Space-time Beamforming for Multiuser Wireless Relay Networks

A.H Phan, H.D. Tuan and H.H. Kha
School of Electrical Engineering and Telecommunications, The University of New South Wales, Sydney, Australia; Email: {z3261071,h.d.tuan,h.k.ha}@unsw.edu.au

Abstract—The paper is concerned with a multiuser communication network, which is assisted by multiple relays. It has been observed through our previous related works that the conventional simultaneous beamforming at parallel amply-and-forward (AF) relays is not quite effective and often infeasible to target practically desirable signal-to-interference-and-noise ratio (SINR) at the destinations. To overcome this shortage, we propose the time-division for multiple-user transmission to the relays so the later can perform beamforming on signals received from the individuals and then parallelly forward its combinations at once to the destinations. The optimal beamforming problem is a non-convex quadratically constrained optimization, which is globally solved by our tailored algorithm of nonsmooth optimization. Its found global optimal solutions are shown very effective and outperform other possible multi-user relay beamformings.

I. INTRODUCTION

Relay-assisted communication has been considered as a very prominent research area [2], [3]. In such communication, relays are used to create spatial diversity between the transceivers and thus enhance the quality of service significantly. Among the relay protocols, the amplify-and-forward relay (AFR) is most popular due to its low complexity of implementation. Relay beamforming is a simple but very effective technique to substantially increase the relay capacity for collaborative communication [2]. Many relay beamforming schemes have been considered for the single user case only, i.e. there only are single source and single destination, which communicate with assist of multiple AFRs (see e.g. [2] and references therein). All relays perform beamforming on the received signal from the source and then forward the beamformed signals at the same time to the destination. Thus, an appropriate optimization is to minimize the beamforming power under the single targeted signal-to-noise (SNR) at the destination [2]. Only very recently, multi-user relay beamforming has been considered in [4], [8] ( [11] also considered multi-user relay as well but in term of designing a precoder matrix of full size instead of beamforming for the signal received at the relays), whereas multiple pairs of source and destination communicate through the same AFRs. All resources send their signals to the relays at the same time, so the signals received at the relays are noisy combinations of all ones sent from the resources. All relays perform beamforming and then broadcast the beamformed signals to all destinations at the same time. Thus at each received destination there is so-called multi-user interference inherent from the signals originated from other sources that are not in pair with the destination.

Therefore, the relay beamforming must be based on the minimization of power beamforming subject to the multiple targeted signal-to-interference-and-noise ratios (SINRs). However, the simulation [4], [8] shows ineffectiveness of this multi-user beamforming. Namely, the required beamforming power has been substantially increased for the case of 3 source-destination pairs with 10 relays. For 4 source-destination pair, there is no feasible beamforming at all, even when SINR constraints are set so moderate 8dB and up to 10 relays are employed. By close scrutiny at the mathematical modeling of these multi-user relay beamforming, one can see that there are lots of conflicting constraints caused by SINRs that restrict the feasibility of beamforming solution. In other words, such parallel AFRs are not feasible enough for multi-user relays.

From computation view point, the above mentioned multi-user relay beamforming optimization problem is nonconvex quadratically constrained quadratic program (NQCQP) [4] The conventional method for solving this NQCQP is semidefinite programming (SDP) relaxation in tandem with randomization. Our previous works [5], [6] particularly analyzed inconsistent performance of this conventional method for NQCQP and proposed new nonsmooth optimization techniques for effective computation of its global optimal solutions.

The purpose of the present paper is two-fold:

- To propose a time-division for the multi-user sources in Section II that allows ARFs to perform beamforming individually for each signal from the sources. After beamforming all these signals from the sources, ARFs just combine them for parallelly broadcasting to all destinations. By making the SINRs constraints sparsely connected in its optimization formulation, this scheme (called by space-time amplify-and-forward (STAF)) greatly increases the feasibility of multi-user beamforming solutions while keeping the broadcasting rate of ARFs the same as for single user beamforming.
- To propose a nonsmooth optimization algorithm tailored for the global optimal solution of NQCQP corresponding to STAF scheme. The numerical simulation results in Section IV show the viability of STAF as well as of this its solution algorithm.

