Abstract—In this paper, the source and relay transmit covariance matrices are jointly optimized for a fading multiple antenna relay channel when the transmitters only have partial channel state information (CSI) in the form of covariance feedback. For full-duplex transmission, we evaluate lower and upper bounds on the ergodic channel capacity. These bounds require a joint optimization over the source and relay transmit covariance matrices. The methods utilized in the previous literature fail to find fast and efficient algorithms for the system model considered in this paper. Therefore, we utilize matrix differential calculus in order to solve the joint optimization problem. In this method, there is no need to specify the eigenvectors of the transmit covariance matrices first. Through simulations, we observe that lower and upper bounds are close to each other.

Index Terms—MIMO relay channel, upper and lower bounds, partial CSI, covariance feedback, optimum power allocation, full-duplex.

I. INTRODUCTION

Utilizing multiple antennas at the transmit and receive terminals of wireless communication systems has been shown to increase the spectral efficiency [1]. In addition, applying cooperative strategies such as adding a relay node to the system can further increase the capacity [2]. On the other hand, exact description of multi-input multi-output (MIMO) relay channel capacity is still an open problem. Several achievable schemes, such as Decode and Forward (DF), Amplify and Forward (AF), and Compress and Forward (CF), can be used as lower bounds to the capacity, while the cut-set theorem provides a valid upper bound.

For single antenna fading relay channels, capacity bounds and power allocations are given in [3] for both full-duplex and half-duplex transmissions, where perfect CSI is available everywhere. A similar setting with individual power constraints at the source and relay is considered in [4], where a max-min type of solution is also introduced. In [5], MIMO relay channels with different fading assumptions are discussed, when only the receivers have the perfect CSI. For full-duplex, fading MIMO relay channels, capacity upper bound and DF achievable rate are found in [6], where only the receivers know the perfect CSI and transmitters do not know the channel.

A more practical channel model, for which the receivers have the perfect CSI and the transmitters have partial CSI, was utilized for point-to-point MIMO and MIMO multiple access channels (MAC) in [7], [8]. In both of these channels, it is possible to find the eigenvectors of the transmit covariance matrices in closed form, and solve a reduced optimization problem over the eigenvalues of the transmit covariance matrices, using an iterative algorithm [8]. However, in relay channels, it is not always possible to find the optimum closed form expression for the eigenvectors of the transmit covariance matrices. In [9], in an attempt to solve the power allocation problem, we picked some sub-optimum eigenvectors, and proposed an algorithm for the eigenvalues only. However, that algorithm is far from achieving the lower bound. Here, we propose a new method for solving the transmit covariance matrices directly (i.e., without the need of finding the eigenvectors first).

In this method, matrix differential calculus [10] is utilized, since it offers a new way for optimizing matrix valued functions by taking derivatives of scalar functions with respect to matrix variables. This eliminates the need for calculating cumbersome partial differentials. At the end, the iterative algorithm updates the entire matrix at once, at each iteration.

In this paper, we consider full-duplex MIMO relay channels where the transmitters have partial CSI in the form of covariance feedback. The source and relay terminals have individual power constraints. We evaluate DF lower bound and cut-set upper bound on the channel capacity that are given in terms of max-min type optimization problems over the source and relay transmit covariance matrices. We solve these joint optimization problems using techniques from [4] and also using matrix differential calculus [10]. The solutions to the optimization problems are in terms of iterative algorithms that find the transmit covariance matrices directly, i.e., without first finding the eigenvectors and then calculating the eigenvalues.

II. SYSTEM MODEL

Consider a relay channel, with source, relay and destination terminals, where the channel between a transmitter and a receiver is represented by a random matrix $H_{xy}$ whose dimensions are the number of receive antennas times the number of transmit antennas. In the case that the receiver has the perfect CSI and the transmitter has only statistical knowledge of channel in terms of covariance feedback, there is a correlation between the signals transmitted by or received at different antenna elements. This channel model is defined as [11]

$$H_{xy} = Z_{xy} \Sigma_{xy}^{1/2}$$

(1)

where subscript $xy$ refers to either $sr$ (source to relay), $sd$ (source to destination), or $rd$ (relay to destination); $Z_{xy}$ is a zero-mean identity covariance random channel matrix, $\Sigma_{xy}$ is
the correlation matrix between the signals transmitted from the antennas on the transmitter.

