This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2014.2324588, IEEE Transactions on Wireless Communications

Joint Optimization of Cooperative Beamforming and Relay Assignment in Multi-User Wireless Relay Networks

Enlong Che[1], H. D. Tuan[1] and H. H. Nguyen[2]

***Abstract*—This paper considers joint optimization of cooperative beamforming and relay assignment for multi-user multi-relay wireless networks to maximize the minimum of the received signal-to-interference-plus-noise ratios (SINR). Separated continuous optimization of beamforming and binary optimization of relay assignment already pose very challenging programs. Certainly, their joint optimization, which involves nonconvex objectives and coupled constraints in continuous and binary variables, is among the most challenging optimization problems. Even the conventional relaxation of binary constraints by continuous box constraints is still computationally intractable because the relaxed program is still highly nonconvex. However, it is shown in this paper that the joint programs fit well in the d.c (difference of two convex functions/sets) optimization framework. Efficient optimization algorithms are then developed for both cases of orthogonal and nonorthogonal transmission by multiple users. Simulation results show that the jointly optimized beamforming and relay assignment not only save transmission bandwidth but can also maintain well the network SINRs.**

***Index Terms*—Cooperative beamforming, power allocation, relay selection, relay assignment, wireless relay networks, d.c. programming.**

I. INTRODUCTION

Multi-user relay-assisted wireless communication is considered as a promising architecture for future-generation multihop cellular networks [1]–[4]. Recent research efforts have focused on distributed relay processing methods and techniques in order to enhance communication coverage, data rates and diversified Quality-of-Service (QoS) of multi-user communication networks. Among these research efforts, it has been recognized that relay assignment and beamforming are very effective techniques, especially in the context of amplify-and-forward (AF) protocol, in which the role of relays is to amplify the signals received from the sources and then forward the results to destinations.

To date, the design problems concerning relay assignment have been examined mainly for single-user scenarios and with a focus on the outage probability analysis of the best relay selection (see e.g. [5]–[10] and references therein). On the other hand, for a multi-relay network supporting a single user, the phase of the optimal beamforming at each relay can be predetermined and the relay beamforming thus reduces to merely power allocation problems (see e.g. [11]–[15]). There appears no efficient procedure for determining jointly-optimized relay assignment and power allocation. For example, reference [16] considered power optimization for each possible subset of relays while reference [4] considered power optimization on a subset of relays, which is heuristically chosen by relaxation-based rounding.

Multi-user beamforming in a multi-relay network is a much harder problem and it has already been thoroughly studied in recent years [17]–[24]. In particular, joint power allocation and relay assignment problems in *orthogonal* multi-user scenarios have recently been considered in [25]–[29], either under a game-theoretical framework or by adopting a maximum-ratio-combining receiver. By relaxing the binary constraints $\{0,1\}$ of relay assignment to a box constraint on $[0,1]$, the optimization problems in these papers become convex programs. So again, the optimization approach in these papers is to first conduct relay assignment by relaxation-based rounding and then re-optimize power for the assigned relays.

From the above discussion and to the authors' best knowledge, the jointly-optimal designs of the beamforming and relay selection have not been adequately addressed. Apparently, the main reason is the hardness of the joint optimization. In fact, finding separately-optimized beamforming and relay selection in multi-user multi-relay networks is already challenging. Specifically, the binary program of relay assignment is a very hard combinatoric one. The only known approach is to relax the binary constraints to box constraints, which would still lead to a highly nonconvex program. The relay beamforming design problems that consider the practical individual power constraints at the relays are also highly nonconvex optimization and they were recently solved in [21]–[23], [30]. Certainly, the joint optimization of beamforming and relay selection to maximize the minimum of the received SINRs constitutes the hardest nonlinear mixed binary program. This is because such a program involves nonconvex objective and coupled nonconvex constraints in continuous variables of beamforming and binary variables of assignment.

The present paper appears to be the first attempt to address this joint SINR optimization in a systematic way, which is also practical and efficient. Similar to our past [31]–[33] and recent [21]–[23], [34], [35] developments we use d.c. (difference of two convex functions/sets) programming [36] as the tool for finding the solutions. The main reason for using this tool comes from its universality [36, Chapter 3]: almost all NP-hard optimization programs can be theoretically represented by d.c programs, which are to minimize d.c. functions over d.c. con-

[1]Faculty of Engineering and Information Technology, University of Technology, Sydney, NSW 2007, Australia. Email: enlong.che@student.uts.edu.au, tuan.hoang@uts.edu.au

[2]Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, Canada; Email: ha.nguyen@usask.ca.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2014.2324588, IEEE Transactions on Wireless Communications

straints. Although these d.c. programs are still very hard, they motivate exploring hidden convex structures of the problem at hand in order to obtain the constructive solutions [36], [37]. It has been recognized for a long time [36] that the d.c. structure of binary constraints has yet to be explored as the binary constraints seem to be more suitable in a d.m. (difference of two monotonic functions/sets) setting [38], [39]. Moreover, different from a binary quadratic problem [40], which has been successfully solved by d.c. optimization, the constructive d.c. representations of the mixed-binary programs considered in the present paper are not available. Through elegant variable changes and effective approximations, our contribution is to reformulate these problems as tractable d.c. programs, which become solvable by d.c. iterations of polynomial complexity and also re-optimization processes. Simulation results show that the proposed computational procedures are capable of locating approximation solutions that are very close to the global optimal solutions. This is evidenced by the fact that the approximation solutions achieve SINR performances that are very close to their upper bounds.

The paper is structured as follows. Section II provides a fundamental background on a class of mixed binary optimization and d.c. programming for finding its solution. Its application to the joint program of beamforming and relay assignment for the case of orthogonal transmission of the users is presented in Section III. Section IV is devoted to the problem of jointly optimized beamforming and relay assignment in nonorthogonal transmission of the users. Simulations are presented in Section V. Section VI concludes the paper.

Needless to say, similar to the majority of previous works on cooperative beamforming and/or precoding (see e.g. [17]–[29], [41], [42]) the perfect channel state information of both communication links connected to a relay is assumed to be available at the relay. The reader is also referred to [43] and references therein for channel estimation techniques in relay networks, [30] for the robustness of cooperative beamforming against channel estimation error, and [44]–[46] for strategies to reduce the overall delay incurred in relaying signals to multiple receivers.

Notation. Denote $\langle a,b\rangle = a^H b$ for $a=(a_1,\ldots,a_N)^T\in\mathbb{C}^N$ and $b=(b_1,\ldots,b_N)\in\mathbb{C}^N$, so $||a||^2=\langle a,a\rangle=\sum_{i=1}^N |a_i|^2$. The notation $a\odot b$ means the element-wise Hadamard product of two vectors a and b, which is obviously commutative, i.e., $a\odot b\odot c$ does not depend on the order of a,b and c in the product. It is also obvious that $\langle a, b\odot c\rangle = \langle a\odot b, c\rangle$. Upper letters are used to represent matrices and variables are typed boldfaced. Denote $[A]_b\in\{0,1\}^{N\times M}$ as the rounded binary values from $A=[a_{ij}]_{i=1,2\ldots,N;j=1,2,\ldots,M}\in[0,1]^{N\times M}$, so $[A]_b=[[a_{i,j}]_b]_{i=1,2,\ldots,N;j=1,2,\ldots,M}$. For $n\ge m$, define $C(n,m)=n!/m!(n-m)!$.

II. Mathematical foundation: Mixed binary optimization by canonical d.c. programming

It will be seen in next two sections that the joint programs of cooperative beamforming and relay assignments in both scenarios of orthogonal and nonorthogonal transmissions by users can be recast to the following mixed binary program:

$$\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{K}} \quad \mathcal{F}(\boldsymbol{z}) := f(\boldsymbol{z}) - g(\boldsymbol{z}) \ : \ \boldsymbol{x}\in\{0,1\}^N, \tag{1}$$

where $\boldsymbol{z}\in\mathcal{Z}$ is the continuous beamforming variable with $\mathcal{Z}$ being $\mathbb{R}^q$ or $\mathbb{C}^q$, $\boldsymbol{x}$ is the binary link variable, $\mathcal{K}$ is a compact and convex set, and f and g are convex functions. The objective function $\mathcal{F}(\boldsymbol{z})$ in (1) depends on continuous beamforming variable $\boldsymbol{z}$ only.

The function $\mathcal{F}(\boldsymbol{z})$ as in (1) is called a d.c. function [36]. Unlike the approach of [47], which attempts to locate the global optimal solution of program (1) by combining d.c. iterations (DCIs) with a customized branch-and-bound technique of high computational complexity, here we shall follow the ideas of [31]–[34], [37] to obtain equivalent d.c. decompositions that make DCIs alone efficient.

The first step is to express binary constraint $\boldsymbol{x}\in\{0,1\}^N$ in (1) by a d.c. constraint.

Proposition 1: Under the definitions

$$D := [0,1]^N, \tag{2a}$$

$$C := \left\{\boldsymbol{x}\in R^N \ : \ \sum_{n=1}^N \boldsymbol{x}_n^2 - \sum_{n=1}^N \boldsymbol{x}_n < 0\right\}, \tag{2b}$$

the binary set $\{0,1\}^N$ is the difference of two convex sets D and C, i.e., $\{0,1\}^N = D\setminus C$.

Proof: It is obvious that $\{0,1\}^N\subset D\setminus C$. Also $\boldsymbol{x}\in\{0,1\}^N$ is equivalent to

$$\boldsymbol{x}_n - \boldsymbol{x}_n^2 = 0, \ n=1,2,\ldots,N. \tag{3}$$

On the other hand, $\boldsymbol{x}_n-\boldsymbol{x}_n^2\ge 0$ for $\boldsymbol{x}\in D$ and $\sum_{n=1}^N \boldsymbol{x}_n^2 - \sum_{n=1}^N \boldsymbol{x}_n \ge 0$ $\boldsymbol{x}\notin C$, so each $\boldsymbol{x}\in D\setminus C$ is feasible to constraint (3), i.e., $D\setminus C\subset\{0,1\}^N$. □

For convex sets $\mathcal{D}:=\mathcal{K}\cap(\mathcal{Z},D)$ and $\mathcal{C}:=\mathcal{K}\cap(\mathcal{Z},C)$ it is seen that (1) can be compactly rewritten as

$$\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}\setminus\mathcal{C}} f(\boldsymbol{z}) - g(\boldsymbol{z}), \tag{4}$$

which is minimization of a d.c. function over a d.c. set, so (1)/(5) belongs to the class of d.c. programming [36]. Since the feasibility set of (4) is obviously disconnected because of the binary values of the variable $\boldsymbol{x}$, there should be no path-following procedure for its iterative solution. Inspired by recent developments [48]–[50] in global optimization as well as [31]–[34], [51]–[53] in local optimization, it is recommended to amend (4) to a canonical d.c. program, which is minimization of a d.c. function over a convex set. This makes it possible to apply the path-following DCI. For this purpose, rewrite (4) by

$$\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} f(\boldsymbol{z}) - g(\boldsymbol{z}) \ : \ \sum_{n=1}^N \boldsymbol{x}_n - \sum_{n=1}^N \boldsymbol{x}_n^2 \le 0. \tag{5}$$

Using the Lagrangian $\mathcal{L}(\boldsymbol{z},\boldsymbol{x},\mu) := f(\boldsymbol{z}) - g(\boldsymbol{z}) + \mu(\sum_{n=1}^N \boldsymbol{x}_n - \sum_{n=1}^N \boldsymbol{x}_n^2)$ with only one Lagrangian multiplier

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2014.2324588, IEEE Transactions on Wireless Communications

μ to handle the single nonconvex constraint [54], (5) is expressed by $\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} \max_{\mu\geq 0} \mathcal{L}(\boldsymbol{z},\boldsymbol{x},\mu)$, while its Lagrangian duality is $\max_{\mu\geq 0} \min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} \mathcal{L}(\boldsymbol{z},\boldsymbol{x},\mu)$. Note that, in general, there is a duality gap, i.e., $\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} \max_{\mu\geq 0} \mathcal{L}(\boldsymbol{z},\boldsymbol{x},\mu) > \sup_{\mu\geq 0} \min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} \mathcal{L}(\boldsymbol{z},\boldsymbol{x},\mu)$ [36]. However, the strong Lagrangian duality holds for (5) as summarized in the following Proposition.