Notations: Matrices and column vectors are denoted by boldfaced uppercase and lowercase characters, respectively. The notation $A \succeq 0$ means $A$ is a (Hermitian) positive semi-definite matrix. We denote $\langle A \rangle = \text{trace}(A)$, $\langle A, B \rangle = \text{trace}(AB)$.
trace($A^H B$), and $\langle a, b \rangle = a^H b$. $\lambda_{max}(X)$ stands for the maximal eigenvalue of $X$. $\odot$ denotes the component-wise Hadamard product. For a vector $x = (x_1, ..., x_m)^T$, either diag$x$ or diag$[x_i]_{i=1}^m$ is a diagonal matrix with diagonal entries $x_i$.

II. SYSTEM MODELLING

Consider a wireless network as shown by Fig. 1, where $M$ sources $S_i, i = 1, 2, ..., M$ (users) communicate in pair to $M$ destinations (receivers) $D_i, i = 1, 2, ..., M$ though $N$ relays $R_j, j = 1, 2, ..., N$. All sources, relays and destinations are supposed to be equipped with single antennas and in addition the relays operate in half-duplex mode.

Let $s = [s_1, s_2, ..., s_M]^T$ be the independent signal symbols that $M$ sources are supposed to communicate to the destinations. Each of $s_i$ is assumed as zero mean with the same variance $\sigma_s = E[|s_i|^2]$. Let $h_i = [h_{i1}, h_{i2}, ..., h_{iM}]^T$ be the channel vector from the sources to the relay $i$ and $l_i = [l_{i1}, l_{i2}, ..., l_{iN}]^T$ denote the channel vector from $N$ relays to the destination $i$.

The transmission of the sources to the relays is arranged to be time orthogonal, i.e. the sources use different time slots for their transmissions to the relays. Without loss of generality, let source $S_1$ transmits its signal $s_1$ at the first time slot, source $S_2$ transmit its signal $s_2$ at the second time slot and so all. As a result, there are $M$ time slots for signal transmission from the sources to the relays. At every time slot $i$, $N$ relays perform beamforming by assigning weight $x_{ji}$, $j = 1, 2, ..., N$ on the received signal. For $x_i := (x_{i1}, x_{i2}, ..., x_{iN})^T$ the signal sent by the relays after beamforming and combining all the beamformed received signals is presented by

$$t = \sum_{i=1}^M \text{diag}[x_i](h_is_i + n_i).$$

Thus, the received signal at destination $j$ is

$$d_j = l_j^T \left[ \sum_{i=1}^M \text{diag}[x_i](h_is_i + n_i) \right] + \bar{n}_j,$$

where $n_i$ and $\bar{n}_j$ are additive Gaussian white noise (AWGN) components.

Since only component $s_j$ is of interest in destination $j$, the signal power at destination $j$ is

$$P_j = \sigma_j^2 |\langle c_{jj}, x_j \rangle|^2 \quad \text{with} \quad c_{jj} = l_j \odot h_j,$$

while the interference and noise power at destination $j$ is

$$INT_j = \sigma_j^2 \sum_{i \neq j} |\langle c_{ji}, x_i \rangle|^2 + \sigma_j^2 \sum_{i=1}^M \langle \text{diag}[l_{ij}]^2 N_j, x_i x_i^H \rangle + \sigma_d^2$$

where $\sigma_j^2$ and $\sigma_d^2$ are white noise variances at relays and destinations respectively, and $c_{ji} = l_j \odot h_i$.