When the relay is allowed to transmit and receive at the same time, the channel is said to be in full-duplex mode. In this case, received signals at the relay and destination are given as

\[ r = H_{rs}x_s + n_r, \quad y = H_{sd}x_s + H_{rd}x_r + n_y \]  \hspace{1cm} (2)

where \( r \) is \( N_r \) long received vector at the relay, \( y \) is \( N_d \) long received vector at the destination, \( x_s \) is an \( M_s \) long transmitted signal from the source and \( x_r \) is an \( M_r \) long transmitted signal from the relay. The covariance matrices of the transmitted signals are \( Q_s = E[x_s x_s^H] \) and \( Q_r = E[x_r x_r^H] \), and there are individual power constraints on the source and relay transmit covariance matrices. Noise vectors at the relay, \( n_r \), and at the destination, \( n_y \), are zero-mean, identity covariance complex Gaussian random vectors.

III. MATRIX DIFFERENTIAL CALCULUS

In this section, we introduce matrix differential calculus [10] that will be useful in later sections. We start by defining the “differential” of a scalar function. Let \( \phi : \mathbb{R} \to \mathbb{R} \) be a real-valued function. The differential is the linear part of the increment of the value of a function, \( \phi(x+u) - \phi(x) \), at a fixed point \( x \) with an increment \( u \). The derivative of the function \( \phi \) at the point \( x \) is found by dividing the differential of the function with the increment \( u \), and by taking the limit as \( u \) goes to 0.

\[ \phi'(x) = \lim_{u \to 0} \frac{\phi(x+u) - \phi(x)}{u}. \]

The differential is denoted by \( d\phi(x; u) \) and it is equal to \( \phi'(x)u \). Similarly, let \( f : \mathbb{R}^n \to \mathbb{R}^m \) be a vector valued function, and \( x, u \in \mathbb{R}^n \). The differential of \( f \) is defined as \( df(x; u) = A(x)u \), where \( m \times n \) dimensional matrix \( A(x) \) is called the first derivative of \( f \) at \( x \). It is important to note here that while the differential of a vector valued function is a vector, derivative of a vector valued function is a matrix. Since dealing with a matrix is cumbersome, partial derivatives are often used in optimization problems involving vector valued functions. In fact, as the first identification theorem in [10] states, the elements of \( m \times n \) matrix \( A(x) \) are the partial derivatives of \( f \) evaluated at \( x \), and \( A(x) \) is called the Jacobian matrix of \( f \). \( DF(x) = A(x) \). As a result of this, if \( f \) is differentiable at \( x \) and \( u \), we have found a differential \( df \) at \( x \), then the value of the partial derivatives at \( x \) can be immediately determined.

Finally, the differential of a matrix valued function can be determined using the vector representation of matrices. Let \( F : \mathbb{R}^{n \times d} \to \mathbb{R}^{m \times p} \) be a matrix function, and differentiable at \( X \in \mathbb{R}^{n \times q} \). Then the differential can be written as \( vec dF(X; U) = A(X)vec X \), where the Jacobian is an \( mp \times nq \) matrix \( DF(X) = A(X) \).

Given a matrix function \( F(X) \), determining the derivative of this function from its differential is carried out as follows: (i) compute the differential of \( F(X) \), (ii) vectorize to obtain

\[ d vec F(X) = A(X)d vec X, \]

and (iii) conclude that \( DF(X) = A(X) \). In this paper, we mainly deal with scalar functions, \( \phi : \mathbb{R}^{n \times q} \to \mathbb{R} \), of matrix variables. In this case, the differential can be written as \( D\phi(X) = \frac{\partial\phi(X)}{\partial vec X} \). However, the idea of arranging the partial derivatives of \( \phi(X) \) into a matrix (rather than a vector) is appealing and sometimes useful, so with a slight abuse of notation we use \( D\phi(X) = \frac{\partial\phi(X)}{\partial X} \). For scalar functions of matrix variable, the differential of \( \phi(X) \) is given as \( \frac{\partial\phi(X)}{\partial X} = (vec A)^T d vec X \), which is also equal to \( d\phi = tr(A^T dX) \), where Jacobian matrix is given as \( D\phi(X) = \frac{\partial\phi(X)}{\partial X} = A \).

