Duality and Optimization for Generalized Multi-hop MIMO Amplify-and-Forward Relay Networks with Linear Constraints

An Liu\(^1\), Vincent K. N. Lau\(^1\), Youjian Liu\(^2\)

\(^1\)Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology

\(^2\)Department of Electrical and Computer Engineering, University of Colorado at Boulder

Abstract

We consider a generalized multi-hop MIMO amplify-and-forward (AF) relay network with multiple sources/destinations and arbitrarily number of relays. We establish two dualities and the corresponding dual transformations between such a network and its dual, respectively under single network linear constraint and per-hop linear constraint. The result is a generalization of the previous dualities under different special cases and is proved using new techniques which reveal more insight on the duality structure that can be exploited to optimize MIMO precoders. A unified optimization framework is proposed to find a stationary point for an important class of non-convex optimization problems of AF relay networks based on a local Lagrange dual method, where the primal algorithm only finds a stationary point for the inner loop problem of maximizing the Lagrangian w.r.t. the primal variables. The input covariance matrices are shown to satisfy a polite water-filling structure at a stationary point of the inner loop problem. The duality and polite water-filling are exploited to design fast primal algorithms. Compared to the existing algorithms, the proposed optimization framework with duality-based primal algorithms can be used to solve more general problems with lower computation cost.

Index Terms

Multi-hop MIMO Networks, Amplify and Forward, Relay, Duality, MIMO Precoder Optimization

The work is funded by Huawei Technologies.
I. INTRODUCTION

The amplify and forward (AF) relay technique is useful in wireless systems for cost-effective throughput enhancement and coverage extension. It has attracted many research works recently. The optimization of the MIMO (Multiple-input multiple-output) AF relay system has been well studied (see the tutorial paper [1] and the reference there in). The optimal relay precoding matrix was established in [2] for a two-hop MIMO relay system and the result was later extended to multi-hop MIMO relay system in [3], where the optimal source and relay precoding matrices are shown to diagonalize the source-relay-destination channel. It was shown in [4] that this diagonalization property also holds for power minimization of MIMO AF relay system under QoS constraints. However, the power minimization problem is still difficult to solve due to its non-convex nature [5]. A low complexity algorithm is proposed in [5] to closely approximate the solution of this non-convex problem for the two-hop case. The source and relay power allocation problem of the multicarrier two-hop AF relay system was studied in [6]. All of the above works focus on the case with a single relay at each hop. In [7], the power minimization under QoS constraints was addressed for a two-hop MIMO relay system with multiple relays. There are also some works on the optimization of multiuser MIMO AF relay systems. For example, the multiuser MIMO MAC (multiaccess channel) AF relay networks are addressed in [8], [9], while the MIMO BC (broadcast channel) AF relay networks are investigated in [9]. Since the optimal structure of the source and relay precoding matrices is still unknown for multi-user MIMO AF relay networks, most of the algorithms in [8], [9] are based on standard optimization methods such as Geometric programming [8] and logarithmic barrier method [10]. For MAC/BC AF relay networks with single-antenna source/destination nodes, two dualities with respect to a total network power constraint and per-hop power constraint have been established in [11]. The results were later generalized to two-hop MIMO MAC/BC AF relay networks [12] and multi-hop MIMO AF relay systems [13]. In some cases, the duality can be used to simplify the network optimization problem, e.g., some applications of the duality can be found in [13], [14].

In this paper, we consider a generalized multi-hop MIMO AF relay network called the B-MAC AF relay network, which has multiple sources, multiple destinations and arbitrarily number of relays. The B-MAC AF relay network topology includes almost all the existing AF relay networks as special cases and the optimization of such general network has not been addressed in the literature. We derive structural properties of the optimization problem (duality and polite water-filling (PWF)) and exploit the duality and PWF structure to design efficient source and relay precoding algorithms for general B-MAC AF relay networks. For the special case of the single-hop B-MAC interference network (B-MAC IFN), the duality
and PWF results have been established in [15], [16]. However, due to the key difference in topology (single-hop versus multi-hop), there are various technical challenges and we cannot extend the existing results and algorithms trivially from [15], [16]. Specifically, due to the multi-hop topology, the achievable rate is a complicated function of the input covariance and relay precoding matrices. This makes both the proof of dualities and the optimization of B-MAC AF relay networks much more challenging than the previous works with simpler network topology. Furthermore, the optimization constraints in the multi-hop B-MAC AF relay network is more complicated than that in B-MAC IFN. For example, the per-hop power constraint and per-relay power constraint are two practically important constraints that only occur in the multi-hop network topology.

We propose several new techniques to solve these challenges. The main contributions are listed below.

- **Structural Properties (Duality) of MIMO AF Relay Network Optimization**: We show that the achievable regions of a B-MAC AF relay network and its dual are the same under *single network linear constraint* (Type I duality) or *per-hop linear constraint* (Type II duality). We also derive explicit *dual transformations* to calculate the dual input covariance and relay precoding matrices. It is difficult to use the brute force computation approach in [11]–[13] to prove the duality for general B-MAC AF relay networks. We propose two new techniques, namely the *network equivalence* and the network dual scaling technique, to prove the Type I and Type II dualities. The proof is not only simpler but also reveals more insight on the duality structure. The duality and the network equivalence results are useful for MIMO precoder optimization in B-MAC AF relay networks. First, they are used in Section IV-B to derive the PWF structure of the optimal input covariance matrices, which can be exploited to design efficient input covariance optimization algorithm as illustrated in Section IV-C2. Second, the duality is used in Section IV-C3 and IV-D to simplify the optimization of relay precoding matrices in BC AF relay networks.

- **Efficient MIMO Precoder Optimization Algorithm for AF Relay Networks**: Based on the duality structure established above, we propose efficient algorithms for MIMO precoder optimization in general MIMO AF relay networks. Since the duality is only established for single network / per-hop linear constraint, we cannot directly exploit the duality structure to handle general power constraints such as per-relay power constraint. To solve this problem, we propose a unified optimization framework to find a stationary point for an important class of B-MAC AF relay network optimization problems. In such a framework, we only need to design a primal algorithm to find a stationary point for the unconstrained inner loop problem of maximizing the Lagrangian. Using the network
equivalence and duality results, we then show that at a stationary point of the inner loop problem, the input covariances have the PWF structure, which is a generalization of the PWF structure in B-MAC IFN \cite{16} to the B-MAC AF Relay network. Finally, we propose several efficient primal algorithms based on duality and PWF structure for two-hop, three-hop, and multi-hop AF relay networks. The proposed algorithms have significant advantages w.r.t. conventional step-size based iterative algorithms (such as the gradient ascent method \cite{17} and the logarithmic barrier algorithms in \cite{9}) in terms of complexity and convergence speed. It can also be used to solve a more general class of problems than the existing AF relay network optimization algorithms in \cite{1}, \cite{8}, \cite{9}.

The rest of the paper is organized as follows. Section II describes the system model, the achievable scheme and the preliminary results on rate duality for B-MAC IFN. Section III presents the main results on the two types of dualities for multi-hop B-MAC AF relay networks. In Section IV, the duality is applied to study the non-convex optimization of B-MAC AF Relay networks. A unified optimization framework together with several primal algorithms are proposed to find a stationary point for an important class of B-MAC AF relay network optimization problems. The performance of the proposed algorithms is verified by simulation in Section V. Section VI concludes.

II. SYSTEM MODEL AND PRELIMINARIES

A. B-MAC AF Relay Network

A B-MAC AF relay network is a general multi-hop AF relay network with multiple sources (transmitters), multiple destinations (receivers) and $Q$ relay clusters as illustrated in Fig. I. Each relay cluster can have multiple relays. Each node can have multiple antennas. We follow the widely used orthogonal assumption \cite{9}, \cite{11}–\cite{13}:

Assumption 1 (Orthogonality among different hops). Due to propagation loss and proper channel reuse, the node in the $q^{th}$ cluster only receive the signals from the $(q-1)^{th}$ cluster, where the sources and destinations are considered as the $0^{th}$ and $(Q+1)^{th}$ cluster respectively.

We consider $L$ data links between the sources and destinations. Each source can have independent data for different destinations, and each destination may want independent data from different sources. Assume the sources are labeled by elements in $\mathcal{T} = \{Tx_1, Tx_2, Tx_3, \ldots\}$ and the destinations are labeled by elements in $\mathcal{R} = \{Rx_1, Rx_2, Rx_3, \ldots\}$. Define $T_l$ ($R_l$) as the label of the source (destination) of the $l^{th}$ data link. The numbers of antennas at $T_l$ and $R_l$ are $L_{T_l}$ and $L_{R_l}$ respectively. The total number of antennas at the $q^{th}$ relay cluster is $L_q$. An example of B-MAC AF relay network is illustrated in Fig.
There are 2 sources, 2 destinations, and 3 data links, of which $T_1 = Tx_1$, $T_2 = Tx_1$, $T_3 = Tx_2$, $R_1 = Rx_1$, $R_2 = Rx_2$, and $R_3 = Rx_2$. The figure also illustrates how to properly allocate sub-channels to achieve the orthogonality among different hops. In Fig. 1, we assume that due to path loss, the aggregate interference from the nodes which are 2 or more hops away from the receiver is negligible. Hence we can avoid interference among the transmissions of different hops by dividing the whole bandwidth into 3 orthogonal sub-channels and allocating different sub-channels to adjacent clusters.

For convenience, we use link based notation for signals. Denote $H_0 \in \mathbb{C}^{L_1 \times L_{T_1}}$ the channel matrix between source $T_1$ and the first relay cluster, $H_q \in \mathbb{C}^{L_q+1 \times L_q}$, $1 \leq q \leq Q$ the overall channel matrix between the $q$th and $(q+1)$th relay cluster, and $H_Q \in \mathbb{C}^{L_R \times L_Q}$ the channel matrix between the $Q$th relay cluster and destination $R_l$. Let $x_k \in \mathbb{C}^{L_{T_k} \times 1}$ denote the transmit signal of link $k$ and assume it is a circularly symmetric complex Gaussian (CSCG) vector with covariance $\Sigma_k$. Let $w_q \in \mathbb{C}^{L_q \times 1}$ denote the overall CSCG noise vector at the $q$th relay cluster with covariance $W_q$. Then the composite received signal at the first relay cluster is $y_{R_1} = \sum_{k=1}^{L_1} H_0^0 x_k + w_1$. The $j$th relay in the $q$th cluster applies linear transformation $F_q(j)$ to the received signal and the resulting composite transmit signal from the $q$th relay cluster is $x_{R_q} = F_q y_{R_q}$, where $F_q = \text{Block Diag} [F_q(1), \ldots, F_q(n_q)] \in \mathbb{C}^{L_q \times L_q}$ is the composite relay precoding matrix, $n_q$ is the number of relays in the $q$th relay cluster, and $y_{R_q} = H_{Q-1}^q x_{R_{Q-1}} + w_q$ is the composite received signal at the $q$th relay cluster. For convenience, define

$$B_{q,q'} = F_q' H_{q-1}^q \ldots H_q^q, \forall q' \geq q$$

$$\tilde{H}_{l,k} = H_{Q}^l B_{1,Q} H_0^k, \forall k,l.$$ (1)

Then using induction, it can be shown that the received signal at $R_l$ is

$$y_l = \sum_{k=1}^{L} \tilde{H}_{l,k} x_k + \sum_{q=1}^{Q} H_Q^l B_{q,Q} w_q + w_{Q+1}^l,$$ (3)

where $w_{Q+1}^l \in \mathbb{C}^{L_R \times 1}$ is the CSCG noise vector at $R_l$ with covariance $W_{Q+1}^l$. As an example, the equivalent channel model for the B-MAC AF relay network in Fig. 1 is illustrated in the lower sub-figure.