Proposition 2: The strong Lagrangian duality holds for (5), i.e.,

$$\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} \max_{\mu\geq 0} \mathcal{L}(\boldsymbol{z},\boldsymbol{x},\mu) = \sup_{\mu\geq 0} \min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} \mathcal{L}(\boldsymbol{z},\boldsymbol{x},\mu). \quad (6)$$

If the supremum of the right hand side of (6) attains at $0 < \mu_0 < +\infty$ then the mixed binary program (5) is equivalent to the following convex constrained program for $\mu \geq \mu_0$ in the sense that they share the same optimal value as well as the optimal solutions,

$$\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} \mathcal{L}(\boldsymbol{z},\boldsymbol{x},\mu). \quad (7)$$

Proof: See Appendix A. □

Proposition 2 shows that mixed binary program (1)/(4)/(5) can be solved by means of program (7) for appropriately chosen $\mu > 0$. It can be easily seen that the objective of (7) is the d.c. function $(f(\boldsymbol{z})+\mu\sum_{n=1}^{N}\boldsymbol{x}_n)-(g(\boldsymbol{z})+\mu\sum_{n=1}^{N}\boldsymbol{x}_n^2)$, so indeed (7) is a canonical d.c. program, which is minimization of a d.c. function over a convex set [36].

We are now in a position to outline three algorithmic steps toward finding a solution of the mixed binary program (5) by means of solutions of the canonical d.c. program (7).

A. Step 1: Initialization by box relaxation

For any iterative procedure used to obtain a solution of program (7), one needs a good initial solution (z^*, x^*), which plays a crucial role in the computational efficiency. Obviously, an initial feasible solution to (7) can be easily located but it may be far away from its optimal solution and thus is not efficient. In our approach, we take (z^*, x^*) as the optimized solution of the following box-relaxed program of (1), which is also (7) when setting $\mu = 0$:

$$\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} \mathcal{F}(\boldsymbol{z}) := f(\boldsymbol{z}) - g(\boldsymbol{z}), \quad (8)$$

The above is minimization of a d.c. function subject to convex constraints.

Suppose that $z^{(\kappa)}$ is feasible to (8). Since g is convex, its gradient $\nabla g(z^{(\kappa)})$ at $z^{(\kappa)}$ is also a subgradient [36]. Therefore,

$$f(\boldsymbol{z}) - g(\boldsymbol{z}) \leq f(\boldsymbol{z}) - g(z^{(\kappa)}) - \langle \nabla g(z^{(\kappa)}), \boldsymbol{z} - z^{(\kappa)} \rangle \quad \forall \boldsymbol{z}.$$

It follows that for any feasible $z^{(\kappa)}$ to (8), the following convex program provides a global upper bound minimization for d.c. program (8):

$$\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} [f(\boldsymbol{z}) - g(z^{(\kappa)}) - \langle \nabla g(z^{(\kappa)}), \boldsymbol{z} - z^{(\kappa)} \rangle] \quad (9)$$

Moreover, for the optimal solution $z^{(\kappa+1)}$ of (9), one has

$$\begin{aligned} & f(z^{(\kappa+1)}) - g(z^{(\kappa+1)}) && \leq \\ & f(z^{(\kappa+1)}) - g(z^{(\kappa)}) - \langle \nabla g(z^{(\kappa)}), z^{(\kappa+1)} - z^{(\kappa)} \rangle && \leq \\ & f(z^{(\kappa)}) - g(z^{(\kappa)}) - \langle \nabla g(z^{(\kappa)}), z^{(\kappa)} - z^{(\kappa)} \rangle && = \\ & f(z^{(\kappa)}) - g(z^{(\kappa)}), \end{aligned}$$

which means that $z^{(\kappa+1)}$ is better than $z^{(\kappa)}$ toward optimizing (8) as long as $z^{(\kappa+1)} \neq z^{(\kappa)}$, i.e., convex program (9) generates a proper solution $z^{(\kappa+1)}$. Thus, initialized from a feasible $(z^{(0)}, x^{(0)})$, recursively generating $(z^{(\kappa)}, x^{(\kappa)})$ for $\kappa = 0, 1, \ldots$, by the optimal solution of convex program (9) is a path-following algorithm. In summary, a d.c. procedure for the generic d.c. program (8) is sketched below.

D.C. Iterations (DCIs):

- *Initialization:* Choose an initial feasible solution $(z^{(0)}, x^{(0)})$ of (8).
- *κ-th DC iteration (DCI):* Solve convex program (9) to obtain the optimal solution (z^*, x^*) and set $\kappa \to \kappa+1$, $(z^{(\kappa)}, x^{(\kappa)}) \to (z^*, x^*)$. Given a tolerance level $\epsilon > 0$, stop if

$$\frac{|f(z^{(\kappa)}) - f(z^{(\kappa-1)}) - g(z^{(\kappa)}) + g(z^{(\kappa-1)})|}{f(z^{(\kappa-1)}) - g(z^{(\kappa-1)})|} \leq \epsilon. \quad (10)$$

Our previous works [31]–[34], [51]–[53] demonstrated successful applications of the above DCIs to various nonconvex programs including the beamforming-only optimizations, where their effective d.c. representations in the form of (8) can be found. Since such DCIs will be repeatedly explored in this paper, let us make a new observation on its efficiency. For $(z^{(0)}, x^{(0)}) \neq (z^{(1)}, x^{(1)}) \neq \ldots \neq (z^{(\kappa+1)}, x^{(\kappa+1)})$ it is true that

$$\mathcal{F}(z^{(0)}) > \mathcal{F}(z^{(1)}) > \ldots > \mathcal{F}(z^{(\kappa+1)}). \quad (11)$$

We now show that $(z^{(\kappa+1)}, x^{(\kappa+1)})$ is in fact the optimal solution of the following program

$$\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} \mathcal{F}_\kappa(\boldsymbol{z}) := f(\boldsymbol{z}) - \max_{\nu=0,1,\ldots,\kappa} \{g(z^{(\nu)}) + \langle \nabla g(z^{(\nu)}), \boldsymbol{z} - z^{(\nu)} \rangle\}. \quad (12)$$

Indeed, suppose $(\bar{z}, \bar{x})$ is the optimal solution of (12) and $\bar{\nu} = \arg\max_{\nu=0,1,\ldots,\kappa} \{g(z^{(\nu)}) + \langle \nabla g(z^{(\nu)}), \bar{z} - z^{(\nu)} \rangle\}$. Then $(\bar{z}, \bar{x}) = (z^{(\bar{\nu}+1)}, x^{(\bar{\nu}+1)})$, i.e., $(\bar{z}, \bar{x})$ is the optimal solution of

$$\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} \mathcal{F}(\boldsymbol{z}; z^{(\bar{\nu})}) := f(\boldsymbol{z}) - (g(z^{(\bar{\nu})}) + \langle \nabla g(z^{(\bar{\nu})}), \boldsymbol{z} - z^{(\bar{\nu})} \rangle) \quad (13)$$

because otherwise $\min\ (13) < \mathcal{F}(\bar{z}; z^{(\bar{\nu})}) = \mathcal{F}_\kappa(\bar{z}) = \min\ (12) \leq \min\ (13)$, which is a contradiction. Moreover, as $(z^{(\bar{\nu}+1)}, x^{(\bar{\nu}+1)})$ is the optimal solution of (12), it follows that

$$\mathcal{F}(z^{(\bar{\nu}+1)}) \leq \min\ (12) < \min_{\nu=1,2,\ldots,\kappa} \mathcal{F}(z^{(\nu)}) = \mathcal{F}(z^{(\kappa)})$$

which together with (11) show that $(z^{(\bar{\nu}+1)}, x^{(\bar{\nu}+1)}) = (z^{(\kappa+1)}, x^{(\kappa+1)})$.

Now, it is obvious that $\mathcal{F}(\boldsymbol{z}) \leq \mathcal{F}_{\kappa+1}(\boldsymbol{z}) \leq \mathcal{F}_\kappa(\boldsymbol{z}) \leq \ldots \leq \mathcal{F}_0(\boldsymbol{z}) \quad \forall\ (\boldsymbol{z},\boldsymbol{x}) \in \mathcal{D}$, so convex functions $\mathcal{F}_\kappa$ are iteratively better global approximations of nonconvex function $\mathcal{F}$. Consequently, DCIs by (9) not only generate improved solutions but also provide successively better convexifications for d.c. program (8). As the set $\mathcal{D}$ is compact, the sequence $\{(z^{\kappa}, x^{\kappa})\}$ is bounded and thus by Cauchy theorem there is a convergent subsequence $\{(z^{\kappa_\nu}, x^{\kappa_\nu})\}$, so $\lim\limits_{\nu\to+\infty}[\mathcal{F}(z^{(\kappa_\nu)}) - \mathcal{F}(z^{(\kappa_{\nu+1})})] = 0$. For every κ there is ν such that $\kappa_\nu \le \kappa$ and $\kappa+1 \le \kappa_{\nu+1}$. By (11), $0 \le \lim\limits_{\kappa\to\infty}[\mathcal{F}(z^{(\kappa)}) - \mathcal{F}(z^{(\kappa+1)})] \le \lim\limits_{\kappa\to\infty}[\mathcal{F}(z^{(\kappa_\nu)}) - \mathcal{F}(z^{(\kappa_{\nu+1})})] = 0$, showing that $\lim\limits_{\kappa\to\infty}[\mathcal{F}(z^{(\kappa)}) - \mathcal{F}(z^{(\kappa+1)})] = 0$. Therefore, given a tolerance $\epsilon > 0$, the above DCIs will terminate after finitely many iteration under the stop criterion $\mathcal{F}(z^{(\kappa)}, x^{(\kappa)}) - \mathcal{F}(z^{(\kappa+1)}, x^{(\kappa+1)}) \le \epsilon$, which is normalized by (10). Each accumulation point $(\bar{z}, \bar{x})$ of the sequence $\{(z^{(\kappa)}, x^{(\kappa)}\}$ obviously satisfies $f(z) - f(\bar{z}) - \langle \nabla g(\bar{z}), z - \bar{z}\rangle \ge 0 \ \forall (z, x) \in \mathcal{D}$, which by the convexity of f and $\mathcal{D}$ also includes the minimum principle necessary optimality condition $\langle \nabla f(\bar{z}) - \nabla g(\bar{z}), \boldsymbol{z} - \bar{z}\rangle \ge 0 \quad \forall\ (\boldsymbol{z}, \boldsymbol{x}) \in \mathcal{D}$. In contrast to conditional gradient algorithms, which may be slow in the neighborhood of a local solution and prone to zigzagging [55], our simulation results show that the stop criterion (10) is satisfied after a few iterations.