Next, we give some important differentials that will be useful later. Differential with respect to \( X \) of the trace of a matrix, \( XX^T \), can be calculated as \( d tr(XX^T) = 2tr(XdX) \). Therefore, the derivative of \( XX^T \) is \( DXX^T = 2X \).

Given a matrix \( H \), the differential with respect to \( X \) of the expression \( log[I + HXX^T ] \) can be calculated as \( d log[I + HXX^T] = tr(2X^T H(I - HXX^T)^{-1}H) \). Therefore, the derivative of the expression is \( D log[I + HXX^T] = 2H^T (I + HXX^T)^{-1}H \).

IV. CAPACITY BOUNDS IN FULL-DUPLEX RELAYING

Although the capacity is not known for general fading MIMO relay channels, capacity bounds can be derived using [2]. Similar bounds are derived for fading MIMO systems before [5], [6], however, our main contribution in this section is solving the max-min optimization problems for full-duplex transmission, in order to evaluate the DF achievable rate and the cut-set upper bound. Solutions to these optimization problems are not trivial when the transmitters have the covariance information of the channel, which is assumed in this paper.

A. Lower Bound on the Capacity

The DF achievable rate can be derived from the mutual information expressions in [2]. These expressions are evaluated for MIMO relay systems in [6] for the case where the transmitters do not have any information about the channel, and the source and relay share the same power constraint. In that case, lower bound maximizing transmit covariance matrices are identity matrices, and the cross-correlation matrices are zero [5]. In this section, we evaluate the DF achievable rate when the transmitters have the covariance information of the channel, and the source and relay have individual power constraints as it is assumed for single antenna systems in [4]. This rate is in terms of a max-min type optimization problem over the source and relay transmit covariance matrices. Next, we solve this optimization problem and propose an iterative algorithm that gives the transmit covariance matrices.

**Theorem 1:** When there is only channel covariance information at the transmitters and perfect CSI at the receivers, DF achievable rate of a full-duplex MIMO relay channel is given as

\[ C_{fd} \geq \max_{tr(Q_s) \leq P_s, tr(Q_r) \leq P_r} \min(I_{mac}, I_{sr}) \]  \hspace{1cm} (3)

where
\begin{align}
I_{mac} &= E \left[ \log \left| I + H_{sd}Q_sH_{sd}^\dagger + H_{rd}Q_rH_{rd}^\dagger \right| \right] \\
I_{sr} &= E \left[ \log \left| I + H_{sr}Q_sH_{sr}^\dagger \right| \right]. 
\end{align}

The key point of the proof is the independence of the source and relay input signals. This is concluded due to [5], where it is stated that the cross-correlation matrices that maximize the mutual information values, $Q_{sr} = E[x_s x_r]$ and $Q_{sr} = E[x_s x_r]$, are zero when the transmitters do not know more than the statistics of the channel.

Theorem 1 gives the DF achievable rate in terms of a max-min type optimization problem that is still needs to be solved. The solution to this problem requires a joint optimization over the source and relay transmit covariance matrices. Because, the optimum $Q_s$, that maximizes $I_{mac}$ in (4) and the optimum $Q_r$ that maximizes $I_{sr}$ in (5) are different. If we maximize $I_{mac}$, that choice of $Q_s$ will result in a low $I_{sr}$. As a result, $I_{sr}$ will come out of the minimization in (3), and the achievable rate will attain a lower value. In order to solve this trade-off, $Q_s$ and $Q_r$ should be found jointly. First, we utilize a method that is proposed in [4]. The following function $R_{lb} \alpha$ and $Q$ is defined as

$$R_{lb}(\alpha, Q) = \alpha I_{mac}(Q) + (1 - \alpha)I_{sr}(Q), \quad 0 \leq \alpha \leq 1$$

where $Q = [Q_s, Q_r]$. The max-min problem in (3) corresponds to first maximizing $R_{lb}(\alpha, Q)$ over $Q$ for a fixed $\alpha$, and then taking the minimum over $\alpha$ [4]. Let us define $V_{lb}(\alpha)$ as $V_{lb}(\alpha) = \max_Q R(\alpha, Q)$ and suppose that $\alpha^*$ provides the minimum value of $V_{lb}(\alpha)$.