The total beamforming power is

$$PT = \mathbb{E}\{ \langle t, t \rangle \} = \sum_{k=1}^M \langle \text{diag}[\sigma_j^2 |h_{ik}|^2 + \sigma_d^2 N_j], x_k x_k^H \rangle. \quad (1)$$

The beamforming design is to minimize the beamforming power $PT$ under constraints of SINR at each destination being above some thresholds. Thus, the optimization formulation is following NQCQP

$$\min_{\mathbf{x}_k \in \mathbb{C}^N, k = 1, 2, ..., M} \sum_{k=1}^M \langle \mathbf{D}_k, \mathbf{x}_k \mathbf{x}_k^H \rangle \quad \text{s.t.} \quad \sigma_j^2 |\langle c_{ji}, x_i \rangle|^2 \geq \alpha_i, \quad (2a)$$

$$\sigma_j^2 \sum_{j \neq i} |\langle c_{ji}, x_i \rangle|^2 + \sum_{j=1}^M \langle \mathbf{G}_j, x_i x_j^H \rangle + \sigma_d^2 \geq \alpha_i, \quad i = 1, 2, ..., M$$

where $\mathbf{D}_k = \text{diag}[\sigma_j^2 |h_{1k}|^2 + \sigma_j^2, ..., \sigma_j^2 |h_{NK}|^2 + \sigma_j^2]$ and $\mathbf{G}_j = \sigma_j^2 \text{diag}[|l_{1j}|^2, ..., |l_{NJ}|^2]$.

For convenience, set $\bar{x}_k := \mathbf{D}_k^{1/2} \mathbf{x}_k$, $\bar{c}_{ji} := \sigma_j \mathbf{D}_j^{-1/2} c_{ji}$ and $\bar{G}_j := \mathbf{D}_j^{-1} \mathbf{G}_j$ and then equivalently express (2) by

$$\min_{\mathbf{x}_k \in \mathbb{C}^N, k = 1, 2, ..., M} \sum_{k=1}^M \langle \bar{c}_{ki}, \bar{x}_k \rangle \quad \text{s.t.} \quad |\langle \bar{c}_{ki}, \bar{x}_k \rangle|^2 \geq \alpha_i, \quad (3a)$$

$$\sum_{j \neq i} |\langle \bar{c}_{ji}, \bar{x}_j \rangle|^2 + \sum_{j=1}^M \langle \bar{G}_j, \bar{x}_i \bar{x}_j^H \rangle + \sigma_d^2 \geq \alpha_i, \quad i = 1, 2, ..., M$$

One can see that SINR constraints (3b) look similar to that in [4], [8]. However, (3b) involves $M$ vector variable $\mathbf{x}_k, k = 1, 2, ..., M$ of dimension $N$ while that in [4], [8] involve only single vector variable $\mathbf{x}$ if the same dimension $M$. Thus, in comparison with [4], [8], the formulation (3b) offers much feasibility of solution sets.

For $\mathbf{C}_{ji} := \bar{c}_{ji} \bar{c}_{ji}^H$, that (3) can be also represented by the
The following rank-one constrained optimization
\[ \min_{\mathbf{X}_k \in \mathbb{C}^{N \times N}, k = 1, 2, \ldots, M} \sum_{k=1}^{M} \langle \mathbf{X}_k \rangle \quad \text{s.t.} \quad (4a) \]
\[ \sum_{j \neq i}^{M} \langle \mathbf{G}_{ji}^H \mathbf{X}_j \rangle + \sum_{j=1}^{M} \langle \mathbf{X}_j^H \mathbf{G}_j \rangle + \sigma^2_i \geq \alpha_i \quad (4b) \]
\[ \mathbf{X}_k \geq 0, k = 1, 2, \ldots, M \quad (4c) \]
\[ \mathbf{X}_k = \tilde{a}_k \tilde{a}_k^H, i, k = 1, 2, \ldots, M \quad (4d) \]

As mentioned, the conventional approach (see e.g. [1]) is to relax (4) by dropping only nonconvex rank-one constraint (4d) and then employ randomization to generated feasible solutions from the optimal solution of the relaxed program. Some analysis on this conventional approach have been made in [5], [6], which particularly show that as far as the optimal solution of the relaxed program is not rank-one, such randomization is too narrow and hardly bring the optimal solution of the original rank-one constrained optimization. A novel approach for locating the global optimal solution of (3) is proposed in the next section.