B. Step 2: Mixed binary solution

Although (7) is already in a d.c. canonical form like (8) with $f(\boldsymbol{z}) \to f(\boldsymbol{z}) + \mu \sum_{n=1}^{N} \boldsymbol{x}_n$ and $g(\boldsymbol{z}) \to g(\boldsymbol{z}) + \mu \sum_{n=1}^{N} \boldsymbol{x}_n^2$, there is an infinite number of d.c. representations for the same nonconvex program in (7), which lead to quite different convex programs (9) for κ-th DCI, i.e., quite different sequences of feasible solutions are generated. In other words, the efficiency of the above generic DCIs is dependent on the choice of a particular d.c. representation. In our approach, we use the following equivalent d.c. representation for (7)

$$\begin{aligned}\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} & [f(\boldsymbol{z}) + \mu(\sum_{n=1}^{N}\boldsymbol{x}_n + (\sum_{n=1}^{N}\boldsymbol{x}_n)^2) - \\ & (g(\boldsymbol{z}) + \mu(\sum_{n=1}^{N}\boldsymbol{x}_n^2 + (\sum_{n=1}^{n}\boldsymbol{x}_n)^2))].\end{aligned} \tag{14}$$

Accordingly, initialized by the optimized solution $(z^*, x^*) \to (z^{(0)}, x^{(0)})$ found by Step 1, the κ-th DCI for (14) is the following convex program

$$\begin{aligned}&\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}} [f(\boldsymbol{z}) + \mu(\sum_{n=1}^{N}\boldsymbol{x}_n + (\sum_{n=1}^{N}\boldsymbol{x}_n)^2) \\ &-(g(z^{(\kappa)}) + \mu(\sum_{n=1}^{N}(x_n^{(\kappa)})^2 + (\sum_{n=1}^{n}x_n^{(\kappa)})^2) \\ &+\langle g(z^{(\kappa)}), \boldsymbol{z} - z^{(\kappa)}\rangle + 2\sum_{n=1}^{N}(x_n^{(\kappa)} + \sum_{n=1}^{N}x_n^{(\kappa)})(\boldsymbol{x}_n - x_n^{(\kappa)}))].\end{aligned} \tag{15}$$

For the optimal solution $(z^{(\kappa+1)}, x^{(\kappa+1)})$ of (15), it is often observed that $\sum_{n=1}^{N} x_n^{(\kappa+1)} - \sum_{n=1}^{N}(x_n^{(\kappa+1)})^2 > 0$, i.e., $x^{(\kappa+1)} \notin \{0,1\}^N$ according to Proposition 2, i.e., $(z^{(\kappa+1)}, x^{(\kappa+1)})$ is infeasible to (5). Thus Step 2 with the κ-DCI (15) is interpreted as an infeasible method for the mixed binary program (5). Of course, its performance also depends on the choice of the penalty parameter μ, which may be tricky [52]. Fortunately, we will see in the simulation section that a simple choice of large enough $\mu > 0$ would make κ-DCI (15) generate a sequence of infeasible solutions $(z^{(\kappa)}, x^{(\kappa)})$ to (5), which converges to its good feasible solution for re-optimization of the next step.

C. Step 3: d.c. re-optimization

Suppose (z^*, x^*) is the solution found by Step 2. It should be noted that there is still no guarantee that (z^*, x^*) is the optimal solution of either (7) or (5). Particularly, x^* may be still not binary but only nearly binary. Therefore, by rounding x^* to the binary values $[x^*]_b \in \{0,1\}^N$ we make a final improvement for optimizing (5) by solving the following d.c. program in continuous variable $\boldsymbol{z}$ only:

$$\min_{(\boldsymbol{z},[x^*]_b)\in\mathcal{K}} \ \mathcal{F}(\boldsymbol{z}) := f(\boldsymbol{z}) - g(\boldsymbol{z}). \tag{16}$$

This program is similar to (8) and thus the above DCIs are directly applicable for its optimized solution.

III. Joint Beamforming and Relay Selection with Orthogonal Transmission

Studied in this section is a wireless relay network in which M source nodes communicate in pair with M destinations nodes with the help of N relays (see Figure 1). All relays, source and destination nodes are each equipped with a single antenna and operate in a half-duplex mode. Information transmission from a source to a destination occurs in two time slots. In the first time-slot, the sources send their signal to the relays. The relays amplify their received signals by multiplying with complex weights and then simply forward these processed signals to all the destinations.

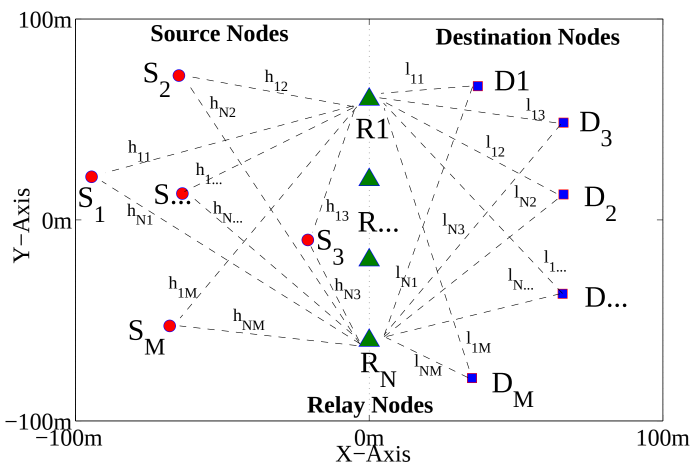

Fig. 1. System Model: Multi-user Multi-relay Wireless Relay Networks.

Let $s = (s_1, s_2, \ldots, s_M)^T \in \mathbb{C}^M$ be the vector of data symbols sent by M sources, which is normalized to zero mean and component-wise independent with variance $\mathbb{E}[|s_i|^2] = 1$. Let $h_n = (h_{n1}, h_{n2}, \ldots, h_{nM})^T \in \mathbb{C}^M,\ n = 1, 2, \ldots, N$ be the vector of uplink channel coefficients between relay n and all users and let $\ell_n = (\ell_{n1}, \ell_{n2}, \ldots, \ell_{nM})^T \in \mathbb{C}^M,\ n = 1, 2, \ldots, N$ be the vector of downlink channel coefficients

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2014.2324588, IEEE Transactions on Wireless Communications

between relay n and all the destinations. As mentioned before, these channel coefficients are assumed to be available at N relays. Communications via source-to-destination links are not taken into account due to severer path loss. Furthermore, let

$$\boldsymbol{x}_m = (\boldsymbol{x}_{1m}, \boldsymbol{x}_{2m}, \dots, \boldsymbol{x}_{Nm})^T \in \{0,1\}^N \subset R^N, \quad m = 1, 2, \dots, M \tag{17}$$

be the vector representing the link connectivity between the user-destination pair m and all the relays. Specifically, $\boldsymbol{x}_{nm} = 1$ means that user m sends its signal s_m to the relay n through channel h_{nm} and this relay is also connected to destination m to forward the processed signal through channel ℓ_{nm}. All coefficients h_{nm} and ℓ_{nm} represent flat-fading channels. Furthermore, h_{nm} and ℓ_{nm} typically remain unchanged over a block of transmitted/relaying signals, i.e., block fading. On the other hand, $\boldsymbol{x}_{nm} = 0$ means that relay n is disconnected from both user m and destination m. To control how the N relays are used to help M source-destination pairs, the number of relays assigned to each user shall be limited to at most N_R [26], [29]:

$$\sum_{n=1}^{N} \boldsymbol{x}_{nm} \le N_R, \; m = 1, 2, \dots, M. \tag{18}$$

This section focuses on a scenario when user-destination pairs operate in orthogonal channels, i.e. the transmissions from the sources to the relays and from the relays to the destinations are orthogonal. For simplicity, time-division multiplexing shall be assumed, but other forms of orthogonal transmission such as frequency-division multiplexing and orthogonal code-division multiplexing can also be applied. Under orthogonal transmission, the received signal at relay n from user m is

$$y_{nm} = \boldsymbol{x}_{nm} h_{nm} s_m + n_n, \tag{19}$$

where n_n represents AWGN at the relay, which is modeled as a zero-mean circularly complex Gaussian random variable with variance σ_R^2. The received signal y_{nm} is then multiplied by a complex beamforming weight $\boldsymbol{\xi}_{nm}$ before being forwarded to destination m. Since relay n can help multiple sources, the constraint on its transmitted power is expressed as

$$\sum_{m=1}^{M} (\boldsymbol{x}_{nm} |\boldsymbol{\xi}_{nm}|)^2 (|h_{nm}|^2 + \sigma_R^2) \le P_n, \quad n = 1, 2, \dots, N, \tag{20}$$

while the total relaying constraint is

$$\sum_{n=1}^{N} \sum_{m=1}^{M} (\boldsymbol{x}_{nm} |\boldsymbol{\xi}_{nm}|)^2 (|h_{nm}|^2 + \sigma_R^2) \le P_T. \tag{21}$$

It should be noted that constraint (20) is related to the hardware power limitation of each relay's transmitter, while constraint (21) expresses the allowable total power that all relays consume in amplifying all of their received signals. Since all the relays transmit to the destination simultaneously, the received signal at destination m is

$$y_{D,m} = \sum_{n=1}^{N} \boldsymbol{x}_{nm} \boldsymbol{\xi}_{nm} \ell_{nm} (h_{nm} s_m + n_n) + n_{D,m}. \tag{22}$$

Using the polar representation $\boldsymbol{\xi}_{nm} = |\boldsymbol{\xi}_{nm}| e^{\jmath \arg(\boldsymbol{\xi}_{nm})}$ it can be easily shown that, in order to maximize the signal-to-noise ratio (SNR) at destination m, the optimal phase of $\boldsymbol{\xi}_{nm}$ is $-\arg(\ell_{nm} h_{nm})$. That is,

$$\boldsymbol{\xi}_{nm} = |\boldsymbol{\xi}_{nm}| e^{-\jmath \arg(\ell_{nm} h_{nm})}. \tag{23}$$

It then follows that equation (22) reduces to the following form, which involves only the magnitudes $|\boldsymbol{\xi}_{nm}|$:

$$\begin{aligned} y_{D,m} = & \sum_{n=1}^{N} \boldsymbol{x}_{nm} |\boldsymbol{\xi}_{nm}| (|\ell_{nm} h_{nm}| s_m + e^{-\jmath \arg(\ell_{nm} h_{nm})} n_n) \\ & + n_{D,m}, \quad n = 1, 2, \dots, N. \end{aligned} \tag{24}$$