Determining on the value of $\alpha^*$, we have three cases. Optimum source and relay covariance matrices may be different in all three cases. In the first case ($\alpha^* = 0$), $R_{lb}(0, Q) = I_{sr}(Q)$ and the condition $I_{mac}(Q) \geq I_{sr}(Q)$ should be satisfied [4]. Since the achievable rate is found by maximizing $I_{sr}(Q)$ only, we find the source transmit covariance matrix, $Q_s$, as a solution to the point-to-point problem from source to relay. When the receiver knows perfect CSI and the transmitter knows partial CSI, point-to-point problem is already solved in [8]. Then, we find the relay transmit covariance matrix, $Q_r$, by maximizing $I_{mac}(Q)$ with a fixed $Q_s$. This is also equivalent to a single user problem, which is solved in [8].

In the second case, $\alpha^* = 1$, $R_{lb}(1, Q) = I_{mac}(Q)$ and the condition $I_{mac}(Q) \leq I_{sr}(Q)$ should be satisfied. In this case, the achievable rate is found by maximizing $I_{mac}(Q)$, which is a MAC problem that is already solved in [8].

In the third case, $0 < \alpha^* < 1$, $R_{lb}(\alpha^*, Q) = \alpha^* I_{mac}(Q) + (1 - \alpha^*)I_{sr}(Q)$ and the condition $I_{mac}(Q) = I_{sr}(Q)$ should be satisfied. In this case, we find the transmit covariance matrices of the source and relay as functions of $\alpha^*$. The third case is the most interesting case as the solution is not trivial. In that case, $Q_s$ and $Q_r$ must be optimized jointly since objective function $R_{lb}(\alpha, Q)$ includes both $I_{sr}$ and $I_{mac}$. However, this joint optimization problem cannot be solved by using the methods from the previous literature. In studies like [7], [8], the transmit covariance matrices are always found by determining their eigenvectors first. This reduces the problem of finding the eigenvalues of the transmit covariance matrix, from a matrix variable to a vector (and sometimes scalar) variable problem. Since the eigenvectors cannot be determined in closed form in this joint optimization, we will use matrix differential calculus.

First, (6) will be maximized over $Q$ for a fixed $\alpha^* > 0 < \alpha^* < 1$.

$$V_{lb}(\alpha^*) = \max_{\text{tr}(Q_s) \leq P_s, \text{tr}(Q_r) \leq P_r} \alpha^* I_{mac}(Q) + (1 - \alpha^*)I_{sr}(Q) \quad (7)$$

Note that, transmit covariance matrices that will result from this optimization will depend on $\alpha^*$. The Lagrangian of (7) can be written as

$$L = R_{lb}(\alpha^*, Q) - \mu_s(\text{tr}(Q_s) - P_s) - \mu_r(\text{tr}(Q_r) - P_r) \quad (8)$$

where $\mu_s$ and $\mu_r$ are Lagrange multipliers corresponding to source and relay power constraints, respectively. Here, we will take the derivative of (8) with respect to $Q_s$ and $Q_r$ directly. However, in order to enforce their positive semi-definiteness, we rewrite them as $Q_s = A A^\dagger$ and $Q_r = B B^\dagger$. Using matrix differential calculus by referring to the examples in Sects. 2, one can take the derivative of (8) with respect to $A$ and $B$ to obtain the following KKT conditions

$$E \left[ \alpha^* H_{sd}^\dagger D_{mac} H_{sd} A + (1 - \alpha^*) H_{sr}^\dagger D_{sr} H_{sr} B \right] \leq \mu_s A \quad (9)$$
$$E \left[ \alpha^* H_{rd}^\dagger D_{mac} H_{rd} B \right] \leq \mu_r B \quad (10)$$

where $D_{mac}$ is the expression inside the determinant in (4) and $D_{sr}$ is the expression inside the determinant in (5). Let us denote the left hand side of (9) as $E_1$ and the left hand side of (10) as $E_2$. Note that when we multiply both sides of (9) with $A^\dagger$ from the right hand side and both sides of (10) with $B^\dagger$ from the right hand side, the inequalities become equalities. This can be seen as an extension to the reasoning in [8]. Consider spectrally decomposing both sides of (9), the inequalities corresponding to non-zero singular values of $A$ are satisfied with equalities due to complementary slackness conditions. The inequalities corresponding to zero singular values of $A$ were originally strict inequalities. However, after multiplying both sides with $A^\dagger$, both sides of those strict inequalities are multiplied with zero, and therefore those strict inequalities become equalities with both sides being zero. The same reasoning goes for (10) as well. Finally, by applying the trace operator, Lagrange multipliers are calculated as

$$\mu_s = \frac{\text{tr}(E_1 A^\dagger)}{P_s}, \quad \mu_r = \frac{\text{tr}(E_2 B^\dagger)}{P_r}. \quad (11)$$