### B. Transmit/Receive Scheme and Achievable Rate

We consider a centralized optimization scheme in which a central node computes all the input covariance and relay precoding matrices and then transmits them to the source and relay nodes.

**Assumption 2 (Assumptions on the CSI).** The central node has the knowledge of the global channel state information (CSI) $H_0^l$'s $H_q$'s and $H_Q^l$'s.
Remark 1. In two-hop AF relay systems, both the source-relay and relay-destination channels can be estimated directly at the destination node by transmitting pilot signals [18]. However, acquiring global CSI in multi-user multi-hop AF relay networks is difficult and requires further study. In this paper, our main focus is to understand the first order optimization problem and solution properties in multi-user multi-hop AF networks. An interesting future work is to study efficient decentralized solution which does not require global CSI knowledge. The centralized solution in this paper provides the basis for studying the decentralized solutions.

For fixed relay precoding matrix $F_q$'s, the source nodes and destinations see an equivalent channel in
which is essentially a B-MAC IFN studied in [15], [19]. Therefore, we can adopt the same set of transmit/receive schemes as in [15], [19].

**Assumption 3** (Assumptions on the transmit/receive scheme).

1) The transmit signal is CSCG.

2) Each signal is decoded by no more than one destination and can be cancelled at this destination for the decoding of other signals.

3) The interference among the source-destination pairs can be specified by some binary coupling matrix $\Phi \in \{0, 1\}^{L \times L}$. That is, $\forall k, l$, after (possible) interference cancellation, the interference from $\mathbf{x}_k$ to $\mathbf{x}_l$ is given by $\phi_{l,k} \hat{H}_{l,k}\mathbf{x}_k$, where $\phi_{l,k}$ is the $(l,k)$th element of $\Phi$. Because $\phi_{l,k} \in \{0, 1\}$, each interference is either completely cancelled or treated as noise.

The set of allowed transmit/receive schemes includes most interference management techniques as special cases, such as spatial interference reduction through beamforming, dirty paper coding (DPC) [20] at sources and successive interference cancellation (SIC) at destinations [15]. For example, in the B-MAC AF relay network in Fig. 1, suppose DPC is used at $\text{Tx}_1$ with $\mathbf{x}_1$ encoded after $\mathbf{x}_2$, and SIC is used at $\text{Rx}_2$ with $\mathbf{x}_3$ decoded after $\mathbf{x}_2$. Then $\phi_{1,3}, \phi_{2,1}, \phi_{2,3}$ and $\phi_{3,1}$ are 1 and all other elements of $\Phi$ are 0.

For given input covariance matrices $\Sigma = (\Sigma_1, \Sigma_2, ..., \Sigma_L)$ and relay precoding matrices $\mathbf{F} = (\mathbf{F}_1, \mathbf{F}_2, ..., \mathbf{F}_Q)$, the covariance matrix of the transmit signal at the $q$th relay cluster is

$$
\Sigma^R_q = \sum_{l=1}^{L} \mathbf{B}_{1,q} \hat{H}_{l,0} \Sigma_l \hat{H}_{l,0}^\dagger \mathbf{B}_{1,q}^\dagger + \sum_{q'=1}^{q} \mathbf{B}_{q',q} \hat{W}_{q'} \mathbf{B}_{q',q}^\dagger.
$$

The interference-plus-noise covariance at $\text{R}_l$ is

$$
\Omega_l = \mathbf{W}_l' + \sum_{k=1}^{L} \phi_{l,k} \hat{H}_{l,k} \Sigma_k \hat{H}_{l,k}^\dagger,
$$

where $\phi_{l,l} = 0$ and

$$
\mathbf{W}_l' = \sum_{q=1}^{Q} \mathbf{H}_{q}^{\dagger} \mathbf{B}_{q,Q} \mathbf{W}_{q} \mathbf{B}_{q,Q}^{\dagger} \mathbf{H}_{q}^{\dagger} + \mathbf{W}_{Q+1}.
$$

The achievable rate of the $l$th data link is [21]

$$
\mathcal{I}_{l}^b (\Sigma, \mathbf{F}) = \log \left| \mathbf{I} + \hat{H}_{l,l} \Sigma_l \hat{H}_{l,l}^\dagger \Omega_l^{-1} \right|,
$$

on which the achievable regions in this paper are based.

We consider two types of linear constraints: Define the single network linear constraint as

$$
\sum_{l=1}^{L} \text{Tr} \left( \Sigma_l \hat{W}_l \right) + \sum_{q=1}^{Q} \text{Tr} \left( \Sigma^R_q \hat{W}_q \right) \leq P_T,
$$

(3)
where $\hat{W}_0^l$ and $\hat{W}_q$ are positive semidefinite; the per-hop linear constraint is defined as
\[
\sum_{l=1}^L \text{Tr} \left( \Sigma_l \hat{W}_0^l \right) \leq P_0, \quad \text{Tr} \left( \Sigma_q^R \hat{W}_q \right) \leq P_q, \quad 1 \leq q \leq Q. \tag{8}
\]

The above linear constraints are more general than the sum/individual power constraints considered in [11]–[13]. The per-hop linear constraint is more relevant to practical systems. However, single network linear constraint is also important in practice because it can be used to handle the more general multiple linear constraints using Lagrange multiplier method as will be shown in Section IV.

C. The Dual of B-MAC AF Relay Network

Definition 1 (Dual of a B-MAC AF relay network). The dual of a B-MAC AF relay network is obtained by the following operations. 1) Reverse the transmission directions, i.e., in the dual network, $R_l$'s become the sources, $T_l$'s become the destinations, and the transmit signal of $R_l$ travels through the $Q$th relay cluster, the $(Q-1)$th relay cluster, ..., the 1th relay cluster, and finally reaches the destination $T_l$. 2) Replace all the channel matrices by their conjugate transposes, i.e., the channel matrix between source $R_l$ and the $Q$th relay cluster is $H_{l,q}^\dagger$, the channel matrix between the $(q+1)$th relay cluster and the $q$th relay cluster is $H_q^\dagger$, $q = Q - 1, ..., 1$, and the channel matrix between the 1th relay cluster and destination $T_l$ is $H_0^\dagger$. 3) In the dual network, the covariance of the noise at the $q$th relay cluster is set as $\hat{W}_q$ and the covariance of the noise at the destination $T_l$ is set as $\hat{W}_0^l$, where $\hat{W}_q$'s and $\hat{W}_0^l$'s are the linear constraint matrices in (7) of the original network. 4) The coupling matrix for the dual network is set as the transpose of that for the original network, i.e., $\Phi^T$.

As an example, the dual of the B-MAC AF relay network in Fig. 1 is illustrated in Fig. 2.

We use the notation $\hat{\cdot}$ to denote the terms in the dual network. Given input covariance matrices $\hat{\Sigma} = (\hat{\Sigma}_1, \hat{\Sigma}_2, ..., \hat{\Sigma}_L)$ at the sources and relay precoding matrices $\hat{F} = (\hat{F}_1, \hat{F}_2, ..., \hat{F}_Q)$ at the relays, the covariance matrix of the transmit signal at the $q$th relay cluster is
\[
\hat{\Sigma}_q^R = \sum_{l=1}^L \hat{B}_{q,q}^\dagger H_Q^\dagger \hat{\Sigma}_l H_l^\dagger \hat{B}_{q,Q} + \sum_{q'=q}^Q \hat{B}_{q,q'}^\dagger \hat{W}_q \hat{B}_{q,q'}, \quad \forall q' \geq q. \tag{9}
\]

The interference-plus-noise covariance at $T_l$ is
\[
\hat{\Omega}_l = \hat{W}_0^l + \sum_{q=1}^Q H_0^\dagger \hat{B}_{1,q}^\dagger \hat{W}_q \hat{B}_{1,q} H_0^\dagger + \sum_{k=1}^L \phi_{k,l} \hat{H}_{l,k} \hat{\Sigma}_k \hat{H}_{l,k}^\dagger, \tag{10}
\]
where \( \mathbf{H}_{l,k} = \mathbf{H}_{0}^{l} \mathbf{B}_{l,Q}^{k} \mathbf{H}_{Q}^{k} \). Hence, the achievable rate of the \( l \)-th data link in the dual network is [21]

\[
\hat{I}_{l}^{\phi} (\mathbf{\Sigma}, \mathbf{\hat{F}}) = \log \left| \mathbf{I} + \mathbf{H}_{l,l} \mathbf{\hat{S}}_{l} \mathbf{H}_{l,l}^{\dagger} \mathbf{\hat{\Omega}}_{l}^{-1} \right|.
\] (11)

Similarly, in the dual network, the two types of linear constraints are defined as

\[
\sum_{l=1}^{L} \text{Tr} \left( \mathbf{\hat{S}}_{l} \mathbf{W}_{l}^{l} \right) + \sum_{q=1}^{Q} \text{Tr} \left( \mathbf{\hat{S}}_{q}^{R} \mathbf{W}_{q} \right) \leq P_{T},
\] (12)

\[
\sum_{l=1}^{L} \text{Tr} \left( \mathbf{\hat{S}}_{l} \mathbf{W}_{Q+1}^{l} \right) \leq P_{Q}, \text{Tr} \left( \mathbf{\hat{S}}_{q}^{R} \mathbf{W}_{q} \right) \leq P_{q-1}, \forall q.
\] (13)

D. Rate Duality for One-hop B-MAC IFN

The proof of the duality results in this paper relies on the rate duality for B-MAC IFN [15], which is a special case of the B-MAC AF relay network when \( Q = 0 \). Let

\[
\left( [\mathbf{H}_{l,k}], \sum_{l=1}^{L} \text{Tr} \left( \mathbf{\Sigma}_{l} \mathbf{W}_{l} \right) \leq P_{T}, [\mathbf{W}_{l}] \right),
\] (14)

denote a B-MAC IFN where the channel matrix between \( T_{k} \) and \( R_{l} \) is \( \mathbf{H}_{l,k} \); the input covariance satisfies \( \sum_{l=1}^{L} \text{Tr} \left( \mathbf{\Sigma}_{l} \mathbf{W}_{l} \right) \leq P_{T} \); the noise covariance at \( R_{l} \) is \( \mathbf{W}_{l} \). For fixed coupling matrix \( \Phi \), the interference-plus-noise covariance and the rate of link \( l \) is given by

\[
\mathbf{\Omega}_{l} = \mathbf{W}_{l} + \sum_{k=1}^{L} \phi_{l,k} \mathbf{H}_{l,k} \mathbf{\Sigma}_{k} \mathbf{H}_{l,k}^{\dagger},
\] (15)

and \( \hat{I}_{l}^{\phi} (\mathbf{\Sigma}) = \log \left| \mathbf{I} + \mathbf{H}_{l,l} \mathbf{\Sigma}_{l} \mathbf{H}_{l,l}^{\dagger} \mathbf{\hat{\Omega}}_{l}^{-1} \right| \) respectively. By definition, the dual network or reverse links is

\[
\left( [\mathbf{H}_{k,l}], \sum_{l=1}^{L} \text{Tr} \left( \mathbf{\hat{S}}_{l} \mathbf{W}_{l} \right) \leq P_{T}, [\mathbf{\hat{W}}_{l}] \right).
\] (16)

The interference-plus-noise covariance and the rate of reverse link \( l \) is given by

\[
\mathbf{\hat{\Omega}}_{l} = \mathbf{\hat{W}}_{l} + \sum_{k=1}^{L} \phi_{k,l} \mathbf{H}_{k,l}^{\dagger} \mathbf{\hat{S}}_{k} \mathbf{H}_{k,l},
\] (17)

and \( \hat{I}_{l}^{\phi^{*}} (\mathbf{\hat{\Sigma}}) = \log \left| \mathbf{I} + \mathbf{H}_{l,l}^{\dagger} \mathbf{\hat{S}}_{l} \mathbf{H}_{l,l} \mathbf{\hat{\Omega}}_{l}^{-1} \right| \) respectively.