Define variables

$$\begin{aligned} \boldsymbol{\xi}_m = (\boldsymbol{\xi}_{1m}, \dots, \boldsymbol{\xi}_{Nm})^T \in \mathbb{C}^N, m = 1, 2, \dots, M, \\ \Xi = (\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_M) \in \mathbb{C}^{N \times M}, \boldsymbol{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_M) \in \mathbb{R}^{N \times M}. \end{aligned}$$

Then the SNR at destination m can be written as

$$\mathbf{SNR}_m(\boldsymbol{x}_m, \boldsymbol{\xi}_m) = \frac{\left(\sum_{n=1}^{N} \boldsymbol{x}_{nm} |\boldsymbol{\xi}_{nm}| |\ell_{nm} h_{nm}|\right)^2}{\sigma_R^2 \sum_{n=1}^{N} (\boldsymbol{x}_{nm} |\boldsymbol{\xi}_{nm}|)^2 |\ell_{nm}|^2 + \sigma_D^2}, \tag{25}$$

where σ_D^2 is the variance of AWGN at the destination. In order to improve the overall throughput of the relay network as well as to offer fairness among the users, we consider the following SNR maximin program:

$$\max_{\boldsymbol{X}, \Xi} \min_{m=1,2,\dots,M} \mathbf{SNR}_m(\boldsymbol{x}_m, \boldsymbol{\xi}_m) \quad \text{s.t.} \quad (17), (18), (20), (21). \tag{26}$$

With the following variable change

$$\begin{aligned} \boldsymbol{x}_{nm} |\boldsymbol{\xi}_{nm}| \to \bar{\boldsymbol{\xi}}_{nm}, \; \bar{\boldsymbol{\xi}}_m = (\bar{\boldsymbol{\xi}}_{1m}, \dots, \bar{\boldsymbol{\xi}}_{Nm})^T \in \mathbb{C}^N, \\ \bar{\Xi} = (\bar{\boldsymbol{\xi}}_1, \dots, \bar{\boldsymbol{\xi}}_M) \in \mathbb{R}^{N \times M}, \end{aligned} \tag{27}$$

and additional variable $\boldsymbol{q} = (\boldsymbol{q}_1, \boldsymbol{q}_2, \dots, \boldsymbol{q}_M)^T \in \mathcal{R}^M$ for the additional constraints

$$\sigma_R^2 \sum_{n=1}^{N} \bar{\boldsymbol{\xi}}_{nm}^2 |\ell_{nm}|^2 \le \boldsymbol{q}_m, \; m = 1, 2, \dots, M, \tag{28}$$

and using the fact that $\boldsymbol{x}_{nm}^2 = \boldsymbol{x}_{nm}$ for $\boldsymbol{x}_{nm} \in \{0, 1\}$, the SNR maximin program in (26) is

$$\max_{\boldsymbol{X}, \bar{\Xi}, \boldsymbol{q}} \min_{m=1,2,\dots,M} \mathbf{SNR}_m(\bar{\boldsymbol{\xi}}_m, \boldsymbol{q}_m) \quad :=$$

$$\frac{\left(\sum_{n=1}^{N} \bar{\boldsymbol{\xi}}_{nm} |\ell_{nm} h_{nm}|\right)^2}{\boldsymbol{q}_m + \sigma_D^2} \quad \text{s.t.} \quad (17), (18), (28), \tag{29a}$$

$$\sum_{m=1}^{M} \bar{\boldsymbol{\xi}}_{nm}^2 (|h_{nm}|^2 + \sigma_R^2) \le P_n, \tag{29b}$$

$$\sum_{m=1}^{M} \sum_{m=1}^{M} \bar{\boldsymbol{\xi}}_{nm}^2 (|h_{nm}|^2 + \sigma_R^2) \le P_T, \tag{29c}$$

$$\bar{\boldsymbol{\xi}}_{nm}^2 (|h_{nm}|^2 + \sigma_R^2) \le \boldsymbol{x}_{nm} P_n, \tag{29d}$$

$$n = 1, 2, \dots, N \,; m = 1, 2, \dots, M.$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2014.2324588, IEEE Transactions on Wireless Communications

Now, (29) is equivalent to

$$\max_{\boldsymbol{X},\bar{\Xi},\boldsymbol{q}} \min_{m=1,2,\dots,M} [2\ln(\sum_{n=1}^{N} \bar{\boldsymbol{\xi}}_{nm}|\ell_{nm}h_{nm}|) - \ln(\boldsymbol{q}_m + \sigma_D^2)] \quad \text{s.t.} \quad (17),(18),(28),(29b)-(29d). \tag{30}$$

One can see that the maximin program (30) is more favorable than the maximin program (29) because the coupled nonlinear terms $\boldsymbol{x}_{nm}|\boldsymbol{\xi}_{nm}|^2$ in (26) have been decoupled by the variable change in (27). The difficulty of the maximin program (30) now lies on its nonconvex objective function (with respect to only the beamforming power variable $\bar{\Xi}$) and the binary constraint in (17) (with respect to only the assignment variable $\boldsymbol{x}$). More importantly, using [36, Prop. 3.1] program (30) can be represented by

$$-\min_{\boldsymbol{X},\bar{\Xi},\boldsymbol{q}} [f_1(\bar{\boldsymbol{\xi}},\boldsymbol{q}) - f_2(\boldsymbol{q})] \;:\; (17),(18),(28)(29b)-(29d), \tag{31}$$

where $f_1(\bar{\boldsymbol{\xi}},\boldsymbol{q})$ and $f_2(\bar{\boldsymbol{\xi}})$ are defined as

$$\begin{aligned} f_1(\bar{\Xi},\boldsymbol{q}) &= \max_{m=1,2,\dots,M}[-2\ln(\sum_{n=1}^{N}\bar{\boldsymbol{\xi}}_{nm}|\ell_{nm}h_{nm}|) \\ &\quad -\sum_{i\neq m}\ln(\boldsymbol{q}_i+\sigma_D^2)] \\ f_2(\boldsymbol{q}) &= -\sum_{m=1}^{M}\ln(\boldsymbol{q}_m+\sigma_D^2), \end{aligned} \tag{32}$$

which are convex functions as the maximum and summation of convex functions [36]. Therefore (31) is already in the form of (1) with

$$\begin{aligned} &(\boldsymbol{z},\boldsymbol{x}) \to ((\bar{\Xi},\boldsymbol{q}),\boldsymbol{X}), \\ &\mathcal{K} \to \{((\bar{\Xi},\boldsymbol{q}),\boldsymbol{X}) \;:\; (18),(28),(29b)-(29d)\}, \\ &f(\boldsymbol{z}) \to f_1(\bar{\Xi},\boldsymbol{q}),\; g(\boldsymbol{z}) \to f_2(\boldsymbol{q}), \end{aligned}$$

and its according d.c. expression (7) is

$$-\min_{\bar{\Xi},\boldsymbol{q},\boldsymbol{X}} [f_1(\bar{\Xi},\boldsymbol{q}) - f_2(\boldsymbol{q}) + \mu(\sum_{m=1}^{M}\sum_{n=1}^{N}\boldsymbol{x}_{nm} - \sum_{m=1}^{M}\sum_{n=1}^{N}\boldsymbol{x}_{nm}^2)] \quad : \quad (18),(28)(29b)-(29d), \tag{33a}$$

$$\boldsymbol{x}_{nm} \in [0,1], n=1,2,\dots,N;\; m=1,2,\dots,M. \tag{33b}$$

Concerning the above optimization problem, the generic Step 1 to Step 3 described before are detailed next for finding optimized solutions of program (33).

A. Step 1: Initialization by all relay beamforming optimization

The box relaxation (8) corresponding to (31) is

$$-\min_{\boldsymbol{X},\bar{\Xi},\boldsymbol{q}} [f_1(\bar{\boldsymbol{\xi}},\boldsymbol{q}) - f_2(\boldsymbol{q})] \;:\; (18),(28),(29b)-(29d),(33b), \tag{34}$$

which is just the corresponding beamforming design for all relays.

According to our computational experience, the following initialization is good with respect to the convergence of its DCIs:

$$\begin{aligned} \bar{\xi}_{nm}^{(0)} &= [\min\{\frac{P_n}{M(|h_{nm}|^2+\sigma_R^2)}, P_n\}]^{1/2}, \\ q_m^{(0)} &= \sigma_R^2\sum_{n=1}^{N}(\bar{\xi}_{nm}^{(0)})^2|\ell_{nm}|^2, \\ n &= 1,\dots,N;\; m=1,\dots,M. \end{aligned}$$

For $\kappa=0,1,\dots,$ the κ-th iteration (9) for program (34) that yields an iterative solution $\bar{\xi}_{nm}^{(\kappa+1)}$ and $q_m^{(\kappa+1)}$ is

$$\min_{\boldsymbol{X},\bar{\Xi},\boldsymbol{q}} [f_1(\bar{\Xi},\boldsymbol{q}) - f_2(q^{(\kappa)}) + \langle\nabla f_2(q^{(\kappa)}),\boldsymbol{q}-q^{(\kappa)}\rangle] \quad \text{s.t.} \quad (29b)-(29d),(33b), \tag{35}$$

where

$$\langle\nabla f_2(q^{(\kappa)}),\boldsymbol{q}-q^{(\kappa)}\rangle = \sum_{m=1}^{M}\frac{1}{q_m^{(\kappa)}+\sigma_D^2}(\boldsymbol{q}_m - q_m^{(\kappa)}). \tag{36}$$

There are totally $2NM+M$ decision variables and $3NM+N+1$ convex and linear constraints in the the convex program (35). Thus its computational complexity is $O((2NM+M)^3(3NM+N+1))$ (see e.g. [56]).

Alternatively, the optimal solution of d.c. program (34) can be found through bisection in parameter $\boldsymbol{t}$ for the following parametric second-order cone program:

$$\begin{aligned} &\max_{\boldsymbol{X},\bar{\Xi},\boldsymbol{t}} \boldsymbol{t} \;:\; (18),(29b)-(29d), \\ &\sum_{n=1}^{N}\bar{\boldsymbol{\xi}}_{nm}|\ell_{nm}h_{nm}| \geq \boldsymbol{t}[\sigma_R^2\sum_{n=1}^{N}\bar{\boldsymbol{\xi}}_{nm}^2|\ell_{nm}|^2+\sigma_D^2]^{1/2}, \\ &m=1,2,\dots,M, \end{aligned} \tag{37}$$

which requires a second-order cone program (SOCP) solver to check for the feasibility of the linear constraints (37) when $\boldsymbol{t}$ is held fixed.

It should be noted that theoretically the parametric SOCP (37) yields its global optimal solution, while DCIs (35) of local search yield a solution that is not necessarily optimal of the d.c. program (34). Our simulation results suggest that the optimized solution located by these two different procedures are the same and thus the DCIs (35) are indeed capable of locating the global optimal solution of (34).