III. NONSMOOTH OPTIMIZATION APPROACH

In this section, we use the nonsmooth optimization method of [6], [7], which has been shown very efficient in locating the global optimization solution of nonconvex optimization problems like (3). First, the non-convex constraints (4d) are equivalently expressed by the following nonsmooth but still continuous constraint
\[ \sum_{k=1}^{M} (\langle \mathbf{X}_k \rangle - \lambda_{\max}(\mathbf{X}_k)) \leq 0. \quad (5) \]

Therefore (4) actually is the following convex program with an additional reverse convex constraint [9]
\[ \min_{0 \leq \mathbf{X}_k \in \mathbb{C}^{N \times N}, k = 1, \ldots, M} \sum_{k=1}^{M} \langle \mathbf{X}_k \rangle, \text{ s.t.} (4b), (5) \quad (6) \]

For iterative purpose, (6) is converted to the following, which is a concave program [9]
\[ \min_{0 \leq \mathbf{X}_k \in \mathbb{C}^{N \times N}} \sum_{k=1}^{M} \langle \mathbf{X}_k \rangle + \mu \left[ \langle \mathbf{X}_k \rangle - \lambda_{\max}(\mathbf{X}_k) \right] \quad (4b) \quad (7) \]

where \( \mu > 0 \) is a penalty parameter.

After initializing from a feasible point \( \mathbf{X}_k^{(\kappa)}, k = 1, \ldots, M \) (which satisfies (4b)) whose maximum eigenvalue is \( \lambda(\mathbf{X}_k^{(\kappa)}) \) and with the corresponding normalized eigenvector \( \mathbf{x}_k^{(\kappa)} \), the following SDP program provides a feasible solution \( \mathbf{X}_k^{(\kappa+1)} \) that is better than \( \mathbf{X}_k^{(\kappa)} \) of (7):
\[ \min_{0 \leq \mathbf{X}_k \in \mathbb{C}^{N \times N}} \sum_{k=1}^{M} \langle \mathbf{X}_k \rangle + \mu \left[ \langle \mathbf{X}_k \rangle - \lambda_{\max}(\mathbf{X}_k^{(\kappa)}) \right] - \langle \mathbf{x}_k^{(\kappa)} \mathbf{x}_k^{(\kappa)H} \rangle \mathbf{X}_k - \mathbf{X}_k^{(\kappa)} \quad (4b) \]

Algorithm 1 summarizes a two-stage penalty function method that can be used to solve the target optimization problem (3). In the initialization stage, a value of \( \mu \) is determined together with rank-one solutions \( \mathbf{X}_k^{(0)}, k = 1, \ldots, M \). In the optimization stage, \( \mu \) and \( \mathbf{X}_k^{(0)}, k = 1, \ldots, M \) can serve as the input to find out the optimal solutions. This algorithm allow the solutions to converge at the final result iteratively.

Algorithm 1 PenFun algorithm to compute optimal solution \( \mathbf{X}_k^{\text{opt}}, k = 1, \ldots, M \) of (4)