The key of the rate duality is a covariance transformation [15] defined as follows.

**Definition 2** (Covariance transformation for B-MAC IFN). The covariance transformation of the input covariance \( \mathbf{\Sigma} \) for a B-MAC IFN with parameters \( \left\{ [\mathbf{H}_{l,k}], [\mathbf{W}_{l}], [\mathbf{\hat{W}}_{l}] \right\} \) as in (14) is defined by 4 steps.

**Step 1**: Decompose the signal of each link \( l \) to \( M_{l} \) streams as

\[
\mathbf{\Sigma}_{l} = \sum_{m=1}^{M_{l}} p_{l,m} \mathbf{t}_{l,m} \mathbf{t}_{l,m}^{\dagger}, l = 1, \ldots, L,
\] (18)

where \( \mathbf{t}_{l,m} \in \mathbb{C}^{L_{i} \times 1} \) is a transmit vector with \( \| \mathbf{t}_{l,m} \| = 1 \) and \( p_{l,m} \)’s are the transmit powers.
Step 2: \( \forall l, m \), compute the MMSE-SIC receive vector as \[ r_{l,m} = \alpha_{l,m} \left( \sum_{i=m+1}^{M_l} H_{l,i} p_{l,i} t_{l,i}^\dagger H_{l,i}^\dagger + \Omega_l \right)^{-1} H_{l,l} t_{l,m}, \] where \( \alpha_{l,m} \) is chosen such that \( \| r_{l,m} \| = 1 \). Calculate the SINR \( \gamma_{l,m} \) as a function of \( \{ p_{l,m}, t_{l,m}, r_{l,m} \} \). 

Step 3: In the reverse links, use \( \{ r_{l,m} \} \) as transmit vectors and \( \{ t_{l,m} \} \) as receive vectors. Apply the SINR duality \[ q = (D^{-1} - \Psi^T)^{-1} \hat{n}, \] where \( \Psi \in \mathbb{R}_{\sum l=1}^{M_l} \times \sum l=1 M_l \) and the \( \left( \sum_{i=1}^{l-1} M_i + m \right)^{th} \) row and \( \left( \sum_{i=1}^{k-1} M_i + n \right)^{th} \) column of \( \Psi \) is \[ \Psi_{k,m}^{l,n} = \begin{cases} \frac{r_{l,m}^\dagger H_{l,l} t_{l,n}}{\phi_{l,k} |r_{l,m}^\dagger H_{l,k} t_{k,n}|^2} & k = l, m < n, \\ 0 & k = l, m \geq n, \end{cases} \] \( D \in \mathbb{R}_{\sum l=1}^{M_l} \times \sum l=1 M_l \) is a diagonal matrix with the \( \left( \sum_{i=1}^{l-1} M_i + m \right)^{th} \) diagonal element given by \[ D_{\sum_{i=1}^{l-1} M_i + m, \sum_{i=1}^{l-1} M_i + m} = \frac{\gamma_{l,m}}{\left( r_{l,m}^\dagger H_{l,l} t_{l,m} \right)^2} \] \( \hat{n} = [\hat{n}_{1,1}, ..., \hat{n}_{1,M_1}, ..., \hat{n}_{L,1}, ..., \hat{n}_{L,M_L}]^T \) with \( \hat{n}_{l,m} = t_{l,m}^\dagger \hat{W}_l t_{l,m} \) \( \forall l, m \). 

Step 4: The Covariance Transformation from \( \Sigma \) to \( \hat{\Sigma} \) is \[ \hat{\Sigma}_l = \sum_{m=1}^{M_l} q_{l,m} r_{l,m} r_{l,m}^\dagger, \quad l = 1, ..., L. \]

We restate the rate duality in \([15]\) as follows.

**Theorem 1** (Rate duality for B-MAC IFN). For any \( \Sigma \) satisfying \( \sum_{l=1}^L \text{Tr} \left( \Sigma_l \hat{W}_l \right) \leq P_T \) and achieving a rate point \( r \) in the network \([14]\), its covariance transformation \( \hat{\Sigma}_{1:L} \) achieves a rate point \( \hat{r} \geq r \) in the dual network \([16]\) and satisfies \( \sum_{l=1}^L \text{Tr} \left( \hat{\Sigma}_l \hat{W}_l \right) = \sum_{l=1}^L \text{Tr} \left( \Sigma_l \hat{W}_l \right) \). Thus, the achievable regions in the forward and reverse links are the same under a single linear constraint.

### III. Duality Results for Multi-hop B-MAC AF Relay Network

In this section, we establish two types of duality for multi-hop B-MAC AF relay network respectively under single network linear constraint and per-hop linear constraint.
A. Type I Duality

The type I duality can be established using the following network equivalence result.

**Theorem 2 (Network Equivalence).** Fixing the relay precoding matrices $F$, the B-MAC AF relay network in Section II-A under constraint (7) is equivalent to the following B-MAC IFN

$$
\left( [\hat{H}_{l,k}], \sum_{l=1}^{L} \text{Tr} \left( \Sigma_l \hat{W}_l^r \right) \leq P_T - P_C, \left[ \hat{W}_l^r \right] \right),
$$

(24)

where $\hat{H}_{l,k}, \forall k, l$ is defined in (2), $W_l^r$ is defined in (5), and

$$
\hat{W}_l = \sum_{q=1}^{Q} H_0^l B_{1,q} \hat{W}_q B_{1,q}^+ H_0^l + \hat{W}_0^l,
$$

(25)

$$
P_C = \sum_{q=1}^{Q} \sum_{q' = q}^{Q} b_{q,q'} \triangleq \text{Tr} \left( B_{q,q'} W_q B_{q,q'}^+ \hat{W}_q \right) \forall q' \geq q.
$$

(26)

Similarly, fixing the dual relay precoding matrices $\hat{F} = (F_1^\dagger, ..., F_Q^\dagger)$, the dual B-MAC AF relay network in Definition 1 under constraint (12) is equivalent to the dual network of (24) given by

$$
\left( [\hat{H}_{k,l}]^\dagger, \sum_{l=1}^{L} \text{Tr} \left( \hat{\Sigma}_l \hat{W}_l^r \right) \leq P_T - P_C, \left[ \hat{W}_l^r \right] \right).
$$

(27)

Please refer to Appendix A for the proof. Define the **Type I dual transformation** as follows.

**Definition 3** (Type I dual transformation). For any input covariance and relay precoding matrices $\Sigma, F$, let $\hat{\Sigma}$ be the covariance transformation of $\Sigma$ in Definition 2 with parameters $\{[\hat{H}_{l,k}], [W_l^r], [\hat{W}_l^r]\}$. Then $\hat{\Sigma}$ and $\hat{F} \triangleq (F_1^\dagger, ..., F_Q^\dagger)$ is called the **Type I dual transformation** of $\Sigma, F$.

Then it follows from Theorem 1 and 2 that the following theorem holds.

**Theorem 3** (Type I duality). For any input covariance and relay precoding matrices $\Sigma, F$ satisfying

$$
\sum_{l=1}^{L} \text{Tr} \left( \Sigma_l \hat{W}_0^r \right) + \sum_{q=1}^{Q} \text{Tr} \left( \Sigma_q^R \hat{W}_q \right) \triangleq P_T^\alpha \leq P_T \text{ and achieving a rate point } r \text{ in a B-MAC AF relay network, its Type I dual transformation } \hat{\Sigma}, \hat{F} \text{ achieves a rate point } \hat{r} \geq r \text{ in the dual network and satisfy }
$$

$$
\sum_{l=1}^{L} \text{Tr} \left( \hat{\Sigma}_l \hat{W}_l^r \right) + \sum_{q=1}^{Q} \text{Tr} \left( \hat{\Sigma}_q^R \hat{W}_q \right) = P_T^\alpha. \text{ Thus, the achievable regions in a B-MAC AF relay network under constraint (7) and that in the dual network under constraint (12) are the same.}
$$

B. Type II Duality

Type II duality can be proved from type I duality using the conception of network dual scaling.

**Definition 4** (Scaled dual network). A scaled dual network with scaling vector $d = [d_1, ..., d_Q, d_{Q+1}]^T$, where $d_q > 0, q = 1, ..., Q$, and $d_{Q+1} = 1$, is obtained by scaling the covariance of the noise vectors in
the dual network defined in Section II-C as follows: the covariance of the noise at the $q^{th}$ relay cluster is scaled to be $d_{q+1}\hat{W}_q$, and the covariance of the noise at the destination $T_l$ is scaled to be $d_1\hat{W}_0$. Furthermore, the relay precoding matrices is given by $\hat{F}$ where $\hat{\Sigma}(19), (21)$ and $(22)$ in Definition 2 with parameters $(Duality for scaled dual network) \text{Corollary 1}$

Then $\hat{\Sigma}(d)$ and $\hat{F} = \left(F_1^+, ..., F_Q^+\right)$ achieves a rate point $\hat{r} \geq r$ in the scaled dual network with scaling vector $d$ and satisfy the constraint

$$\sum_{l=1}^{L} Tr\left(\hat{\Sigma}(d) W_{Q+1}^{l}\right) + \sum_{q=1}^{Q} Tr\left(\hat{\Sigma}_{q}^{R}(d) W_{q}\right) = \sum_{q=0}^{Q} d_{q+1}P_{q}^{x}, \quad (29)$$

where

$$\hat{\Sigma}_{q}^{R}(d) = \sum_{l=1}^{L} B_{q,q}^{l} \hat{H}_{q}^{l} \hat{\Sigma}(d) H_{q}^{l} B_{q,q}^{l} + \sum_{q'=q}^{Q} d_{q'+1} B_{q,q'}^{l} \hat{W}_{q'} B_{q',q''}$$

$$P_{q}^{x} = \sum_{l=1}^{L} Tr\left(\hat{\Sigma}_{l} W_{l}^{0}\right), \quad P_{q}^{0} = Tr\left(\hat{\Sigma}_{q}^{R}(d) W_{q}\right), \quad q = 1, ..., Q.$$

Another theorem is needed to prove the Type II duality. Obtain $\{t_{l,m}\}, \{r_{l,m}\}, \Psi$ and $D$ using (18), (19), (21) and (22) in Definition 2 with parameters $\{[\hat{H}_{l,k}], [W_{l}], [\hat{W}_{l}^{l}]\}$. Define

$$a_{q,q'} = n_{q}^T (D^{-1} - \Psi^T)^{-1} \hat{n}_{q'}, \forall q, q' \in \{0, 1, ..., Q\}, \quad (30)$$

where $n_{q} = \left[n_{1,1,q}, ..., n_{1,M_1,q}, ..., n_{L,1,q}, ..., n_{L,M_L,q}\right]^T$ and $\forall l, m,$

$$n_{l,m}^{q-1} = r_{l,m}^{q} H_{q}^{l} B_{q,q}^{l} W_{q} B_{q,q}^{l} H_{q}^{l} t_{l,m}, \quad q = 1, ..., Q,$$

$$n_{l,m}^{Q} = r_{l,m}^{q} W_{Q+1}^{l} t_{l,m}, \quad (31)$$

$$\hat{n}_{q} = \left[\hat{n}_{q,1,1}, ..., \hat{n}_{q,1,M_1}, ..., \hat{n}_{q,L,1}, ..., \hat{n}_{q,L,M_L}\right]^T$$

and $\forall l, m,$

$$\hat{n}_{l,m}^{q} = t_{l,m}^{q} H_{q}^{l} B_{q,q}^{l} \hat{W}_{q} B_{q,q}^{l} H_{q}^{l} t_{l,m}, \quad q = 1, ..., Q,$$

$$\hat{n}_{l,m}^{Q} = t_{l,m}^{q} \hat{W}_{q}^{l} t_{l,m}. \quad (32)$$
Define $A \in \mathbb{R}^{(Q+1) \times (Q+1)}_+$ whose $(q, q')$th element is
\[
\begin{cases} 
(a_{q-1,q'-1} + b_{q,q'-1}) / P_{q-1}^a, & \text{if } q' > q \\
(a_{q-1,q'-1} / P_{q-1}^a) & \text{otherwise}
\end{cases}
\] (33)