In summary, as an initial solution $(\bar{\Xi}^*, X^*)$ used in Step 2 of the solution algorithm for the mixed binary program (31), we take $\bar{\Xi}^*$ found through either the above DCIs (35) for d.c program (34) or the parametric linear program (37), while

$$\begin{aligned} x_{nm}^* &= (\bar{\xi}_{nm}^*)^2(|h_{nm}|^2+\sigma_R^2)/P_n, \\ n &= 1,2,\dots,N;\; m=1,2,\dots,M. \end{aligned} \tag{38}$$

B. Step 2: Mixed beamforming and assignment optimization

Using the solution $(\bar{\Xi}^*, q^*, X^*)$ found from Step 1 as the initial solution $(\bar{\Xi}^{(0)}, q^{(0)}, X^{(0)})$, for $\kappa=0,1,\dots,$ the corresponding κ-th DCI (15) that gives the iterative solution $(\bar{\Xi}^{(\kappa+1)}, q^{(\kappa+1)}, X^{(\kappa+1)})$ of (33) is the following convex program:

$$\begin{aligned} \min_{\boldsymbol{X},\bar{\Xi},\boldsymbol{q}} & [f_1(\bar{\Xi},\boldsymbol{q}) + \mu h_1(\boldsymbol{X}) - f_2(q^{(\kappa)}) - \mu h_2(X^{(\kappa)}) \\ & -\langle\nabla f_2(q^{(\kappa)}),\boldsymbol{q}-q^{(\kappa)}\rangle] - \mu\langle\nabla h_2(X^{(\kappa)}),\boldsymbol{X}-X^{(\kappa)}\rangle] \\ & \text{s.t.} \quad (18),(28),(29b)-(29d),(33b), \end{aligned} \tag{39}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2014.2324588, IEEE Transactions on Wireless Communications

where $\langle \nabla f_2(q^{(\kappa)}), \boldsymbol{q} - q^{(\kappa)} \rangle$ is defined by (36), while

$$h_1(\boldsymbol{X}) = \sum_{m=1}^{M} (\sum_{n=1}^{N} \boldsymbol{x}_{nm} + (\sum_{n=1}^{N} \boldsymbol{x}_{nm})^2),$$
$$h_2(\boldsymbol{X}) = \sum_{m=1}^{M} (\sum_{n=1}^{N} \boldsymbol{x}_{nm}^2 + (\sum_{n=1}^{N} \boldsymbol{x}_{nm})^2)$$

and

$$\langle \nabla h_2(X^{(\kappa)}), \boldsymbol{X} - X^{(\kappa)} \rangle = 2 \sum_{m=1}^{M} (\sum_{n=1}^{N} (x_{nm}^{(\kappa)} + \sum_{n=1}^{N} x_{nm}^{(\kappa)})(\boldsymbol{x}_{nm} - x_{nm}^{(\kappa)})).$$

The computational complexity of (39) is $O((2NM + M)^3(2NM + 2M + N + 1))$.

C. Step 3: Assigned relay beamforming re-optimization

Suppose that $(\bar{\Xi}^*, X^*)$ is a solution found by Step 2 as above. The corresponding d.c. program (16) to improve the solution further by substituting the rounded binary values $[X^*]_b$ in constraints (29b)-(29c) is the following program in only continuous variable $\bar{\Xi}$ of power allocation:

$$-\min_{\bar{\Xi}, \boldsymbol{q}} [f_1(\bar{\Xi}, \boldsymbol{q}) - f_2(\boldsymbol{q})] \quad : \quad (28), (29b) - (29c), \quad \text{(40a)}$$
$$\bar{\boldsymbol{\xi}}_{nm}^2 (|h_{nm}|^2 + \sigma_R^2) \leq [x_{nm}^*]_b P_n, \quad \text{(40b)}$$
$$n = 1, 2, \ldots, N; \ m = 1, 2, \ldots, M.$$

Constraint (40b) given $[X^*]_b$ implies that program (40) involves only $N_R M$ variables $\bar{\boldsymbol{\xi}}_{nm}$, which correspond to $[x_{nm}^*]_b = 1$ in (40b), i.e., program (40) is assigned relay beamforming re-optimization. Therefore the DCIs adapted to the d.c. program (40) is to initialize from $\bar{\Xi}^{(0)} = \bar{\Xi}^*$ and for $\kappa = 0, 1, \ldots$, κ-th DCI to output the iterative solution $\bar{\Xi}^{(\kappa+1)}$ is the following convex program:

$$\min_{\bar{\Xi}, \boldsymbol{q}} [f_1(\bar{\Xi}, \boldsymbol{q}) - f_2(q^{(\kappa)}) - \langle \nabla f_2(q^{(\kappa)}), \boldsymbol{q} - q^{(\kappa)} \rangle] \quad \text{s.t.}$$
$$(28), (29b) - (29c), (40b), \quad (41)$$

with $\langle \nabla f_2(\bar{\Xi}^{(\kappa)}), \bar{\Xi} - \bar{\Xi}^{(\kappa)} \rangle$ defined by (36). Its computational complexity is $O((N_R + M)^3(2NM + N + M + 1)$. The stopping criterion in (10) can also be used to output the final solution.

Alternatively, relay beamforming based on $[X^*]_b \in \{0,1\}^{N \times M}$ can also be done by a parametric second-order cone program:

$$\max_{\bar{\Xi}, \boldsymbol{t}} \ \boldsymbol{t} \quad \text{s.t.} \quad (29b) - (29c), (40b),$$
$$\sum_{n=1}^{N} \bar{\boldsymbol{\xi}}_{nm} |\ell_{nm} h_{nm}| \geq \boldsymbol{t} [\sigma_R^2 \sum_{n=1}^{N} \bar{\boldsymbol{\xi}}_{nm}^2 |\ell_{nm}|^2 + \sigma_D^2]^{1/2}, \quad (42)$$
$$m = 1, 2, \ldots, M.$$

In summary, the proposed procedure for finding the solutions of the mixed binary program in (31) consists of the above three sequential steps: the initial Step 1 for locating a good initial solution for Step 2, Step 2 for locating the optimized mixed beamforming and assignment solutions, and Step 3 for re-optimizing the assigned relay beamforming.

D. A Case Study with MRC

It is worthwhile to point out that the above procedure is also applicable to the following optimization problem concerning the maximal-ratio-combining (MRC) receiver as studied in [26]:

$$\max_{\boldsymbol{X}, \bar{\Xi}} \min_{m=1,2,\ldots,M} \varphi_m^{\text{MRC}}(\bar{\Xi}) \ : \ (17), (18), (29b) - (29d), \quad (43)$$

where

$$\begin{aligned}\varphi_m^{\text{MRC}}(\bar{\Xi}) &:= \sum_{n=1}^{N} \frac{\bar{\boldsymbol{\xi}}_{nm} |\ell_{nm} h_{nm}|^2}{\sigma_R^2 \bar{\boldsymbol{\xi}}_{nm} |\ell_{nm}|^2 + \sigma_D^2} \\ &= \sum_{n=1}^{N} \frac{|\ell_{nm} h_{nm}|^2}{\sigma_R^2 |\ell_{nm}|^2} - \\ &\quad \sum_{n=1}^{N} \frac{\sigma_D^2 |\ell_{nm} h_{nm}|^2}{\sigma_R^2 |\ell_{nm}|^2 (\sigma_R^2 |\ell_{nm}|^2 \bar{\boldsymbol{\xi}}_{nm} + \sigma_D^2)}. \quad (44)\end{aligned}$$

In contrast to the objective functions in (29a), which are not concave, the objective functions in (44) are clearly seen concave in $\bar{\boldsymbol{\xi}}_{nm} = |\boldsymbol{\xi}_{nm}|^2$. It can also be seen that (44) is an upper bound of the objective function in (30a) as well. The sum-rate maximization considered in [27] is actually

$$\max_{\boldsymbol{X}, \bar{\Xi}} \sum_{m=1}^{M} \varphi_m^{\text{MRC}}(\bar{\Xi}) \ : \ (17), (18), (29b) - (29d),$$

which is quite similar to (43) but appears less insightful.

Obviously, the mixed binary program in (43) is a special case of the mixed binary program in (31) with

$$f_1(\bar{\Xi}) := \max_{m=1,2,\ldots,M} [-\varphi_m^{\text{MRC}}(\bar{\Xi})], \quad f_2(\bar{\Xi}) \equiv 0. \quad (45)$$

Accordingly, the realizations of Step 1 to Step 3 to obtain the solution of (43) are simplified as follows.

- Step 1: With $f_2(\bar{\Xi}) \equiv 0$, program (34) becomes the following convex program:

$$-\min_{\bar{\Xi}} \max_{m=1,2,\ldots,M} [-\varphi_m^{\text{MRC}}(\bar{\Xi})] \ : \ (18), (29b) - (29c),$$
$$0 \leq \bar{\boldsymbol{\xi}}_{nm} \leq P_n, \ n = 1, 2, \ldots, N; \ m = 1, 2, \ldots, M, \quad (46)$$

  which can be easily solved by convex programming solvers instead of DCIs.
- Step 2: Initialized from $\bar{\Xi}^{(0)} = \bar{\Xi}^*$ and $x_{nm}^{(0)} = P_n / \bar{\xi}_{nm}^{(0)}$, where $\bar{\Xi}^*$ is the optimal solution of (46), for $\kappa = 0, 1, 2, \ldots$ use DCIs to solve program (39) with $f_1(\cdot)$ and $f_2(\cdot)$ defined by (45) to generate a sequence of improved solutions $\{(\bar{\Xi}^{(\kappa)}, X^{(\kappa)})\}$ for the mixed binary program (43).
- Step 3: With the solution x^* found from Step 2, substituting $\boldsymbol{X} = [X^*]_b$ in constraint (30c) of program (43) to result in the following convex program in only continuous variables $\boldsymbol{\xi}_{nm}$ (with $[x_{nm}^*]_b = 1$):

$$\max_{\bar{\Xi}} \min_{m=1,2,\ldots,M} \varphi_m^{\text{MRC}}(\bar{\Xi}) \ : \ (29b) - (29c), (40b). \quad (47)$$

The above procedure is different from that presented in [26] in the crucial Step 2. A joint optimization in mixed binary optimization $(\Xi, \boldsymbol{X})$ is treated in [26]. From the optimal solution $\bar{\Xi}^*$ of Step 1, Step 2 in [26] is to take N_U largest

$\bar{\xi}^*_{nm}$ for each m and then assign $x_{nm} = 1$ for these relays to solve (47) in $\bar{\boldsymbol{\xi}}$. The simulation results presented in Section V will show the superior performance of our solutions compared to that presented in [26].

IV. Joint Beamforming and Relay Selection with Nonorthogonal Transmission

In a multi-user wireless relay network, the bandwidth resource is most efficiently utilized if all the users' signals are allowed to be transmitted over the same channel to and from each relay. The price to be paid for such a higher spectral efficiency transmission scheme is the existence of inter-user interference at each relay and each destination node.

