aligned}
\min_{\mathbf{F}_1, \beta} & \quad c_1 \beta^{-2} \| \mathbf{G}^H \mathbf{F}_1 \|_F^2 - 2 a \sqrt{\rho_1} \beta^{-1} \mathbb{E} \left\{ \text{Tr} \left( \mathbf{G}^H \mathbf{F}_1 \right) \right\} + c_3 \beta^{-2} \\
\text{s.t.} & \quad \text{Tr} \left( \mathbf{F}_1 \mathbf{F}_1^H \right) = P'_2 \\
\end{aligned}
\] (23)

where we define \( c_3 = N (1 + \tau_y P_2) \).

Now that the simplified problem in (23) can be solved with a closed-form solution by exploiting the KKT conditions [13]. The details are provided in Appendix B, and the final optimal solution to \( \mathbf{F}_1 \) is given by
\[
\mathbf{F}_1 = \gamma \left( \mathbf{G} \hat{\mathbf{G}}^H + \frac{(1 + \tau_y P_2)}{P_2 (1 - \tau_y)} \mathbf{I} \right)^{-1} \mathbf{G} 
\] (24)

where \( \gamma \) is a scaling factor given in (39) to guarantee the transmit power constraint in (23). It is interesting to see that the optimal solution has the same structure as the MMSE precoding developed in [8] for systems without any relay. From (24), we find that \( \mathbf{F}_1 \) is affected by the BS transmit power \( P_1 \) and the quantization error of the BS precoding only through the scaling factor \( \gamma \).

Moreover, by comparing the MMSE precoder with conventional ZF precoding scheme, we find that both schemes have similar computational complexity because their difference only lies in the calculation of a regular factor.

IV. NUMERICAL RESULTS

In this section, we present the numerical results of the proposed precoding scheme via computer simulations. Both the achievable sum rate and the bit error rate (BER) performance are provided for comparison with the conventional ZFBF-based relay precoding design in [11].

Fig. 2 depicts the achievable sum rate of different precoding schemes under \( M = 4 \) and \( N = 4 \). This figure shows that the proposed MMSE precoding scheme outperforms the conventional ZF precoding scheme under different cases and the performance improvement increases with the transmit power from low to moderate SNR values. At high SNRs, the sum rate performance of both schemes are upper bounded due to the interference-limited effect in limited feedback systems [7]. Notice that, for reference, the performance achieved by the proposed scheme with perfect CSI at the transmitters (i.e., infinite bit number \( B_1 \) and \( B_2 \) of feedback ) is also provided in this figure. Fig. 3 compares the BER performance of different schemes under \( M = 4 \) and \( N = 2 \). Similar observations can be found from this figure as from Fig. 2. Finally, it is also necessary to remark that the proposed precoding strategy has the same computational complexity order as the conventional one.

V. CONCLUSION

This paper presents an improved MMSE-based relay precoding strategy for multiuser relay downlinks. The precoding is designed by minimizing the expected MMSE of the system over two-hop channel realizations conditioned on known quantized channel information. Computer simulations verify that the proposed scheme improves both sum rate and BER performance with comparable complexity.

APPENDIX A

CALCULATION OF AN EXPECTATION

Lemma 1: If \( \mathbf{V} \) is the quantized version of the column unitary matrix \( \mathbf{V} \), then we have
\[
\mathbb{E} \left[ \mathbf{V}^H \mathbf{V} \right] = \mathbb{E} \left[ \mathbf{1} - \delta_i \mathbf{v}_i \right] \mathbf{I}. 
\] (25)

Proof: We calculate the expectation of each \((i, j)\)th entry of the matrix. Denote \( \mathbf{v}_i \) as the \( i \)th column of \( \mathbf{V} \). Start with the decomposition in (10) as follows
\[
\mathbf{v}_i = \sqrt{1 - \delta_i} \mathbf{r}_i + \sqrt{\delta_i} \mathbf{e}_i \tag{26}
\]
where \( \delta_i = 1 - |\mathbf{v}_i^H \mathbf{v}_i| \) and \( \mathbf{e}_i \) is an isotropically distributed unit-norm vector perpendicular to \( \mathbf{v}_i \).
Consider the $i$th diagonal elements, it follows
\[ E[v_i^H \hat{v}_i] = E[\sqrt{1 - \delta_i}] = E[\sqrt{1 - \delta_i}]. \tag{27} \]
As for the $(i,j)$th element with $i \neq j$:
\[ E[v_j^H \hat{v}_i] = E[\sqrt{1 - \delta_i} v_j^H v_i + \sqrt{\delta_i} v_i^H e_i] = E[\sqrt{\delta_i} v_i^H e_i]. \tag{28} \]
Note that since $v_j$ and $e_i$ are uniformly distributed vectors and both are orthogonal to $v_i$, the new vector $-e_i$ follows the same distribution as $e_i$ and is also orthogonal to $v_i$. We have
\[ -E[v_j^H e_i] = E[v_j^H (-e_i)] = E[v_j^H e_i] \tag{29} \]
which implies $E[v_j^H \hat{v}_i] = 0$. By substituting this into (28), it yields $E[v_j^H \hat{v}_i] = 0$. Therefore, by combining the element-wise results, we complete the proof. 