**Theorem 4** (Network Dual Scaling). Consider the following eigensystem
\[
\mathbf{A} \tilde{\mathbf{d}}' = \lambda_{\text{max}} \tilde{\mathbf{d}}',
\] (34)

where $\tilde{\mathbf{d}}'$ is the dominant eigenvector and $\lambda_{\text{max}}$ is the maximum eigenvalue of $A$. Then $\lambda_{\text{max}}$ and $\tilde{\mathbf{d}}'$ must be strictly positive. Let $\hat{\mathbf{d}} = [\hat{d}_1, ..., \hat{d}_Q, 1]$ be the dominant eigenvector of $A$ with the last component scaled to one. Then in the scaled dual network with scaling vector $\hat{\mathbf{d}}$, $\tilde{\mathbf{\Sigma}} (\hat{\mathbf{d}})$ defined in Corollary 1 and $\hat{F} = (\hat{F}_1^\dagger, ..., \hat{F}_Q^\dagger)$ satisfies
\[
\begin{align*}
\sum_{l=1}^L \text{Tr} \left( \hat{\mathbf{\Sigma}}_l \left( \hat{\mathbf{d}} \right) \mathbf{W}_{q+1}^l \right) &= P_Q^a \leq P, \\
\text{Tr} \left( \hat{\mathbf{\Sigma}}_q^R \left( \hat{\mathbf{d}} \right) \mathbf{W}_q \right) &= d_q P_{q-1}^a \leq d_q P_{q-1}^a, \quad q = 1, ..., Q.
\end{align*}
\] (35)

The proof is given in Appendix B. Define the **Type II dual transformation** as follows.

**Definition 5** (Type II dual transformation). For any input covariance and relay precoding matrices $\mathbf{\Sigma}, \mathbf{F}$, let $\hat{\mathbf{\Sigma}} \triangleq \hat{\mathbf{\Sigma}} (\hat{\mathbf{d}})$ be the covariance transformation of $\mathbf{\Sigma}$ in Definition 2 with parameters $\{[\mathbf{H}_l,k], [\mathbf{W}_l], [\mathbf{W}'_l (\hat{\mathbf{d}})]\}$, where $\hat{\mathbf{d}} = [\hat{d}_1, ..., \hat{d}_Q, 1]$ is the dominant eigenvector of the eigensystem in (34). Then $\hat{\mathbf{\Sigma}}$ and the dual relay precoding matrices $\hat{\mathbf{F}} \triangleq \hat{\mathbf{F}} (\hat{\mathbf{d}}) = (\hat{c}_1 \mathbf{F}_1^\dagger, ..., \hat{c}_Q \mathbf{F}_Q^\dagger)$, where $\hat{c}_Q = 1 / \sqrt{\hat{d}_Q}$, $\hat{c}_q = \sqrt{\hat{d}_q+1 / \hat{d}_q}$, $q = 1, ..., Q - 1$, is called the **Type II dual transformation** of $\mathbf{\Sigma}, \mathbf{F}$.

For the same input covariance $\hat{\mathbf{\Sigma}} \triangleq \hat{\mathbf{\Sigma}} (\hat{\mathbf{d}})$, the rate point in the scaled dual network with scaling vector $\hat{\mathbf{d}}$ and relay precoding matrices $\{\hat{\mathbf{F}}_1^\dagger, ..., \hat{\mathbf{F}}_Q^\dagger\}$ is equal to that in the original dual network with relay precoding matrices $\hat{\mathbf{F}} \triangleq \hat{\mathbf{F}} (\hat{\mathbf{d}})$. Furthermore, it follows from (35) that the Type II dual transformation $\hat{\mathbf{\Sigma}}, \hat{\mathbf{F}}$ satisfies (13). Combining the above and Corollary 1 the following theorem is proved.

**Theorem 5** (Type II duality). For any input covariance and relay precoding matrices $\mathbf{\Sigma}, \mathbf{F}$ satisfying $\sum_{l=1}^L \text{Tr} \left( \hat{\mathbf{\Sigma}}_l \hat{\mathbf{W}}_0^l \right) \triangleq P_0^a \leq P_0$, $\text{Tr} \left( \hat{\mathbf{\Sigma}}_q^R \hat{\mathbf{W}}_q \right) \triangleq P_q^a \leq P_q$, $\forall q$, and achieving a rate point $r$ in the B-MAC AF relay network defined in Section II-A, its Type II dual transformation $\hat{\mathbf{\Sigma}}, \hat{\mathbf{F}}$ achieves a rate point $\hat{r} \geq r$ in the dual network and satisfy $\sum_{l=1}^L \text{Tr} \left( \hat{\mathbf{\Sigma}}_l \hat{\mathbf{W}}_{q+1}^l \right) = P_Q^a$, $\text{Tr} \left( \hat{\mathbf{\Sigma}}_q^R \hat{\mathbf{W}}_q \right) = P_{q-1}^a$, $\forall q$. Thus, the achievable regions in a B-MAC AF relay network under constraint (8) and that in the dual network under constraint (13) are the same.
Remark 2 (Generality of Type I/II duality). The previous duality results for various special cases are compared to illustrate the generality of the Type I/II duality. In [11] and [12], the duality was established respectively for multi-hop MAC/BC AF relay networks with single-antenna source/destination nodes and two-hop MIMO MAC/BC AF relay networks. However, the approach in [11], [12] cannot be easily extended to the general B-MAC AF relay network. The duality for multi-hop MIMO AF relay system was established in [13]. Although the proof can be extended to multi-user AF relay networks by using block diagonal precoding / receiving matrices and the notion of independent streams, the duality in [3] cannot cover the duality in this paper as elaborated below.

1) The B-MAC AF relay network is not captured by the system model in [13]. In [13], there are $N_b$ independent data streams in one transmission. Similar to the coupling matrix defined in Assumption 3 we can use a binary inter-stream coupling matrix $\Phi^s \in \{0, 1\}^{N_b \times N_b}$ to specify the interference among the data streams. In [13], $\Phi^s$ can only have two forms: a) if linear transceivers are used at the source and destination nodes, all the diagonal elements of $\Phi^s$ are zero, and all the off-diagonal elements of $\Phi^s$ are one; b) if non-linear transceivers are used, $\Phi^s$ is a triangular matrix. However, for general B-MAC AF relay networks, if we decompose the equivalent MIMO links into independent data streams, the inter-stream coupling matrix $\Phi^s$ can be any binary matrix with zero diagonal elements.

2) Explicit dual transformations are not part of [13] for the multi-user case with general inter-stream coupling matrix $\Phi^s$. The dual transformation between the DPC-based and SIC-based MIMO AF relay systems in [13] cannot be generalized to B-MAC AF relay networks because the matrix “$\Phi$” in (14) of [13] is no longer upper-triangle.

Furthermore, the above special cases consider power constraints only. By using the techniques of network equivalence and network dual scaling, we are able to establish the dualities and explicit dual transformations for B-MAC AF relay networks under more general linear constraints. The duality proof based on these new techniques is not only simpler but also reveals more insight on the duality structure that can be exploited to design MIMO precoder optimization algorithms as shown in the next section.

IV. APPLICATIONS IN NETWORK OPTIMIZATION PROBLEMS

We propose an unified optimization framework to find a stationary point for a class of AF relay network optimization problems based on the local Lagrange dual method (LDM) [23], where the primal algorithm only finds a stationary point for the inner loop problem of maximizing the Lagrangian. The duality is first used to characterize the PWF structure of the input covariance matrices at a stationary point. Then the duality and PWF are exploited to design efficient primal algorithm for general B-MAC AF relay networks.
and structured primal algorithms for BC AF Relay network.

A. Unified Optimization Framework based on Local LDM

A general optimization problem in a B-MAC AF Relay network can be expressed as

$$\max f(\Sigma, F), \text{s.t. } g_n(\Sigma, F) \geq 0, \ n = 1, \ldots, N, \text{ and } \Sigma_l \succeq 0, \ l = 1, \ldots, L,$$

(36)

where $f(\Sigma, F)$ and $g_n(\Sigma, F)$, $\forall n$ are real valued functions of $\Sigma, F$. In this paper, we consider a class of problems whose Lagrangian can be written in the following form

$$L(\Sigma, F, \lambda) \triangleq f(\Sigma, F) + \sum_{n=1}^{N} \lambda_n g_n(\Sigma, F)$$

$$= \sum_{l=1}^{L} w_l \mathcal{I}_l^b(\Sigma, F) - \sum_{l=1}^{L} \text{Tr} \left( \Sigma_l \hat{W}_l^0 \right) - \sum_{q=1}^{Q} \text{Tr} \left( \Sigma_q^R \hat{W}_q \right), \ \Sigma_l \succeq 0, \ \forall l,$$

(37)

where $\lambda = [\lambda_n \geq 0]_{n=1,\ldots,N}$ are Lagrange multipliers, $\hat{W}_l^0, \ \forall l$ are positive semidefinite, $\hat{W}_q, \ \forall q$ are positive definite, and $w_l \geq 0, \ \forall l$. The Lagrangian of many important optimization problems can be written in the form of (37). Two typical problems are listed below.

P1: Maximize the weighted sum-rate under multiple linear constraints:

$$\max \sum_{l=1}^{L} \mu_l \mathcal{I}_l^b(\Sigma, F)$$

(38)

$$\text{s.t. } \sum_{l=1}^{L} \text{Tr} \left( \Sigma_l \hat{W}_l^l \right) + \sum_{q=1}^{Q} \text{Tr} \left( \Sigma_q^R \hat{W}_{n,q} \right) \leq \delta_n, \ n = 1, \ldots, N, \ \text{and } \Sigma_l \succeq 0, \ \forall l,$$

where $\mu_l \geq 0, \ \forall l$, $\hat{W}_l^l$’s and $\hat{W}_{n,q}$’s are positive semidefinite and $\delta_n \geq 0, \ \forall n$.

P2: Minimize a single linear cost under individual rate constraints:

$$\min \sum_{l=1}^{L} \text{Tr} \left( \Sigma_l \hat{W}_l^0 \right) + \sum_{q=1}^{Q} \text{Tr} \left( \Sigma_q^R \hat{W}_q \right)$$

(39)

$$\text{s.t. } \mathcal{I}_l^b(\Sigma, F) > I_l^0, \ l = 1, \ldots, L, \ \text{and } \Sigma_l \succeq 0, \ \forall l,$$

where $I_l^0 \geq 0$ is the rate constraint for the $l$th data link.

Remark 3 (Per-relay power constraint). The multiple linear constraints in P1 include the per-relay power constraint as special cases. Let $n_q$ denote the number of relays in the $q$th relay cluster, and let $L_{q,j}$ denote the number of antennas at the $j$th relay of the $q$th relay cluster. Then the power constraint $P_{q,j}$ for the $j$th relay in the $q$th relay cluster can be expressed in the form of the general multiple linear constraints in P1 as $\sum_{l=1}^{L} \text{Tr} \left( \Sigma_l \hat{W}_l^l \right) + \sum_{q=1}^{Q} \text{Tr} \left( \Sigma_q^R \hat{W}_{n,q} \right) \leq \delta_n$ by setting $\delta_n = P_{q,j}$, $\hat{W}_n^l = 0, \ \forall l$,.
\( \hat{W}_{n,q'} = 0, \forall q' \neq q, \) and \( \hat{W}_{n,q} = \text{BlockDiag} \left[ 0_1, \ldots, 0_{j-1}, I_j, 0_{j+1}, \ldots, 0_{n_q} \right] \), where \( I_j \) is an \( L_{q,j} \times L_{q,j} \) identity matrix and \( 0_{j'}, \forall j' \neq j \) is an \( L_{q,j'} \times L_{q,j'} \) zero matrix.