Using the idea in [8], we propose the following iterative algorithm to solve for the fixed point equations that are obtained from (9)-(10)

$$Q_s(n+1) = \frac{E_1(n)A(n)P_s}{\text{tr}(E_1(n)A(n)^\dagger)}, \quad Q_r(n+1) = \frac{E_2(n)B(n)^\dagger P_r}{\text{tr}(E_2(n)B(n)^\dagger)} \quad (12)$$
This iterative algorithm finds the optimum transmit covariance matrices of the source and relay for Case 3. After running this algorithm for different $\alpha$ values, a minimization over $\alpha$ is performed in order to find the lower bound. It is important to note that the algorithm in (12) updates every element of the transmit covariance matrices at once. As mentioned before, the eigenvectors of the transmit covariance matrices were not determined beforehand, they are found implicitly after the algorithm in (12) converges.

B. Upper Bound on the Capacity

Having derived the DF achievable rate and jointly optimized the source and transmitter covariance matrices, in this section we consider the cut-set upper bound. This bound is introduced in [2] and evaluated for different channel model assumptions in the literature. For example, when the receivers have perfect CSI and the transmitters have no CSI, cut-set cut-set upper bound on MIMO relay channel capacity is found in [6]. In this paper, we consider a case where there is transmit covariance information at the transmitters. In this case, similar to the lower-bound development, we first evaluate the mutual information expressions in the cut-set bound, and then optimize the upper bound over $Q_s$ and $Q_r$.

**Theorem 2:** When there is only channel covariance information at the transmitters and perfect CSI at the receivers, cut-set upper bound of a full-duplex MIMO relay channel is given as

$$C_{fd} \leq \max_{\text{tr}(Q_s) \leq P_s, \text{tr}(Q_r) \leq P_r} \min(I_{mac}, I_{bc})$$

where $I_{bc} = E \left[ \log \left| I + H_{bc} Q_r H_{bc}^\dagger \right| \right]$, $I_{mac}$ is given in (4), and $H_{bc} = \left[ H_{sd} \, H_{sr} \right]^\dagger$.

Note that the DF achievable rate and the cut-set upper bound expressions both involve the same $I_{mac}$. Therefore, lower and upper bounds meet and provide the capacity if $I_{mac}$ comes out of the minimization in both cases. As the case of the lower bound, we have a max-min problem to solve in the upper bound as well. The method for this solution is similar to the lower bound solution and utilizes matrix differential calculus. We will skip some of the development where they can easily be obtained from lower-bound analysis. This time, we define $R_{ub}$ as

$$R_{ub}(\alpha, Q) = \alpha I_{mac}(Q) + (1 - \alpha) I_{bc}(Q), \quad 0 \leq \alpha \leq 1$$

Note that unlike the DF achievable rate, the upper bound, $R_{ub}$ depends on $I_{bc}$, not on $I_{sr}$. Depending on the value of minimum $\alpha^*$, the solution again has three cases. In the first case ($\alpha^* = 0$), $R(0, Q) = I_{bc}(Q)$ and the condition $I_{mac}(Q) \geq I_{bc}(Q)$ should be satisfied. For this case, the Lagrangian can be written as

$$L = I_{bc}(Q) - \mu_s (\text{tr}(A A^\dagger) - P_s)$$

Using matrix differential calculus, and by taking the derivative of (15) with respect to $A$, we obtain the KKT conditions. Then, similar to the lower bound we derive the following algorithm

$$Q_s(n + 1) = \frac{E_3(n) A(n)^\dagger}{\text{tr}(E_3(n) A(n)^\dagger)} P_s$$

where $E_3 = E \left[ H_{bc} D_{bc}^{-1} H_{bc} A \right]$, and $D_{bc}$ is the matrix inside the determinant of $I_{bc}$. Next, $Q_r$ is found by maximizing $I_{mac}$ using fixed $Q_s$. This is equivalent to a single user problem that is solved in [8].