APPENDIX B
SOLVING (23) BY USING KKT CONDITIONS
Start with the Lagrangian penalty function defined by
\[ L(F_1, \beta, \mu) = c_1 \beta^{-2} \| \hat{G} H F_1 \|_F^2 - 2a \sqrt{\rho_1 \beta^{-1}} \Re \{ \text{Tr} (\hat{G} H F_1) \} + c_3 \beta^{-2} + \mu (\text{Tr}(F_1 P_1^H) - P_2^H) \tag{30} \]
where $\mu$ is the lagrange multiplier. Then, the KKT conditions for (23) are as follows
\[ \text{Tr}(F_1 P_1^H) = P_2^H \tag{31a} \]
\[ \nabla_{F_1} L = c_1 \beta^{-2} \hat{G} G H F_1 - a \sqrt{\rho_1 \beta^{-1}} \hat{G} + \mu F_1 = 0 \tag{31b} \]
\[ \nabla_{\beta} L = -2c_1 \beta^{-3} \| \hat{G} H F_1 \|_F^2 + 2a \sqrt{\rho_1 \beta^{-1}} \Re \{ \text{Tr}(\hat{G} H F_1) \} - 2c_3 \beta^{-3} = 0. \tag{31c} \]
From (31b), we have
\[ F_1 = a c_1^{-1} \beta \sqrt{\rho_1} (\hat{G} G H + c_1^{-1} \beta^2 \mu I)^{-1} \hat{G}. \tag{32} \]
By substituting (32) into (31c), it yields
\[ \nabla_{\beta} L = a^2 \beta^{-2} c_1^{-1} \rho_1 \text{Tr} \left( \hat{G} G H \left( \hat{G} G H + c_1^{-1} \beta^2 \mu I \right)^{-1} \right) \times \left( I - \hat{G} G H \left( \hat{G} G H + c_1^{-1} \beta^2 \mu I \right)^{-1} \right) - 2c_3 \beta^{-3} \]
\[ = a^2 \beta c_1^{-2} \rho_1 \text{Tr} \left( \hat{G} G H \left( \hat{G} G H + c_1^{-1} \beta^2 \mu I \right)^{-1} \right) \times \left( \hat{G} G H + c_1^{-1} \beta^2 \mu I \right) \]
\[ - 2c_3 \beta^{-3} \times \left( \hat{G} G H + c_1^{-1} \beta^2 \mu I \right)^{-1} \tag{33} \]
\[ = 0 \]
where (33) uses the equality $I - AA^H (AA^H + \alpha I)^{-1} = \alpha (AA^H + \alpha I)^{-1}$ for any $\alpha \neq 0$. Thus, we have
\[ \mu = a^{-2} \beta^{-4} c_2 c_1 \rho_1 \text{Tr} \left( \hat{G} G H \left( \hat{G} G H + c_1^{-1} \beta^2 \mu I \right)^{-1} \left( \hat{G} G H + c_1^{-1} \beta^2 \mu I \right) \right)^{-1}. \tag{34} \]
In order to further simplify the result of $\mu$, we then substitute (32) into (31a) and obtain
\[ \text{Tr} \left( \hat{G} G H \left( \hat{G} G H + c_1^{-1} \beta^2 \mu I \right)^{-1} \left( \hat{G} G H + c_1^{-1} \beta^2 \mu I \right) \right)^{-1} \]
\[ = \mu = \beta^{-2} c_1^2 P_2^H. \tag{35} \]
Then, from (35) and (34), the final value of $\mu$ is given by
\[ \mu = \beta^{-2} c_1^2 P_2^H. \tag{36} \]
Using this result of $\mu$, the solution to $F_1$ in (32) can be rewritten as
\[ F_1 = a c_1^{-1} \beta \sqrt{\rho_1} (\hat{G} G H + c_1^{-1} \beta^2 \mu I)^{-1} \hat{G}. \tag{37} \]
Thus far, by considering the transmit power constraint (31a) in the KKT conditions and then substituting $p_1$, $c_1$, $P_2$, and $c_3$ into the solution, the optimal $F_1$ to the problem in (23) can be finally expressed by
\[ F_1^* = \gamma \left( \hat{G} G H + \left( 1 + \tau_2 P_2 \right) \left( P_2 (1 - \tau_2) \right)^{-1} \hat{G} \right)^{-1} \tag{38} \]
where $\gamma$ is a scaling factor determined by
\[ \gamma = \frac{P_2^H}{\left\| \left( \hat{G} G H + \left( 1 + \tau_2 P_2 \right) \left( P_2 (1 - \tau_2) \right)^{-1} \hat{G} \right)^{-1} \right\|_F^2}. \tag{39} \]

REFERENCES