It is well known that the problems in (38) and (39) are usually non-convex problems with non-zero duality gap. Hence, the standard Lagrange dual method (LDM) [10] cannot be used to solve these problems. In [23], we proposed a local LDM to find a stationary point for a non-convex problem. We apply the local LDM to obtain the following algorithm for problem (36).

**Algorithm LLDM** (for Finding a Stationary Point of Problem (36)):

**Initialization:** Choose initial \( \lambda^{(0)} \) such that \( \lambda^{(0)}_n > 0, \forall n \). Let \( i = 1 \).

**Step 1** (Primal update in the inner loop): For fixed \( \lambda = \lambda^{(i-1)} \), find a stationary point \( \bar{\Sigma}^{(i-1)}, \bar{F}^{(i-1)} \) of

\[
\max_{\Sigma, F} \left\{ L_\lambda (\Sigma, F) \right\} = L (\Sigma, F, \lambda), \text{ s.t. } \Sigma_l \succeq 0, \forall l,
\]

using some primal algorithm with initial point \( \bar{\Sigma}^{(i-2)}, \bar{F}^{(i-2)} \).

**Step 2** (Dual update in the outer loop): Update \( \lambda \) as

\[
\lambda^{(i)} = \lambda^{(i-1)} + t^{(i)} z^{(i)},
\]

where \( t^{(i)} \) is the step size, and \( z^{(i)} \) is the optimal solution of the following quadratic programming problem:

\[
\min_z z^T g^{(i)} + \frac{1}{2} z^T J^{(i)} z, \text{ s.t. } \lambda^{(i-1)} + z \geq 0,
\]

where \( g^{(i)} = \left[ g_1^{(i)}, \ldots, g_N^{(i)} \right]^T \) with \( g_n^{(i)} = g_n (\bar{\Sigma}^{(i-1)}, \bar{F}^{(i-1)}) \), \( \forall n \), and \( J^{(i)} \in \mathbb{R}^{N \times N} \) is positive definite.

**Return to Step 1 until convergence.**

**Choice of the Matrix \( J^{(i)} \):** \( J^{(i)} \) is obtained by the well known BFGS update as follows [24]

\[
J^{(i+1)} = \begin{cases} 
J^{(i)} + \frac{q_i q_i^T}{t_i^T p_i} - \frac{J^{(i)} p_i p_i^T J^{(i)}}{t_i^T p_i}, & q_i^T p_i > 0, \\
J^{(i)}, & \text{otherwise},
\end{cases}
\]

\[
p_i = \lambda^{(i)} - \lambda^{(i-1)}, \quad q_i = g^{(i+1)} - g^{(i)}.
\]

The initial \( J^{(1)} = 2m_b I \), where \( m_b \) is the smallest positive integer such that \( \max \frac{|\check{z}_n|}{\check{z}_n} < 0.5 \), and \( \check{z} = [\check{z}_1, \ldots, \check{z}_N] \) is the optimal solution of the problem

\[
\max_{\check{z}} z^T g^{(1)} + \frac{1}{2} z^T J^{(1)} z.
\]

\(^1\)Note that \( \bar{\Sigma}^{(-1)}, \bar{F}^{(-1)} \) is randomly generated.
Choice of the Step Size $t^{(i)}$: Set $t^{(i)} = \alpha^{(i)}2^{-mt}$, where $m_t$ is an integer which is initialized as 0 and is incremented until one of the following conditions is satisfied: 1) $L (\Sigma^{(i)}, F^{(i)}, \lambda^{(i)}) \leq L (\Sigma^{(i-1)}, F^{(i-1)}, \lambda^{(i-1)})$. 2) $m_t = m_0^t$, where $m_0^t$ is a small positive integer, e.g., we let $m_0^t = 2$ in the simulations. 3) $r_{ei} \leq r_{ei-1}$, where

$$r_{ei} = \max_n \left| \lambda_n \left( \Sigma^{(i)}, F^{(i)} \right) \right| + \left( \max_n \left\{ g_n \left( \Sigma^{(i)}, F^{(i)} \right) \right\} \right)^+,$$

is defined as the residual error after the $i$th iteration. Finally, $\alpha^{(i)}$ is given by

$$\alpha^{(i+1)} = (1 - \beta) \alpha^{(i)} + \beta t^{(i)},$$

where $0 < \alpha^{(0)} \leq 0.5$, $0 < \beta < 1$ and are set as $\alpha^{(0)} = 0.25$, $\beta = 0.2$ in the simulations.

Algorithm LLDM converges to a stationary point of Problem (46) under mild conditions [23]. In the subsequent sections, we first characterize the PWF structure at a stationary point of Problem (40). Then we design several efficient primal algorithms based on the PWF and duality.

B. Structural Properties of a Stationary Point

At any stationary point of Problem (40) with fixed relay precoding matrices $F$, the input covariance matrices $\Sigma$ must satisfy a PWF structure as stated in the following corollary.

Corollary 2 (PWF for fixed relay precoding). Let $\tilde{\Sigma}$ be a stationary point of problem (40) with fixed $F$, i.e., $\tilde{\Sigma}$ satisfies the following KKT conditions

$$\nabla_{\Sigma} L_\lambda (\Sigma, F) \preceq 0, \forall l,$$

$$\text{Tr} \left( \Sigma \nabla_{\Sigma} L_\lambda (\Sigma, F) \right) = 0, \forall l.$$

Then the Type I dual transformation of $\Sigma, F$ is $\tilde{\Sigma}, \tilde{F} = \left( F_1^T, ..., F_L^T \right)$, where $\tilde{\Sigma} = \left( \tilde{\Sigma}_1, ..., \tilde{\Sigma}_L \right)$ with

$$\tilde{\Sigma}_l = w_l \left( \Omega_l^{-1} - \left( \Omega_l + \tilde{H}_{l,1} \Sigma_l \tilde{H}_{l,1}^T \right)^{-1} \right), \forall l.$$

Obtain the interference-plus-noise covariance matrices $\tilde{\Omega}_l$'s from (4) using $\tilde{\Sigma}, F$ and the dual ones $\tilde{\tilde{\Omega}}_l$'s from (10) using $\tilde{\tilde{\Sigma}}, \tilde{\tilde{F}}$. For each $l$, create an equivalent channel: $\tilde{H}_{l,1} = \tilde{\Omega}_l^{-1/2} \tilde{H}_{l,1} \tilde{\Omega}_l^{-1/2}$. Perform the thin SVD $\tilde{H}_{l,1} = E_l \Delta_l G_l^T$, where $E_l \in \mathbb{C}^{L_{l,1} \times N_l}$, $G_l \in \mathbb{C}^{L_{l,1} \times N_l}$, $\Delta_l \in \mathbb{R}^{N_l \times N_l}$, and $N_l = \text{Rank} (\tilde{H}_{l,1})$. Then, $\tilde{\Sigma}$ must have a polite water-filling structure, i.e.,

$$Q_l \triangleq \tilde{\Omega}_l^{1/2} \Sigma_l \tilde{\Omega}_l^{1/2} = G_l (w_l I - \Delta_l^{-2})^+ G_l^T, \forall l.$$

Furthermore, $\tilde{\Sigma}$ also possesses the polite water-filling structure, i.e.,

$$\hat{Q}_l \triangleq \tilde{\Omega}_l^{1/2} \Sigma_l \tilde{\Omega}_l^{1/2} = E_l (w_l I - \Delta_l^{-2})^+ E_l^T, \forall l.$$
On the other hand, if certain $\Sigma$ has the above polite water-filling structure for a given $F$, it must be a stationary point of problem (40) for this fixed $F$. 

Corollary 2 follows straightforward from the PWF results for B-MAC IFN in Theorem 3 of [16] and the network equivalence result in Theorem 2. The detailed proof is omitted due to limited space.

In some cases, better algorithms can be designed by considering the following dual network problem.

**Definition 6 (Dual network problem).** The dual network problem of (40) is defined as

$$
\max_{\Sigma, \hat{F}} \hat{L}_{\lambda}(\Sigma, \hat{F}) \triangleq \sum_{l=1}^{L} w_l \hat{X}_l^{\phi_p} \left( \Sigma_l, \hat{F} \right) - \sum_{l=1}^{L} \text{Tr} \left( \Sigma_l W_{l+1} \right) - \sum_{q=1}^{Q} \text{Tr} \left( \hat{\Sigma}_q^R W_q \right), \text{ s.t. } \hat{\Sigma}_l \succeq 0, \forall l. \tag{48}
$$

**Theorem 6 (Equivalence of problem (40) and its dual).** For any stationary point $\hat{\Sigma}, \hat{F}$ of dual network problem (48), its Type I dual transformation $\bar{\Sigma}, \bar{F}$ must also be a stationary point of problem (40).

Please refer to Appendix C for the proof.

**C. Primal Algorithms for two-hop B-MAC AF Relay networks**

We present several primal algorithms for two-hop B-MAC AF Relay networks. In this case, the relay precoding matrices is $F = \text{BlockDiag} \left[ F(1), \ldots, F(n_1) \right]$, where $F(j) \in \mathbb{C}^{L_{1,j} \times L_{1,j}}$ is the relay precoding matrix at the $j$th relay, $L_{1,j}$ is the number of antennas at the $j$th relay, and $n_1$ is the number of relays.

1) **Gradient Ascent Primal Algorithm :** The gradient ascent (GA) algorithm with the step size determined by Armijo rule [17] can be used to find a stationary point of Problem (40) as summarized in Table I. In Algorithm GA, the gradient update for source node is performed over the precoding matrices determined by Armijo rule [17] can be used to find a stationary point of Problem (40) as summarized in Table I. The detailed proof is omitted due to limited space.

$$
\nabla_{\Sigma_l}^{\phi_p} \left( \Sigma_l, \hat{F} \right) = \sum_{l=1}^{L} w_l \hat{X}_l^{\phi_p} \left( \Sigma_l, \hat{F} \right) - \sum_{l=1}^{L} \text{Tr} \left( \Sigma_l W_{l+1} \right) - \sum_{q=1}^{Q} \text{Tr} \left( \hat{\Sigma}_q^R W_q \right), \text{ s.t. } \hat{\Sigma}_l \succeq 0, \forall l. \tag{48}
$$

**Please refer to Appendix C for the proof.**
Initialize $T_l$, $\forall l$ and $F = \text{BlockDiag}[F(1),...,F(n_1)]$.

**Do**

1. Calculate the gradient over $T_l$'s using (49): $G_{T_l} = \nabla_{T_l} L_{\lambda}(\Sigma, F), l = 1,...,L$.
2. Choose step size $\alpha$ via Armijo rule and update $T_l$'s as $T_l = T_l + \alpha G_{T_l}, l = 1,...,L$.
3. Calculate the gradient over $F(j)$'s using (50): $G_{F_j} = \nabla_{F(j)} L_{\lambda}(\Sigma, F), j = 1,...,n_1$.
4. Choose step size $\alpha$ via Armijo rule and update $F$ as $F(j) = F(j) + \alpha G_{F_j}, j = 1,...,n_1$.

**until** all the norms of $G_{T_l}, \forall l$ and $G_{F_j}, \forall j$ are small enough.

2) Polite Water-filling based Primal Algorithm: We exploit the duality and PWF structure to design Algorithm PWF as summarized in Table II. By Corollary 2 and the property of gradient update for $F(j)$'s, if Algorithm PWF converges, the solution must be a stationary point of problem (40). Simulations show that Algorithm PWF has a faster convergence speed than Algorithm GA.

3) Structured Primal Algorithm for two-hop BC AF relay networks: Consider sum rate maximization under multiple linear constraints (i.e., problem (38) with $\mu_l = 1, \forall l$) in a two-hop BC AF relay network applying DPC. We apply the duality and Corollary 2 to design a structured primal algorithm to find a stationary point of the inner loop problem (40) for this special case.