Specifically, the received signal at relay n is now given by

$$y_{R,n} = \boldsymbol{x}_n(\sum_{m=1}^{M} h_{nm}s_m + n_n),$$

where, as before, n_n represents AWGN at each relay whose variance is σ_R^2 and $\boldsymbol{x}_n$ is the link connectivity between relay n and the users and destinations. $x_n = 1$ means that relay n is selected for the connection and $x_n = 0$ means that relay n is not selected. The relay selection is restricted by

$$\sum_{n=1}^{N} \boldsymbol{x}_n \le N_R,\ \boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_N)^T \in \{0,1\}^N. \tag{48}$$

Relay n then amplifies its received signal by a complex gain $\boldsymbol{\xi}_n$, $n = 1, 2, \dots, N$. Thus, relay n sends the following signal to destination m:

$$y_n = \boldsymbol{x}_n\boldsymbol{\xi}_n(\sum_{m=1}^{M} h_{nm}s_m + n_n) \tag{49}$$

The complex gain $\boldsymbol{\xi}_n$ is constrained by the following individual relaying power

$$\boldsymbol{x}_n|\boldsymbol{\xi}_n|^2(\sum_{m=1}^{M}|h_{nm}|^2 + \sigma_R^2) \le P_n,\ n = 1, 2, \dots, N, \tag{50}$$

as well as the total relaying power

$$\sum_{n=1}^{N}\boldsymbol{x}_n|\boldsymbol{\xi}_n|^2(\sum_{m=1}^{M}|h_{nm}|^2 + \sigma_R^2) \le P_T. \tag{51}$$

Accordingly, the received signal at destination m is

$$y_{D,m} = \sum_{n=1}^{N}\boldsymbol{x}_n\boldsymbol{\xi}_n\ell_{nm}(\sum_{m=1}^{M} h_{nm}s_m + n_n) + n_{D,m} \tag{52}$$

where $n_{D,m}$ is AWGN at destination m, whose variance is σ_D^2.

As only signal $\sum_{n=1}^{N}\boldsymbol{\xi}_n\boldsymbol{x}_n^2\ell_{nm}h_{nm}s_m = \sum_{n=1}^{N}\boldsymbol{\xi}_n\boldsymbol{x}_n\ell_{nm}h_{nm}s_m$ is of interest in (52), the signal-to- interference-plus-noise ratio (SINR) at destination m is given by

$$\mathbf{SINR}_m(\boldsymbol{\xi}, \boldsymbol{x}) = \frac{|\sum_{n=1}^{N}\boldsymbol{x}_n\boldsymbol{\xi}_n\ell_{nm}h_{nm}|^2}{\sum_{j\neq m}^{M}|\sum_{n=1}^{N}\boldsymbol{x}_n\boldsymbol{\xi}_n\ell_{nm}h_{nj}|^2 + \sigma_R^2\sum_{n=1}^{N}\boldsymbol{x}_n|\boldsymbol{\xi}_n|^2|\ell_{nm}|^2 + \sigma_D^2}. \tag{53}$$

The joint optimization in beamforming weight $\boldsymbol{\xi}_n$ and relay assignment $\boldsymbol{x}$ is formulated as follows

$$\max_{\boldsymbol{\xi},\boldsymbol{x}} \min_{m=1,2,\dots,M} \mathbf{SINR}_m(\boldsymbol{\xi}, \boldsymbol{x})\ :\ (48), (50), (51). \tag{54}$$

For each fixed binary x^* that is feasible to (48), (54) is still a (nonconvex) maximin program, which has been successfully addressed in [30]. However, it is not practical to address (54) by solving $C(N, N_R)$ such maximin programs.

The objective functions in program (54), which are fractions of multivariate polynomials, are highly complex. The next Theorem is the first step to recognize its partial convexities that will be useful for the later development.

Theorem 1: With the following variable change

$$\boldsymbol{x}_n\boldsymbol{\xi}_n \to \tilde{\boldsymbol{\xi}}_n,\ \tilde{\boldsymbol{\xi}} := (\tilde{\boldsymbol{\xi}}_1, \dots, \tilde{\boldsymbol{\xi}}_N)^T \tag{55}$$

the mixed binary program (54) is equivalent to the following mixed binary program

$$\max_{\boldsymbol{x},\tilde{\boldsymbol{\xi}},\boldsymbol{q}} \min_{m=1,2,\dots,M} \varphi_m(\tilde{\boldsymbol{\xi}}, \boldsymbol{q}) := \frac{|\sum_{n=1}^{N}\tilde{\boldsymbol{\xi}}_n\ell_{nm}h_{nm}|^2}{\boldsymbol{q}_m + \sigma_D^2}\ :\ (48), \tag{56a}$$

$$\sum_{j\neq m}^{M}|\sum_{n=1}^{N}\tilde{\boldsymbol{\xi}}_n\ell_{nm}h_{nj}|^2 + \sigma_R^2\sum_{n=1}^{N}|\tilde{\boldsymbol{\xi}}_n|^2|\ell_{nm}|^2 \le \boldsymbol{q}_m, \tag{56b}$$

$$|\tilde{\boldsymbol{\xi}}_n|^2(\sum_{m=1}^{M}|h_{nm}|^2 + \sigma_R^2) \le P_n, \tag{56c}$$

$$\sum_{n=1}^{N}|\tilde{\boldsymbol{\xi}}_n|^2(\sum_{m=1}^{M}|h_{nm}|^2 + \sigma_R^2|) \le P_T, \tag{56d}$$

$$|\tilde{\boldsymbol{\xi}}_n|^2(\sum_{m=1}^{M}|h_{nm}|^2 + \sigma_R^2|) \le \boldsymbol{x}_nP_T, \tag{56e}$$

$$m = 1, 2, \dots, M;\quad n = 1, 2, \dots, N,$$

where each function φ_m is convex in $(\tilde{\boldsymbol{\xi}}, \boldsymbol{q})$.

Proof: Any feasible solution (ξ, x) to (54) leads to the obviously feasible solution $(\tilde{\xi}, q, x)$ to (56) with $\tilde{\xi}$ defined by (55) and q_m accordingly defined by the left hand side of (56b). Therefore $\max\ (54) \le \max\ (56)$. On the other hand, for the optimal solution $(\tilde{\xi}^{\rm opt}, q^{\rm opt}, x^{\rm opt})$ of (56) and then $\xi_n^{\rm opt} = x_n\tilde{\xi}_n$, $n = 1, 2, \dots, N$, it is clear that $(\xi^{\rm opt}, x^{\rm opt})$ is feasible to (54) and then by (56b) it is $\mathbf{SINR}_m(\xi^{\rm opt}, x^{\rm opt}) \ge \varphi_m(\tilde{\xi}^{\rm opt}, q^{\rm opt})$, $m = 1, 2, \dots, M$. The last inequalities imply $\max\ (56) \le \max\ (54)$, which together with $\max\ (54) \le \max\ (56)$ show that $\max\ (56) = \max\ (54)$.

Finally, each function $\varphi_m(\tilde{\boldsymbol{\xi}}, \boldsymbol{q})$ is convex as a composite of the convex function $\phi(\boldsymbol{s}, \boldsymbol{q}_m) = |\boldsymbol{s}|^2/(\boldsymbol{q}_m + \sigma_D^2)$ and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2014.2324588, IEEE Transactions on Wireless Communications

linear function $s(\tilde{\boldsymbol{\xi}}_n) = |\sum_{n=1}^{N} \tilde{\boldsymbol{\xi}}_n \ell_{nm} h_{nm}|^2$ [36]. The convexity of $\phi(\boldsymbol{s}, \boldsymbol{q}_m)$ on $\mathcal{C} \times R_+$ immediately follows from positive definiteness of its Hessian. $\square$

Using [36, Prop. 3.1] again to rewrite program (56) as

$$-\min_{\boldsymbol{x}, \tilde{\boldsymbol{\xi}}, \boldsymbol{q}} [f_1(\tilde{\boldsymbol{\xi}}, \boldsymbol{q}) - f_2(\tilde{\boldsymbol{\xi}}, \boldsymbol{q})] \ : \ (56b) - (56e), (48) \qquad (57)$$

with $f_1(\boldsymbol{\xi}, \boldsymbol{q}) := \max_{m=1,2,\ldots,M} \sum_{i \neq m} \phi_i(\boldsymbol{\xi}, \boldsymbol{q})$ and $f_2(\boldsymbol{\xi}, \boldsymbol{q}) := \sum_{m=1}^{M} \phi_m(\boldsymbol{\xi}, \boldsymbol{q})$. Each $\phi_m(\boldsymbol{\xi}, \boldsymbol{q})$ is convex by Theorem 1, so indeed f_1 and f_2 are convex as maximum and sum of convex functions [36]. Therefore, program (57) is in the form of (1) with

$$\begin{array}{c} (\tilde{\boldsymbol{\xi}}, \boldsymbol{q}) \to \boldsymbol{z}, \ \boldsymbol{x} \to \boldsymbol{x}, \ f(\boldsymbol{z}) \to f_1(\tilde{\boldsymbol{\xi}}, \boldsymbol{q}), \\ g(\boldsymbol{z}) \to f_2(\tilde{\boldsymbol{\xi}}, \boldsymbol{q}), \ \mathcal{K} \to \{(\tilde{\boldsymbol{\xi}}, \boldsymbol{q}, \boldsymbol{x}) \ : \ (56b) - (56e), \\ \boldsymbol{x} \in [0,1]^N, \ \sum_{n=1}^{N} \boldsymbol{x}_n \leq N_R\}, \end{array} \qquad (58)$$

and its exact canonical d.c. expression (7) accordingly is

$$\begin{array}{c} \min_{\boldsymbol{x}, \tilde{\boldsymbol{\xi}}, \boldsymbol{q}} [f_1(\tilde{\boldsymbol{\xi}}, \boldsymbol{q}) - f_2(\tilde{\boldsymbol{\xi}}, \boldsymbol{q}) + \mu(\sum_{n=1}^{N} \boldsymbol{x}_n - \sum_{n=1}^{N} \boldsymbol{x}_n^2)] \ : \\ (56b) - (56e), \sum_{n=1}^{N} \boldsymbol{x}_n \leq N_R, \ \boldsymbol{x} \in [0,1]^N. \end{array} \qquad (59)$$

A. Step 1: Initialization by all relay beamforming optimization

By (58), the corresponding box relaxed program (8) corresponding to (57)/(59) is

$$-\min_{\tilde{\boldsymbol{\xi}}, \boldsymbol{q}} [f_1(\tilde{\boldsymbol{\xi}}, \boldsymbol{q}) - f_2(\tilde{\boldsymbol{\xi}}, \boldsymbol{q})] \ : \ (58), \qquad (60)$$

which is simply an all-relay beamforming SINR maximin optimization. Therefore, initializing from $(\tilde{\xi}^{(0)}, q^{(0)})$, for $\kappa = 0, 1, \ldots$, the corresponding κ-th DCI (8) to output iterative $(\tilde{\xi}^{(\kappa+1)}, q^{(\kappa+1)})$ in DCIs is the following convex program

$$\begin{array}{c} \min_{\tilde{\boldsymbol{\xi}}, \boldsymbol{q}} [f_1(\tilde{\boldsymbol{\xi}}) - f_2(\tilde{\xi}^{(\kappa)}, q^{(\kappa)}) \\ -\langle \nabla f_2(\tilde{\xi}^{(\kappa)}, q^{(\kappa)}), (\tilde{\boldsymbol{\xi}} - \xi^{(\kappa)}, \boldsymbol{q} - q^{(\kappa)}) \rangle] \ : \ (58), \end{array} \qquad (61)$$

where

$$\langle \nabla f_2(\tilde{\xi}^{(\kappa)}, q^{(\kappa)}), (\tilde{\boldsymbol{\xi}} - \tilde{\xi}^{(\kappa)}, \boldsymbol{q} - q^{(\kappa)}) \rangle = \sum_{m=1}^{M} [\frac{2\mathrm{Re}\{(\sum_{n=1}^{N} \overline{\tilde{\xi}_n^{(\kappa)} \ell_{nm} h_{nm}}) \sum_{n=1}^{N} \ell_{nm} h_{nm} (\tilde{\boldsymbol{\xi}}_n - \tilde{\xi}_n^{(\kappa)})\}}{q_m^{(\kappa)} + \sigma_D^2} - \frac{|\sum_{n=1}^{N} \tilde{\xi}_n^{(\kappa)} \ell_{nm} h_{nm}|^2 (\boldsymbol{q}_m - y_m^{(\kappa)})}{(q_m^{(\kappa)} + \sigma_D^2)^2}] \qquad (62)$$

The computational complexity of convex program (61) is $O((N+M)^3(4N+M+2)$. Note that the effective capacity of the κ-th DCI (61) for locating the approximately global optimal solutions of (60) has been shown in [30].