The second case is again a MIMO-MAC channel and is already known. In the third case ($\alpha < \alpha^* < 1$), $R_{ub}(\alpha^*, Q) = \alpha^* I_{mac}(Q) + (1 - \alpha^*) I_{bc}(Q)$ and the condition $I_{mac}(Q) = I_{bc}(Q)$ should be satisfied. The Lagrangian for this case as given as

$$L = R_{ub}(\alpha^*, Q) - \mu_s (\text{tr}(A A^\dagger) - P_s) - \mu_r (\text{tr}(B B^\dagger) - P_r)$$

Using matrix differential calculus and by taking the derivative of (17) with respect to $A$ and $B$, we obtain the KKT conditions. Then, using the similar method as in the lower bound we derive the algorithm below

$$Q_s(n + 1) = \frac{E_3(n) A(n)^\dagger}{\text{tr}(E_3(n) A(n)^\dagger)} P_s \quad Q_r(n + 1) = \frac{E_2(n) B(n)^\dagger}{\text{tr}(E_2(n) B(n)^\dagger)} P_r$$

where $E_4 = E \left[ A^\dagger H_{sd} D_{mac}^{-1} H_{sd} A + (1 - \alpha^*) H_{bc} D_{bc}^{-1} H_{bc} A \right]$.

This iterative algorithm finds the transmit covariance matrices of the source and relay nodes that solves the Case 3 of the optimization problem in the upper bound. Finally, a minimization over $\alpha$ is performed in order to find which case results in the upper bound.

V. Numerical Results

We start with a convergence analysis. For all calculations, the power constraints ($P_s$ and $P_r$) are fixed at 10 dB. At each iteration, we calculate the matrix norms of transmit covariance matrices of the source and relay terminals. Then, in Figure 1, we plot the difference in matrix norms between successive iterations. We clearly see that as the iteration index increases, covariance matrices converge to their optimum values.

Second, capacity bounds on full-duplex MIMO relay channel are simulated using the proposed algorithms. Power constraints are chosen to be 10 dB for all cases. Figures 2 and 3 give those bounds for different channel covariance matrices. For the covariance matrix corresponding to Figure 2, lower and upper bounds are given by $\alpha^* = 1$ point (Case 2), which is the minimum value of the curve with respect to $\alpha$. As expected, the lower bound is equal to upper bound at Case 2, and the capacity is in fact achieved for this covariance matrix setting. Similarly, for the covariance matrix corresponding to Figure 3, lower and upper bounds are given by $\alpha^* = 0.9$ point (Case 3), which is the minimum value of the curves with respect to $\alpha$. The difference between the lower and the upper bounds for this case is about 1%. Maximum difference between the bounds happens in Case 1, the point of $\alpha = 0$. At that point, the difference between the rates is 10%. 

\[ \text{Raw text content} \]
Next, with the source power is fixed at 10 dB, we simulate the lower bound algorithm by changing the relay power. In Figure 4, we observe that the channel is subject to Case 2 condition when the relay power is 5-10 dB, to Case 3 condition when the relay power is 10-15 dB, and to Case 1 condition when the relay power is 15-30 dB. The channel saturates with relay power since in Case 1 the relay power is large enough to forward all the information decoded at the relay node to the destination node, and the achievable rate is limited by the capacity of the source to relay link [4].

VI. CONCLUSION

In this paper, we analyzed both full-duplex fading MIMO relay channels when the transmitters have partial CSI and the receivers have the perfect CSI. The channel capacity for such a system is not known in general. We derived DF achievable rates and cut-set upper bounds on the channel capacity which were given in terms of max-min type optimization problems. When the transmitters know the channel covariance information, finding optimum source and relay transmit covariance matrices become important. Because, power allocation over the spatial dimension of the channel has a significant impact on the performance. We use matrix differential calculus to solve the source and relay transmit covariance matrices jointly. In our method, optimum transmit covariance matrices are found directly using a fast and efficient iterative algorithm.

REFERENCES