Note that in this case, $Q = 1$, $H^l_0, \forall l$ are the same, and $\hat{W}_0^l, \forall l$ are the same. For convenience, let

$$
\left( H_b, \begin{bmatrix} H^l_b \end{bmatrix}, W^l_b, \begin{bmatrix} W^l_b \end{bmatrix}, \hat{W}_b, \hat{W}_b^l \right),
$$

(52)
denote a two-hop BC AF relay network where the channel matrix between source and relay is $H^l_b = H_b, \forall l$, the channel matrix between relay and the $j^{th}$ destination is $H^l_Q = H_b$, the covariance of the noise at the relay is $W_1 = W_b$, the covariance of the noise at the $j^{th}$ destination is $W^l_j = W_b$, and the constraint matrices in Lagrangian (37) is $\hat{W}_0^l = \hat{W}_b, \forall l$, $\hat{W}_1 = \hat{W}_b^l$. The dual network of (52) is

$$
\left( \begin{bmatrix} H^l_m \end{bmatrix}, H_m, W^l_m, W_m, \begin{bmatrix} \hat{W}_m^l \end{bmatrix}, \hat{W}_m^l \right),
$$

(53)

which is a two-hop MAC AF relay network where the channel matrix between the $l^{th}$ source and relay is $H^l_m = H^l_b$, $\forall l$, the channel matrix between relay and the destination is $H_m = H_b$, the covariance of the noise at the relay is $W^l_m = \hat{W}_b$, the covariance of the noise at the $l^{th}$ destination is $W_m = \hat{W}_b$, and the constraint matrices in Lagrangian (37) is $\hat{W}_m^l = \hat{W}_b, \forall l$, $\hat{W}_m = \hat{W}_b$. By Theorem 6 Problem (40) for the BC relay network (52) can be equivalently solved by solving the dual network problem (48), whose
Initialize $\Sigma$ and $F = \text{BlockDiag} \{F(1), \ldots, F(n_1)\}$ such that $\Sigma_l \succeq 0, \forall l$.

Obtain the initial $\hat{\Sigma}$ from (45) using the initial $\Sigma$.

**While** not converge **do**

1. For fixed $\hat{\Sigma}$ and $F$, update $\Sigma$ by polite water-filling:

   Obtain $\hat{\Omega}_l = \hat{W}_l + \sum_{k=1}^L \Phi_{k,l} \hat{H}_{k,l}^\dagger \Sigma_k \hat{H}_{k,l}, \forall l$ where $\hat{W}_l$ is given in (25).

   **For** $l = 1$ to $L$

   a. Obtain $\Omega_l$ using (4).

   b. Perform thin SVD $\Omega_l^{-1/2} \hat{H}_{l,l} \Omega_l^{-1/2} = E_l \Delta_l G_l^\dagger, \forall l$.

   c. Update $\Sigma_l$ as $\Sigma_l = \hat{\Omega}_l^{-1/2} E_l (w_l I - \Delta_l^{-2})^+ G_l^\dagger \hat{\Omega}_l^{-1/2}$.

   **End**

2. For fixed $\Sigma$, update $F$ as $F(j) = F(j) + \alpha \nabla F(j) \hat{L}_\lambda(\Sigma, F), j = 1, \ldots, n_1$, where the step size $\alpha$ is obtained by Armijo rule.

3. For fixed $\Sigma$ and $F$, update $\hat{\Sigma}$ by polite water-filling:

   Obtain $\hat{\Omega}_l, \forall l$ using (4).

   **For** $l = 1$ to $L$

   a. obtain $\hat{\Omega}_l = \hat{W}_l + \sum_{k=1}^L \Phi_{k,l} \hat{H}_{k,l}^\dagger \Sigma_k \hat{H}_{k,l}$, where $\hat{W}_l$ is given in (25).

   b. Perform thin SVD $\Omega_l^{-1/2} \hat{H}_{l,l} \Omega_l^{-1/2} = E_l \Delta_l G_l^\dagger, \forall l$.

   c. Update $\hat{\Sigma}_l$ as $\hat{\Sigma}_l = \Omega_l^{-1/2} E_l (w_l I - \Delta_l^{-2})^+ E_l \Omega_l^{-1/2}$.

   **End**

End

objective function, in this special case, is given by

$$\hat{L}_\lambda(\hat{\Sigma}, \hat{F}) = \log \left| \frac{H_m^\dagger F R \hat{F}^\dagger H_m^* + H_m^\dagger \hat{F} W_m^* \hat{F}^\dagger H_m^* + W_m}{H_m^\dagger \hat{F} W_m^* \hat{F}^\dagger H_m^* + W_m} \right|$$

$$- \text{Tr} \left( \hat{F}^\dagger (R + W_m^*) \hat{F}^\dagger \hat{W}_m^* \right) - \sum_{l=1}^L \text{Tr} \left( \hat{\Sigma}_l \hat{W}_m^l \right),$$

where $R = \sum_{l=1}^L H_m^\dagger \hat{\Sigma}_l H_m^*$. 

By exploiting the specific structure of $\hat{L}_\lambda$, we propose an improved algorithm over PWF.

**Algorithm PWFI:**

**Initialization:** Choose proper initial $\hat{\Sigma}, \hat{F}$ such that $\hat{\Sigma}_l \succeq 0, \forall l$.

**Step 1:** For fixed $\hat{F}$ and $\hat{\Sigma}_k, \forall k \neq l$, Problem (48) reduces to the single-user optimization problem:

$$\max_{\Sigma_l} \log \left| \tilde{\Sigma}_l \hat{W}_m^l + \tilde{W}_l \right| - \text{Tr} \left( \hat{\Sigma}_l \hat{W}_m^l \right),$$
where $\tilde{H}_l = H_m \tilde{F} H_m^†$, $\tilde{W}_l = H_m \tilde{F} \sum_{k \neq l} H_m^† \tilde{\Sigma}_k H_m^† \tilde{F} H_m^† + H_m \tilde{F} W_m \tilde{F} H_m^† + W_m$, and $\tilde{W}_l = H_m^† \tilde{F} W_m \tilde{F} H_m^† + \tilde{W}_m$. Applying Corollary [2] to the single-user optimization problem in (55), the optimal solution is given by

$$\tilde{\Sigma}_l = \tilde{W}_l^{-1/2} G_l (I - \Delta_l^{-2}) + G_l^† \tilde{W}_l^{-1/2}, \forall l,$$

(56)
where $G_l$ and $\Delta_l$ are obtained by performing the thin SVD $\tilde{W}_l^{-1/2} \tilde{H}_l \tilde{W}_l^{-1/2} = E_l \Delta_l G_l^†$. Update each $\tilde{\Sigma}_l$ for once using (56) when fixing $\tilde{F}$ and $\tilde{\Sigma}_k$, $\forall k \neq l$.

**Step 2:** For fixed $\tilde{\Sigma}$, $\tilde{F}$ is updated by solving max $\hat{L}_\lambda (\tilde{\Sigma}, \tilde{F})$, which is equivalent to the following problem

$$\max_{\tilde{F}} \log \left| \frac{\tilde{H}_l^† \tilde{F} \tilde{R} \tilde{H} + \tilde{H}_l^† \tilde{F} \tilde{H} + I}{\tilde{H}_l^† \tilde{F} \tilde{H} + I} \right| - \text{Tr} \left( \tilde{F}^† (\tilde{R} + I) \tilde{F} \right),$$

(57)
where $\tilde{H} = (\tilde{W}_m)_{-1/2} H_m^† \tilde{W}_m^{-1/2}$, $\tilde{R} = (\tilde{W}_m)_{1/2} \tilde{F}^† \tilde{W}_m$ and $\hat{L}_\lambda = \min (\text{Rank} (\tilde{H}), \text{Rank} (\tilde{R}))$. Perform the SVD $\tilde{H} = U_h D_h V_h$ and the eigenvalue decomposition $\tilde{R} = U_r D_r V_r$.

Let $\sigma_1, \ldots, \sigma_M$ denote the $M$ largest singular values of $\tilde{H}$ with descending order and $U_h^M$ denote the semi-unitary matrix formed by the singular vectors corresponding to $\sigma_i$'s. Let $\delta_1, \ldots, \delta_M$ denote the $M$ largest eigenvalues of $\tilde{R}$ with descending order and $U_r^M$ denote the semi-unitary matrix formed by the eigenvectors corresponding to $\delta_i$'s. Then the optimal $\tilde{F}$ is given by

$$\tilde{F} = \left( \tilde{W}_m^r \right)^{-1/2} U_r^M D_f U_r^M \left( \tilde{W}_m^r \right)^{-1/2},$$

(58)
where $D_f = \text{diag} (f_1, \ldots, f_M)^{1/2} \geq 0$ with $f_i$, $i = 1, \ldots, M$ given by

$$f_i = \frac{1}{2 \sigma_i^2 (\delta_i + 1)} \left[ \sqrt{\delta_i^2 + 4 \delta_i \sigma_i^2} - \delta_i - 2 \right]^+.$$

**Return to step 1 until convergence.**

Since the objective $\hat{L}_\lambda (\tilde{\Sigma}, \tilde{F})$ is upper bounded and is increased after each update, Algorithm PWFI must converge to a fixed point $\tilde{\Sigma}, \tilde{F}$. It can be shown that the fixed point $\tilde{\Sigma}, \tilde{F}$ must be a stationary point of the dual network problem (48). Then by Theorem [6] the corresponding Type I dual transformation $\tilde{\Sigma}, \tilde{F}$ is a stationary point of Problem (40) in the BC relay network (52).

**D. Primal Algorithms for Multi-hop B-MAC AF Relay Networks**

Algorithm GA/PWF can be easily extended to solve Problem (40) for B-MAC AF relay networks with more than two hops. The only difference is that the gradient update of the relay precoding matrices is generalized to multi-hop case as

$$F_q (j) = F_q (j) + \alpha \nabla_{F_q (j)} L_\lambda (\Sigma, F), q = 1, \ldots, Q, j = 1, \ldots, n_q,$$

where the gradient $\nabla_{F_q (j)} L_\lambda (\Sigma, F)$, $\forall q, j$ can be calculated similar to the one in (50).
The duality can be used to simplify the the sum rate maximization under multiple linear constraints for a three-hop BC relay network. We design Algorithm PWF3 to find a stationary point of the inner-loop problem (40) for this special case, where $\mathbf{H}_l^0 = \mathbf{H}_0$, $\forall l$, $l \neq 1$, $\forall l$ and $\hat{\mathbf{W}}_0^l = \hat{\mathbf{W}}_0$, $\forall l$. Algorithm PWF3 switches the optimization between the original network and the dual network so that at each time, we only need to consider a two-hop network optimization problem solved in Section IV-C3.

Algorithm PWF3:

Initialization: Choose proper initial $\Sigma, \mathbf{F} = (\mathbf{F}_1, \mathbf{F}_2)$ such that $\Sigma_l \succeq 0$, $\forall l$.