B. Step 2: Mixed beamforming and assignment optimization

Suppose $\tilde{\xi}^*$ is found from Step 1 and then $x_n^* = |\tilde{\xi}_n^*|^2 (\sum_{m=1}^{M} |h_{nm}|^2 + \sigma_R^2)/P_T$, $n = 1, 2, \ldots, N$, while q_n^* is defined by the left hand side of (56b) at $\tilde{\xi}^*$. Such $(x^*, \tilde{\xi}^*, q^*)$ is obviously feasible to (57). Initialized by $(x^{(0)}, \tilde{\xi}^{(0)}, q^{(0)}) = (x^*, \tilde{\xi}^*, q^*)$, the corresponding κ-th DCI (15) for program (59) to output iterative solution $(x^{(\kappa+1)}, q^{(\kappa+1)})$ for $\kappa = 0, 1, \ldots,$ is

$$\begin{array}{c} \min_{\boldsymbol{x}, \tilde{\boldsymbol{\xi}}, \boldsymbol{q}} [f_1(\tilde{\boldsymbol{\xi}}, \boldsymbol{q}) + \mu h_1(\boldsymbol{x}) - f_2(\tilde{\xi}^{(\kappa)}, q^{(\kappa)}) \\ -\mu h_2(x^{(\kappa)}) - \langle \nabla f_2(\tilde{\xi}^{(\kappa)}, q^{(\kappa)}), (\tilde{\boldsymbol{\xi}}, \boldsymbol{q}) - (\tilde{\xi}^{(\kappa)}, q^{(\kappa)}) \rangle \\ -\mu \langle \nabla h_2(x^{(\kappa)}), \boldsymbol{x} - x^{(\kappa)} \rangle] \quad : \quad (58), \end{array} \qquad (63)$$

where $\langle \nabla f_2(\tilde{\xi}^{(\kappa)}, q^{(\kappa)}), (\tilde{\boldsymbol{\xi}}, \boldsymbol{q}) - (\tilde{\xi}^{(\kappa)}, q^{(\kappa)}) \rangle$ is defined by (62), while

$$\langle \nabla h_2(x^{(\kappa)}), \boldsymbol{x} - x^{(\kappa)} \rangle = 2 \sum_{n=1}^{N} (x_n^{(\kappa)} + \sum_{n=1}^{N} x_n^{(\kappa)})(\boldsymbol{x}_n - x_n^{(\kappa)}).$$

Its computational complexity is $O((NM+M)^3(N+NM+2M))$.

C. Step 3: Assigned relay re-optimization

Suppose $x^* \in [0,1]^N$ is the optimized solution found from Step 2. The corresponding program (16) for assigned beamforming re-optimization is

$$\begin{array}{c} \max_{\boldsymbol{\xi}, \boldsymbol{q}} \min_{m=1,2,\ldots,M} \varphi_m(\boldsymbol{\xi}, \boldsymbol{q}) \quad \text{s.t.} \\ ([x_1^*]_b \boldsymbol{\xi}_1, \ldots, [x_N^*]_b \boldsymbol{\xi}_N, \boldsymbol{q}, [x^*]_b) \in \mathcal{K}, \end{array} \qquad (64)$$

with $\mathcal{K}$ defined by (58), which is similar to program (60) and thus can be solved by the same DCIs for program (60). In fact, program (64) involves only N_R variables $\boldsymbol{\xi}_n$ corresponding to $[x_n^*]_b = 1$.

In summary, our proposed solution procedure for the mixed binary program (56)/(59) for joint optimization in beamforming weights $\boldsymbol{\xi}$ and relay assignment binary variable $\boldsymbol{x}$ consists of three steps: Based on the initial solution founded by Step 1, Step 2 aims at jointly optimizing continuous variables $\boldsymbol{\xi}$ and binary variables $\boldsymbol{x}$ through d.c. approximation and optimization, while Step 3 re-optimizes the beamforming weight $\boldsymbol{\xi}$.

Before closing this section, it is pointed out again that the formulations (26) and (54) for orthogonal and nonorthogonal transmissions are based on the availability of perfect channel state information of both communication links connected to the relays. Similar to [30], formulations similar to (26) and (54) but taking into account channel uncertainties are possible. Such design problems and solutions would likely require much more computational effort and deserve a separate study.

V. NUMERICAL RESULTS

For simulation, consider a $200\text{m} \times 200\text{m}$ square region as illustrated in Figure 1, where N relays are fixed at coordinates $(0, \frac{200}{N+1})$, $(0, 2 \times \frac{200}{N+1})$,..., $(0, N \times \frac{200}{N+1})$ and M source nodes (destination nodes, resp.) are randomly allocated on the left hand side (right hand side, resp.) of the relay array.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2014.2324588, IEEE Transactions on Wireless Communications

Following [11], [57], fading channels with path loss and a Rayleigh component are generated by $h_{nm} \sim \mathcal{N}(0, \sigma^2_{hnm})$ and $\ell_{nm} \sim \mathcal{N}(0, \sigma^2_{\ell nm})$, where their variances are proportional to the physical link distance [58]: $\sigma^2_{hnm} = C/d^{\beta}_{hnm}$ and $\sigma^2_{\ell nm} = C/d^{\beta}_{\ell nm}$. Here, d_{hnm} ($d_{\ell nm}$, resp.) denotes the distance between the mth source node and the nth relay node in meters (the nth relay node and the mth destination node, resp.). The constant C is given as $C = G_t G_r \eta^2/(4\pi)^2 L$, where G_t is the transmitter antenna gain, G_r is the receiver antenna gain, η is the wavelength in meters, $L \leq 1$ is the system loss factor not related to propagation and β is the path loss exponent. An urban environment with $\beta = 3, G_t = G_r = 1, \eta = 1/3m, L = 1$ is used in all simulations. In addition, the variances of AWGN noise samples at relays and destination nodes are assumed to be 10^{-10}. The value of penalty parameter μ in program (33) of orthogonal source transmission and program (59) for nonorthogonal source transmission is set to be 50 and 10, respectively. In all simulations, the proposed algorithm are indeed capable of locating the approximately global optimal solutions of the mixed binary programs (26) and (54) as they output the optimized values that basically match their upper bounds provided by the optimal values of the relaxed programs (34) and (60), respectively.

A. Orthogonal Source Transmission

For ease of presentation and discussion, we refer to our proposed solution procedure (Step 1 to Step 3 in Section III) as Procedure I. To show the effectiveness of its jointly optimized beamforming and relay assignment capability we also compare its minimum SNR performance with that obtained by the following existing heuristic procedures for the relay assignment:

- Procedure II: Based on the solution $\bar{\xi}^*$ of the box-relaxed program (34) (found by Step 1), for each user m, assign $x^*_{nm} = 1$ to N_R relays corresponding to the N_R highest beamforming gains $\bar{\xi}^*_{nm}$.
- Separate best link selections for each user: (i) "Best S-D channels" selects N_R relays corresponding to the N_R strongest products of uplink and down link, i.e., maximum $|h_{nm}\ell_{nm}|$; (ii) "Best S-R channels" selects N_R relays corresponding to the N_R strongest uplinks, i.e., maximum $|h_{nm}|$; (iii) "Best R-D channels" selects N_R relays corresponding to the N_R strongest downlinks, i.e., maximum $|\ell_{nm}|$.

With these relay assignments, program (40) is then implemented for re-optimization to output the corresponding minimum SNR values.

We start with the single-user case ($M = 1$) in which $N_R = 3$ relays assigned from $N = 5$ and $N = 10$ relays. Figure 2 plots the minimum SNR versus the relay power limit P_n. The total power $P_T = 0.7 N_R P_n$ is set so it is independent from the total relay number N. The proposed Procedure I yields performance curves that almost match the All-relay BF curves of the upper bounds provided by the relaxed program (34). It is clearly observed that the performance of Procedure II is inferior to that of Procedure I by up to 2 dB, although it is still better than performances by the above three mentioned best link selections. The communication coverage is better provided with a larger number of relays, which explains the improvement of all performance curves when increasing N from 5 to 10.

Next, for the multi-user case, $(M, N, N_R) = (5, 10, 3)$, Figure 3 plots the minimum SNR versus the relay power limit P_n, where, as before, the total power is $P_T = 0.7 N_R P_n$. Similarly, the proposed Procedure I outputs the minimum SNRs that match their upper bounds provided by the relaxed program (34). Procedure II performs poorer than Procedure I but it still outperforms all the channel strength-based selection algorithms. It is important to realize that under the same relaying power constraints, a properly selected relay subset for each user is able to retain the minimum SINR performance of the conventional nonselective scheme and thus actually improves the bandwidth efficiency and the life time of the network. Next, with the number of user pairs $M = 5$ and the individual relay power limit $P_n = 0$dB, Figure 4 plots the minimum SNR versus the total relay numbers N with

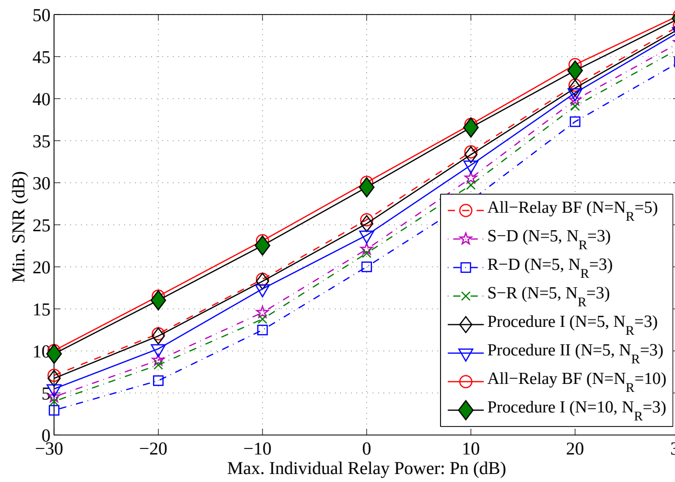

Fig. 2. Minimum SNR versus individual relay power for $M = 1$, $N_R = 3$ and $N \in \{5, 10\}$.