% Initialization stage:
% Initial Step:
% Initialize proper \( \mu \) and a solution \( \mathbf{X}_k^{(0)}, k = 1, \ldots, M \) to satisfy (4b).
% Set \( \kappa := 0 \)
% % Step \( \kappa \):
% Solve (8) to obtain its optimal solutions \( \mathbf{X}_k^{(\kappa+1)}, k = 1, \ldots, M \).
% if \( \langle \mathbf{X}_k^{(\kappa+1)} \rangle \approx \lambda_{\max}(\mathbf{X}_k^{(\kappa+1)}), k = 1, \ldots, M \) (i.e., rank-one solution found) then
% Reset \( \mathbf{X}_k^{(0)} := \mathbf{X}_k^{(\kappa+1)}, k = 1, \ldots, M \).
% Terminate, and output \( \mu \) and \( \mathbf{X}_k^{(0)}, k = 1, \ldots, M \).
% else if \( \mathbf{X}_k^{(\kappa+1)} = \mathbf{X}_k^{(\kappa)}, k = 1, \ldots, M \) (i.e., no improved solution found, no rank-one result) then
% Reset \( \mu := 2 \mu \) and return to the Initial step.
% else
% Reset \( \kappa := \kappa + 1 \) and \( \mathbf{X}_k^{(\kappa)} := \mathbf{X}_k^{(\kappa+1)} \) for the next iteration.
% end if
% % Optimization stage:
% Set \( \kappa := 0 \). Solve (8) to obtain its optimal solution \( \mathbf{X}_k^{(\kappa+1)}, k = 1, \ldots, M \).
% if \( \langle \mathbf{X}_k^{(\kappa+1)} \rangle \approx \langle \mathbf{X}_k^{(\kappa)} \rangle, k = 1, \ldots, M \) (i.e., convergence) then
% Terminate, and output \( \mathbf{X}_k^{\text{opt}} := \mathbf{X}_k^{(\kappa+1)}, k = 1, \ldots, M \).
% else
% Reset \( \kappa := \kappa + 1 \) and \( \mathbf{X}_k^{(\kappa)} := \mathbf{X}_k^{(\kappa+1)}, k = 1, \ldots, M \).
% Continue to the next iteration
% end if
% Output the final solution \( \mathbf{X}_k^{\text{opt}}, k = 1, \ldots, M \).

IV. NUMERICAL RESULTS

This section presents numerical results to verify the performance of our proposed non-smooth optimization approach. In each simulation scenario, the frequency-flat channels are generated according to Rayleigh distribution with normalized channel gain. The final results are then obtained by averaging over 1,000 Monte-Carlo simulation runs.

Several system configurations are simulated. First, the scenario of \( M = 3 \) sources-destination pairs with \( N = 10 \) AFs. Then \( M \) is increased to \( M = 4 \) and \( M = 5 \) but \( N = 10 \) is held fixed.

Two other schemes are also simulated for the performance comparison with STAF. A multiuser AF (MAF) scheme [4],
[8] where the transmission of all users are conducted simultaneously is used when the number of users is \( M = 3 \). MAF model may shows worse performance in term of total beamforming power but the whole transmissions consume only two time slots. Alternatively, another scheme called single-user AF (SAF) model is where each user pair is assigned to transmit and receive independently to the other pairs. This model may show the most effective performance in term of beamforming power. Nevertheless, SAF scheme consumes up to \( 2M \) time slots for the whole transmission. The proposed STAF uses \( M + 1 \) independent time slots for the whole transmission.

As can be observed from Figure 2, a sacrifice in transmission time always result in better beamforming power consumption. However, although using least number of time slots, MAF indicates very poor performance compared with the others. STAF’s performance is slightly less effective than SAF in power consumption but the number of time slots reduced in the transmission is remarkable.

As the number of source-destination pairs increases to \( M = 4 \) and \( M = 5 \) the systems’ performances are compared in Figure 3. MAF performance cannot be included there because of its infeasibility. Although SAF slightly outperforms STAF model, the gaps are unremarkable if the number of used time slots are taken into account.

V. CONCLUSION

We have proposed an appropriate time-division for multi-user sources to facility feasible relay beamforming for multi-user relay wireless network. With the global optimal beamforming solution found by a novel nonsmooth optimization algorithm proposed in the paper, the corresponding beamforming is able to balance the trade-off between the transmission time consumption and the detection performance measured by SINRs at the multi-user destinations.

REFERENCES