Step 1: For fixed $\mathbf{F}_1$, the inner loop problem reduces to Problem (40) in the two-hop BC AF relay network (52) with $\mathbf{H}_b = \mathbf{H}_1 \mathbf{F}_1 \mathbf{H}_0$, $\mathbf{H}_b^l = \mathbf{H}_Q^l$, $\forall l$, $\mathbf{W}_b^l = \mathbf{H}_1 \mathbf{F}_1 \mathbf{F}_1^\dagger \mathbf{H}_1^l + \mathbf{W}_2$, $\mathbf{W}_b^l = \mathbf{W}_Q^{l+1}$, $\forall l$, $\hat{\mathbf{W}}_b = \mathbf{H}_b^l \mathbf{F}_1^\dagger \hat{\mathbf{W}}_1 \mathbf{F}_1 \mathbf{H}_0 + \hat{\mathbf{W}}_0$ and $\hat{\mathbf{W}}_b = \hat{\mathbf{W}}_2$. Obtain the input covariance and relay precoding matrices $\hat{\Sigma}_b = \left( \hat{\Sigma}_b^1, \ldots, \hat{\Sigma}_b^c \right)$, $\hat{\mathbf{F}}_b$ for the dual two-hop MAC AF relay network (53) by the Type I dual transformation of $\Sigma, \mathbf{F}_2$ applied to network (52). Update each $\hat{\Sigma}_b^l$ for once using (56) when fixing $\hat{\mathbf{F}}_b$ and $\hat{\Sigma}_b^l$, $\forall k \neq l$. Then update $\hat{\mathbf{F}}_b$ using (58) for fixed $\hat{\Sigma}_b$. Finally, obtain the updated $\Sigma, \mathbf{F}_2$ by the Type I dual transformation of $\hat{\Sigma}_b, \hat{\mathbf{F}}_b$ applied to network (53).

Step 2: Obtain $\hat{\Sigma}, \hat{\mathbf{F}}_2$ by the Type I dual transformation of $\Sigma, \mathbf{F}$.

Step 3: For fixed $\hat{\mathbf{F}}_2$, update $\hat{\Sigma}, \hat{\mathbf{F}}_1$ by improving the objective value of the dual network problem (48), which reduces to Problem (40) in the two-hop MAC AF relay network (53) with $\mathbf{H}_m^l = \mathbf{H}_1^l \hat{\mathbf{F}}_2 \mathbf{H}_Q^l$, $\forall l$, $\mathbf{H}_m = \mathbf{H}_0$, $\mathbf{W}_m = \mathbf{H}_1^l \hat{\mathbf{F}}_2^\dagger \mathbf{W}_2 \hat{\mathbf{F}}_2^\dagger \mathbf{H}_1 + \hat{\mathbf{W}}_1$, $\mathbf{W}_m = \hat{\mathbf{W}}_0$, $\hat{\mathbf{W}}_m^l = \mathbf{H}_Q^l \hat{\mathbf{F}}_2^\dagger \mathbf{W}_2 \hat{\mathbf{F}}_2^\dagger \mathbf{H}_Q^l + \mathbf{W}_Q^{l+1}$, and $\hat{\mathbf{W}}_m^l = \mathbf{W}_1$. Treating $\hat{\Sigma}, \hat{\mathbf{F}}_1$ as the input covariance and relay precoding matrices in network (53), update $\hat{\Sigma}, \hat{\mathbf{F}}_1$ using a single iteration of Algorithm PWFI applied to network (53).

Step 4: Obtain $\Sigma, \mathbf{F}$ by the Type I dual transformation of $\hat{\Sigma}, \hat{\mathbf{F}}$.

Return to step 1 until convergence.

Using Theorem 3 it can be shown that Algorithm PWF3 monotonically increases the objective after each iteration and converges to a stationary point of Problem (40).

Remark 4. For the special case of B-MAC IFN, the duality and PWF have been established in [15], [16]. However, we cannot extend the existing results and algorithms trivially from [15], [16]. We proposed new techniques, namely the network equivalence technique in Theorem 2 and the network dual scaling technique in Theorem 4 to generalize the duality and PWF to B-MAC AF relay networks. The proof of Theorem 4 is non-trivial. The algorithm design also relies on the new network equivalence result. Moreover, Theorem 6 (equivalence of problem (40) and its dual) is non-trivial and cannot be re-derived from our previous results in [15], [16]. Based on Theorem 6 we proposed structured primal algorithms (PWFI and PWF3) which are entirely different from the single-hop PWF algorithms in [15], [16].
Block fading channel is assumed and each channel matrix has zero-mean i.i.d. Gaussian entries with unit variance. The simulation results in all figures are averaged over 100 random channel realizations.

A. Verify the Convergence of the Proposed Primal Algorithms

The parameters of Problem (40) are set as \( w_l = 1, \hat{W}_0^l = 0.01I, \hat{W}_q = 0.01I, \forall l \). For accuracy comparison, define the convergence error as the sum of the norms of the gradients over \( F \) and all \( T_l \)'s.

We first demonstrate the advantages of Algorithm PWF over GA described in Section IV-C1. Consider the B-MAC AF relay network in Fig. 1 with \( Q = 1 \) (two-hop), \( n_1 = 1 \) relay, and \( L_{T_l} = 4, L_{R_l} = 2, \forall l, L_1 = 4 \) antennas at each node. In Fig. 3 we plot the average objective value and the convergence error versus the number of iterations. It can be observed that PWF has faster convergence speed than GA. It also achieves a higher objective value. The complexity order per iteration of PWF and GA is similar. However, the overall complexity of PWF is lower because it avoids the linear search for the gradient update of \( \Sigma \), and requires less iterations to achieve the same accuracy.

In Fig. 4 we compare the convergence of Algorithm GA, PWF and PWFI for a two-hop MAC relay network with \( L = 4 \) sources, \( n_1 = 1 \) relay and \( L_{T_l} = 2, \forall l, L_1 = 4, L_{R_l} = 4, \forall l \) antennas at each node. The results show that PWFI has the fastest convergence speed. PWFI also has the lowest complexity because it avoids the linear search in the gradient update.

V. SIMULATION RESULTS

Figure 3. Convergence comparison for the B-MAC AF relay network in Fig. 1

Figure 4. Convergence comparison for a two-hop MAC AF relay network
Finally, we verify the convergence of Algorithm PWF3 for a three-hop BC relay network with $n_q = 1, q = 1, 2$ relay at each relay cluster, $L = 4$ destinations and $L_{T_l} = 4, \forall l, L_q = 4, q = 1, 2, L_{R_l} = 2, \forall l$ antennas at each node. Fig. 5 plots the average objective value and objective error, defined as the gap (in logarithmic scale) from the stationary point that the algorithm converges to, versus the number of iterations. Algorithm PWF3 quickly converges to a stationary point of problem (40) with high accuracy.

B. Advantages of the Proposed Local LDM with Duality-based Primal Algorithm

In Fig. 5 and 6 we consider Problem (38) with $\mu_l = 1, \forall l$ (i.e., sum-rate maximization) and illustrate the advantages of the proposed local LDM with duality-based primal algorithms (i.e., Algorithm PWF or PWFI). The following baseline algorithms are compared.

- Baseline 1 (Algorithm 3 in [9]): Algorithm 3 in [9] is based on logarithmic barrier method [10] and is designed for sum-rate maximization under power constraints in two-hop BC relay networks.
- Baseline 2 (logarithmic barrier algorithm): This is an extension of the Algorithm 3 in [9] to B-MAC AF relay networks using the logarithmic barrier method [10].
- Baseline 3: Local LDM with GA as the primal algorithm.

We first evaluate the performance in a two-hop BC relay network with $n_1 = 1$ relay, $L = 4$ destinations and $L_{T_l} = 4, \forall l, L_1 = 4, L_{R_l} = 2, \forall l$ antennas at each node. Assume a sum power constraint at the source: $\sum_{l=1}^{4} \text{Tr}(\Sigma_l) \leq 10$, and a power constraint at the relay: $\text{Tr}(\Sigma_3) \leq P_R$. In Fig. 6 we plot the
average sum-rate, and the CPU time required to achieve the same accuracy, versus the power constraint at relay $P_R$. All algorithms achieve similar sum-rate. However, LLDM with both PWF and PWFI clearly outperform the Algorithm 3 in [9] in terms of CPU time.

In Fig. 7, we evaluate the sum-rate and CPU time performance in the B-MAC AF relay network in Fig. 3 with $Q = 1$ (two-hop), $n_1 = 2$ relays, $L_{1,j} = 2$, $j = 1, 2$ antennas at each relay, and $L_{T_l} = 4$, $L_{R_l} = 2$, $\forall l$ antennas at each source/destination node. Consider individual power constraints at each node:

$$
\sum_{l=1}^{2} \text{Tr} (\Sigma_l) \leq 10, \quad \text{Tr} (\Sigma_3) \leq 10, \quad \text{Tr} (\Sigma^R_l \text{BlockDiag} [I_1, 0_2]) \leq P_R \quad \text{and} \quad \text{Tr} (\Sigma^R_3 \text{BlockDiag} [0_1, I_2]) \leq P_R.
$$

All algorithms achieve similar sum-rate. However, LLDM with PWF requires the least CPU time.

In Fig. 8, we consider Problem (39) with $\hat{W}_l^0 = I, \forall l$, $\hat{W}_q = I, \forall q$ (i.e., sum power minimization), and $\mathcal{I}_l^0 = \mathcal{I}_l^0, \forall l$ for a two-hop MAC relay network with $L = 3$ sources, $n_1 = 1$ relay and $L_{T_l} = 2$, $\forall l$, $L_1 = 4$, $L_{R_l} = 4$, $\forall l$ antennas at each node. The proposed LLDM with PWF is compared to baseline 3 and baseline 4: Algorithm 9 in [9] (joint search based on logarithmic barrier method). We plot the average sum power and the CPU time versus the individual rate constraint ($\mathcal{I}_l^0$). All algorithms achieve similar sum power. However, LLDM with PWF requires the least CPU time.

VI. CONCLUSION

We show that the achievable regions of a multi-hop MIMO B-MAC AF relay network and its dual are the same under single network linear constraint or per-hop linear constraint. Two dual transformations
are provided to calculate the dual input covariance and relay precoding matrices. These results include
the dualities in [11]–[13] as special cases. Furthermore, our proof is simpler and reveals more structural
property of the duality. Based on the established duality structure, we propose efficient algorithms for
MIMO precoder optimization in B-MAC AF relay networks. First, a unified optimization framework is
proposed based on the local Lagrange dual method in [23] so that we only need to focus on designing
a primal algorithm to find a stationary point of the unconstrained inner loop problem. Using duality, we
characterize the polite water-filling (PWF) structure of the input covariance matrices at a stationary point
of the inner loop problem. Then, the duality and PWF are exploited to design efficient primal algorithms.
The proposed local LDM with duality-based primal algorithms has lower computation cost and faster
convergence speed than the conventional step-size based iterative algorithms.

APPENDIX

A. Proof of Theorem 2

It follows from \( \hat{F}_q = F_q^\dagger \) that \( \hat{B}_{q,q'} = B_{q,q'}, \forall q' \geq q \). Hence the equivalent channel of the dual network
is \( \hat{H}_{l,k} = H_{l,k}^\dagger B_{1,q}^\dagger \tilde{H}_k^\dagger, \forall l, k \). By (10), the covariance of the equivalent noise of dual link \( l \)
\( \hat{W}_l^\dagger + \sum_{q=1}^{Q} H_l^\dagger \hat{B}_{1,q}^\dagger \hat{W}_q \hat{B}_{1,q} H_l^\dagger = \hat{W}_l^\dagger + \sum_{q=1}^{Q} H_l^\dagger \hat{B}_{1,q}^\dagger \hat{W}_q \hat{B}_{1,q} H_l^\dagger = \hat{W}_l^\dagger \). Note that
\[
\sum_{q=1}^{Q} \text{Tr} \left( \hat{\Sigma}_q W_q \right) = \sum_{q=1}^{Q} \sum_{l=1}^{L} \text{Tr} \left( \hat{B}_{q,q}^\dagger \hat{H}_q^\dagger \hat{\Sigma}_l^q \hat{H}_q^\dagger \hat{B}_{q,q} W_q \right) + \sum_{q=1}^{Q} \sum_{q'=q}^{Q} \text{Tr} \left( \hat{B}_{q,q'}^\dagger \hat{W}_q \hat{B}_{q,q'} W_q \right)
\]
\[
= \sum_{l=1}^{L} \text{Tr} \left( \hat{\Sigma}_l \sum_{q=1}^{Q} H_q^\dagger B_{q,q} W_q B_{q,q}^\dagger H_q^\dagger \right) + \sum_{q=1}^{Q} \sum_{q'=q}^{Q} \text{Tr} \left( \hat{B}_{q,q'} \hat{W}_q \hat{B}_{q,q'} W_q \right) + \sum_{q=1}^{Q} \sum_{q'=q}^{Q} \text{Tr} \left( \hat{B}_{q,q'}^\dagger \hat{W}_q \hat{B}_{q,q'} W_q \right).
\]
Hence, the linear constraint in (12) can be expressed as \( \sum_{l=1}^{L} \text{Tr} \left( \hat{\Sigma}_l W_l \right) + P_C \leq P_T \). The above proves
the equivalence between the dual B-MAC AF relay network under constraint (12) and the B-MAC IFN
in (27). The equivalence between the B-MAC AF relay network under constraint (7) and the B-MAC
IFN in (24) can be proved similarly.

B. Proof of Theorem 4

Recall that \( t_{l,m}, r_{l,m}, \Psi \) and \( \mathbf{D} \) are obtained using (18), (19), (21) and (22) in Definition 2 with
parameters \( \left\{ \left[ \hat{H}_{l,k} \right], \left[ W_l \right], \left[ \hat{W}_l \right] \right\} \). Since \( \hat{\Sigma}(\mathbf{d}) \) is the covariance transformation of \( \Sigma \) obtained by
Definition 2 with parameters \( \left\{ \left[ \hat{H}_{l,k} \right], \left[ W_l \right], \left[ \hat{W}_l(\mathbf{d}) \right] \right\} \), using the fact that \( t_{l,m}, r_{l,m}, \Psi \) and \( \mathbf{D} \) only
depends on \( \Sigma \) and \( \left\{ \left[ \hat{H}_{l,k} \right], \left[ W_l \right] \right\} \), we have
\[
\hat{\Sigma}_l(\mathbf{d}) = \sum_{m=1}^{M_l} \tilde{q}_{l,m} r_{l,m}^\dagger r_{l,m}^\dagger, l = 1, ..., L,
\]
where \( \{ \bar{q}_{l,m} \} \) is given by

\[
\bar{q} = (D^{-1} - \Psi^T)^{-1} \left( \sum_{q=0}^{Q-1} d_{q+1} \hat{n}_q + \hat{n}_Q \right),
\]

(59)

and \( \hat{n}_q \)'s are defined in (52). Note that

\[
\sum_{l=1}^{L} \text{Tr} \left( \hat{\Sigma}_l (d) W_{q+1}^l \right) = \sum_{l=1}^{L} \sum_{m=1}^{M_l} r_{l,m}^\dagger W_{q+1}^l r_{l,m} \bar{q}_{l,m}
\]

\[
= n_Q^T (D^{-1} - \Psi^T)^{-1} \left( \sum_{q'=0}^{Q-1} d_{q'+1} \hat{n}_{q'} + \hat{n}_Q \right)
\]

(60)

where (60-a) follows from (31) and (59), and (60-b) follows from the definition of \( a_{q,q'} \)'s in (30). Hence the first linear constraint in (35) can be expressed as

\[
\sum_{q'=0}^{Q-1} d_{q'+1} a_{Q,q'} + a_{Q,Q} = P_{Q}^{tx}.
\]

(61)

Similarly, \( \forall q = 1, \ldots, Q \),

\[
\text{Tr} \left( \hat{\Sigma}_q^R (d) W_q \right)
\]

\[
= \sum_{l=1}^{L} \sum_{m=1}^{M_l} r_{l,m}^\dagger H_{q,l}^T B_{q,l} W_q B_{q,l}^T H_{q,l} r_{l,m} \bar{q}_{l,m} + \sum_{q'=0}^{Q-1} d_{q'+1} \text{Tr} \left( B_{q,q'}^T \hat{W}_q B_{q,q'} W_q \right)
\]

\[
= n_{q-1}^T (D^{-1} - \Psi^T)^{-1} \left( \sum_{q'=0}^{Q-1} d_{q'+1} \hat{n}_{q'} + \hat{n}_Q \right) + \sum_{q'=0}^{Q-1} d_{q'+1} b_{q,q'}
\]

(62)

where (62) follows from (30) and (26), and thus the other linear constraints in (35) can be expressed as

\[
\sum_{q'=0}^{Q-1} d_{q'+1} a_{q-1,q'} + a_{q-1,Q} = \sum_{q'=0}^{Q-1} d_{q'+1} b_{q,q'},
\]

(63)

Now replace \( P_{q}^{tx} \)'s in (61) and (63) with \( \lambda P_{q}^{tx} \)'s, we obtain another set of equations

\[
\sum_{q'=0}^{Q-1} d_{q'+1} a_{Q,q'} + a_{Q,Q} = \lambda P_{Q}^{tx},
\]

(64)

\[
\sum_{q'=0}^{Q-1} d_{q'+1} a_{q-1,q'} + a_{q-1,Q} + \sum_{q'=0}^{Q-1} d_{q'+1} b_{q,q'} = d_{q} \lambda P_{q-1}^{tx}, \quad q = 1, \ldots, Q.
\]
It can be verified that the equations in (64) forms the eigensystem $\mathbf{A}d = \lambda d$ in (34), where $\mathbf{A} \in \mathbb{R}_+^{(Q+1)\times Q+1}$ is defined in (33), $d = [d_1, ..., d_Q, 1]^T$. Note that $a_{Q,0} > 0$ due to nonsingularity of $\mathbf{W}_q^L$'s and $\mathbf{W}_q^T$'s. Furthermore, since $\mathbf{W}_q$'s and $\mathbf{W}_q$'s are assumed to be non-singular and $\mathbf{F}_q \neq 0$, $q = 1, ..., Q$, we have $b_{q,q} > 0$, $q = 1, ..., Q$. Using these facts, the following lemma can be proved.

**Lemma 1.** The following is true for $\mathbf{A}$ defined in (33). 1) $\mathbf{A} \succeq 0$ and $\mathbf{A} \neq 0$. 2) For any $\lambda$ and $d$ satisfying $\mathbf{A}d = \lambda d$, if any element of $d$ is zero, then $d = 0$.

It follows from Lemma 1 and the Perron–Frobenius theorem [26, Chp. 8] that the maximum eigenvalue $\lambda_{\text{max}}$ and the associated dominant eigenvector $\tilde{d}$ of $\mathbf{A}$ satisfies $\lambda_{\text{max}} > 0$ and $\tilde{d} > 0$. Then we can always obtain a scaled eigenvector $\tilde{d} = \tilde{d}'/\tilde{d}_{Q+1} = [\tilde{d}_1, ..., \tilde{d}_Q, 1]^T$ that satisfies $\mathbf{A}\tilde{d} = \lambda_{\text{max}}\tilde{d}$. On the other hand, it follows from (29) that $\lambda_{\text{max}}$ is equal to 1, which indicates that $\mathbf{F}_q(\tilde{d})$ and $\tilde{\mathbf{F}}$ satisfies (35).

**C. Proof of Theorem 4**

By Corollary 2, $\hat{\Sigma}$ satisfies the PWF structure for fixed $\tilde{\mathbf{F}}$ and the KKT condition in (44) with $\mathbf{F} = \tilde{\mathbf{F}}$. To prove that $\hat{\Sigma}, \tilde{\mathbf{F}}$ is a stationary point, we only need to further show that $\hat{\Sigma}, \tilde{\mathbf{F}}$ satisfies

$$
\nabla_{\mathbf{F}_q(j)}L_{\lambda}(\hat{\Sigma}, \tilde{\mathbf{F}}) = 0, \forall q = 1, ..., Q, j = 1, ..., n_q. \tag{65}
$$

Since $\hat{\Sigma}, \tilde{\mathbf{F}}$ is a stationary point of the dual problem (48), we have $\nabla_{\mathbf{F}_q(j)}L_{\lambda}(\hat{\Sigma}, \tilde{\mathbf{F}}) = 0, \forall q, j$. In the following, we prove that $\nabla_{\mathbf{F}_q}L_{\lambda}(\hat{\Sigma}, \tilde{\mathbf{F}}) = (\nabla_{\mathbf{F}_q}L_{\lambda}(\hat{\Sigma}, \tilde{\mathbf{F}}))^*, \forall q$, from which (65) follows immediately.

Note that for fixed $\mathbf{F}_q', \forall q' \neq q$ and $\tilde{\mathbf{F}}_q', \forall q' \neq q$, problem (40) and (48) can be equivalent to the inner loop problem and its dual for a two-hop B-MAC AF relay network. Hence, without loss of generality, we only need to consider the two-hop case (i.e., $Q = 1$) and prove that $\nabla_{\mathbf{F}}L_{\lambda}(\hat{\Sigma}, \tilde{\mathbf{F}}) = (\nabla_{\mathbf{F}}L_{\lambda}(\hat{\Sigma}, \tilde{\mathbf{F}}))^*$. Using (45) in Corollary 2 and the expression of $\nabla_{\mathbf{F}}L_{\lambda}(\hat{\Sigma}, \tilde{\mathbf{F}})$ in (51), it can be shown that

$$
\nabla_{\mathbf{F}}L_{\lambda}(\hat{\Sigma}, \tilde{\mathbf{F}}) = -2\sum_{l=1}^{L} \sum_{k=1}^{L} \phi_{l,k} \tilde{S}_l \tilde{F}_l \tilde{S}_k - 2 \sum_{l=1}^{L} \tilde{S}_l \tilde{F}_l \tilde{W}_l - 2 \sum_{l=1}^{L} \tilde{S}_l \tilde{F}_l \tilde{S}_l
$$

$$
+ 2\sum_{l=1}^{L} w_l H_{1}^{l\dagger} \tilde{\Omega}_l^{-1} H_{1}^{l\dagger} \hat{\Sigma}_l H_{0}^{l\dagger} - 2 \sum_{l=1}^{L} \tilde{W}_l \tilde{F}_l \tilde{S}_l, \tag{66}
$$

where $\tilde{S}_k = H_{0}^{k\dagger} \hat{\Sigma}_k H_{0}^{k\dagger}$ and $\tilde{S}_l = H_{1}^{l\dagger} \hat{\Sigma}_l H_{1}^{l\dagger}$. Similar,

$$
\nabla_{\tilde{\mathbf{F}}}L_{\lambda}(\hat{\Sigma}, \tilde{\mathbf{F}}) = -2\sum_{l=1}^{L} \sum_{k=1}^{L} \phi_{l,k} \tilde{S}_l \tilde{F}_l \tilde{S}_k - 2 \sum_{l=1}^{L} \tilde{S}_l \tilde{F}_l \tilde{W}_l - 2 \sum_{l=1}^{L} \tilde{S}_l \tilde{F}_l \tilde{S}_l
$$

$$
+ 2\sum_{l=1}^{L} w_l H_{1}^{l\dagger} \tilde{\Omega}_l^{-1} H_{1}^{l\dagger} \hat{\Sigma}_l H_{1}^{l\dagger} - 2 \sum_{l=1}^{L} \tilde{W}_l \tilde{F}_l \tilde{S}_l. \tag{67}
$$
It follows from (46) and (47) in Corollary 2 that

\[
H_l^\dagger \Omega_l^{-1} H_l^\dagger P H_0^\dagger \Sigma_l H_l^\dagger = H_l^\dagger \Sigma_l H_l^\dagger P H_0^\dagger \Omega_l^{-1} H_l^\dagger.
\] (68)

Combining (66-68) and \( \tilde{F} = \tilde{F}^\dagger \), we have \( \nabla F L_\lambda \left( \tilde{\Sigma}, \tilde{F} \right) = \left( \nabla \tilde{F} L_\lambda \left( \tilde{\Sigma}, \tilde{F} \right) \right)^* \). This completes the proof.

REFERENCES