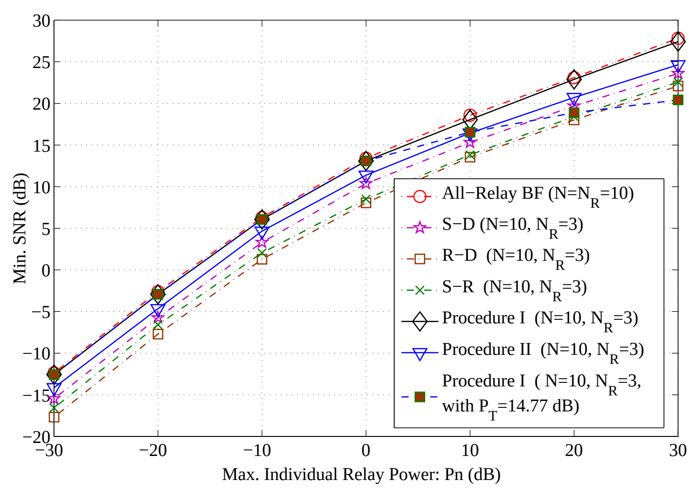

Fig. 3. Minimum SNR versus individual relay power for $(M, N, N_R) = (5, 10, 3)$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2014.2324588, IEEE Transactions on Wireless Communications

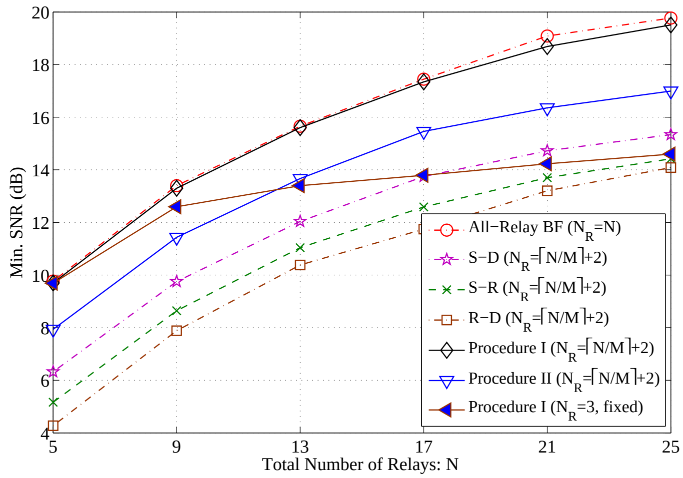

Fig. 4. Minimum SNR versus the number of relays for $M = 5$, $N_R = \lceil \frac{N}{M} \rceil + 2$

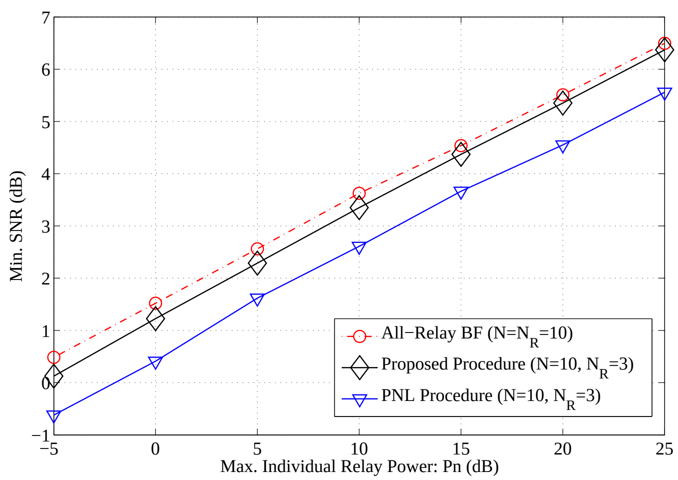

Fig. 5. Minimum SNR versus the individual relay power for MRC receiver: $(M, N, N_R) = (5, 10, 3)$.

$N_R = \lceil \frac{N}{M} \rceil + 2$, where $\lceil \cdot \rceil$ is the ceiling function. Again, the figure clearly shows that only Procedure I is able to perform closely with the upper bound provided by the optimal value of (34). The given choice of selected relays N_R seems to be reasonable. On the contrary, with $N_R = 3$ fixed, the SNR performance becomes saturated although the total number of relays is increased to improve the communication coverage.

The computational complexity of Procedure I is provided in Table I. The iteration numbers for different versions of Step 1 correspond to the number of required SOCP programs for locating solutions of the parametric program (37) and the number of required convex programs (35) for locating solutions of the d.c. program (34), respectively. It is found that both can yield similar optimized solutions, while the former has a guaranteed optimality given some accuracy tolerance (set as 10^{-6}). However, the latter (35) converges even faster to those well optimized solutions. Similarly, the iteration numbers for Step 2 (Step 3, resp.) correspond to the number of required convex programs (39) ((41) or (42), resp.) for finding the solution of program (33) ((40), resp.).

Figure 5 shows the corresponding simulation results for the MRC program (43). Again, the corresponding Procedure I could locate the best optimized solution for (43) as its performance curve is very close to the upper bound. However, the procedure proposed in [26], referred to as "PNL Procedure" (discussed at the end of Section III) has a poorer performance.

B. Non-orthogonal Source Transmission

By Procedure I we refer to the solution procedure proposed in Section IV. Yet another procedure is included for comparison with the proposed Procedure I (still called as Procedure II). Based on the optimized beamforming ξ^* of the box-relaxed program (60) (found by Step 1 in Section IV) Procedure II solves the following binary program for the relay assignment by enumerating all feasible relay assignments $\boldsymbol{x}$

$$\max_{\boldsymbol{x}\in\{0,1\}^N} \ |\sum_{n=1}^{N} \boldsymbol{x}_n \xi_n^* \ell_{nm} h_{nm}|^2 \ : \ \sum_{n=1}^{N} \boldsymbol{x}_n = N_R. \qquad (65)$$

Figure 6 shows the minimum SINR performance versus individual relaying power limit P_n, where the total relaying power $P_T = 0.7 N_R P_n$ is set. For two network settings of $(N, M, N_R) = (25, 5, 20)$ and $(N, M, N_R) = (30, 5, 20)$, the performance of the proposed Procedure I well approximates the upper bound found by the relaxed program (60). It is clear from Figure 6 that the optimized values of both (54) and (60) increase with the individual relaying power P_n.

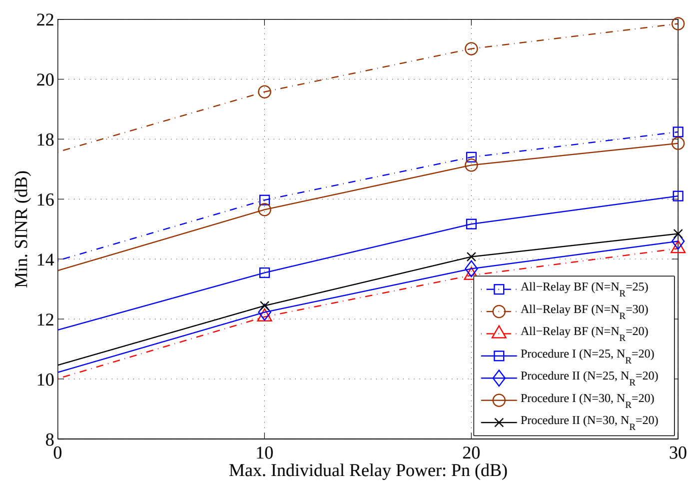

Fig. 6. Minimum SINR versus maximum individual relay power: $M = 5$, $N_R = 20$.

Unlike the orthogonal source transmission, for which relays can be individually deactivated user-wise in its up/down links and thus the size of the relay network is not necessarily reduced, it is important to realize that the performance for the non-orthogonal source transmission is indeed more sensitive on the number N_R of active relays. This contributes to the gap between the performance plots for (54) and (60). Nevertheless, it is seen that the SINR performance can be significantly improved by choosing a fixed number of relays from a larger set of relays, which clearly justifies the advantage of relay assignment. The figure also shows that the performance

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2014.2324588, IEEE Transactions on Wireless Communications

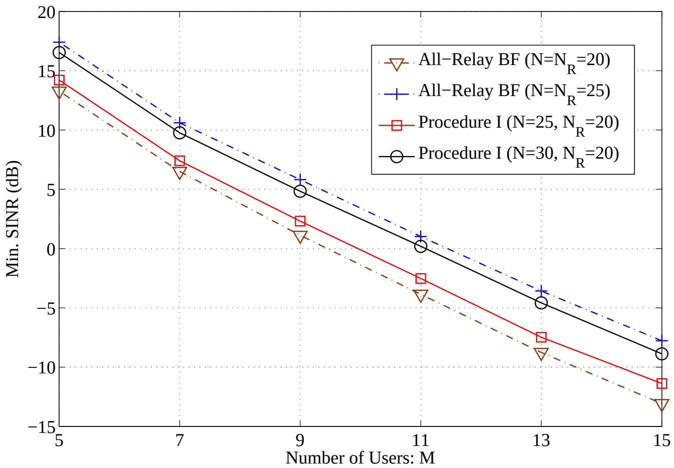

Fig. 7. Minimum SINR versus number of users: $N_R = 20$, $P_n = 20$dB, $P_T = 31.46$dB

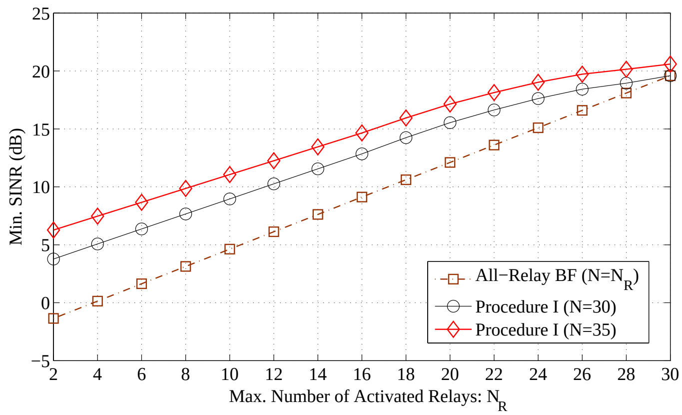

Fig. 8. Minimum SINR versus maximum number of activated relays: $M = 5$, $P_n = 10$dB, $P_T = 0.7P_nN_R$

of Procedure II (separated relay assignment) is much less favorable.

Furthermore, we analyze the SINR performances versus the user number M under fixed relay number $(N, N_R) = (25, 20)$ or $(N, N_R) = (30, 20)$ and power constraints of $P_n = 20$ dB and $P_T = 31.46$ dB. It can be seen in Figure 7 that the minimum SINR performance is degraded as more users share the same relay resource. However, Procedure I is still able to improve the performance by as much as 5 dB against the bottom dash-line, which represents the performance of the nonselective scheme.

The computational experience is provided in Table II. Like Table I, the most computational loads are due to Step 1 (for locating a good enough initial solution) and Step 3 (for beamforming re-optimization).

Finally, the performance difference between selective and nonselective beamforming is further analyzed by Figure 8. With the same number N_R of activated relays, the performance advantage of selective beamforming over nonselective beamforming is clearly observed. It is again found that the performance of selective beamforming is improved by choosing N_R relays from a larger set of N candidate relays. This not only reiterates the above-mentioned findings, but also clearly shows the superior performance of Procedure I over other procedures in the most crucial step of relay assignment.

Table I

THE AVERAGED ITERATION NUMBERS FOR ORTHOGONAL TRANSMISSIONS: $M = 5, P_n = 0$ dB

N	5	9	13	17	21	25
Step 1 by SOCP	12.03	12.38	13.27	13.94	14.14	14.37
Step 1 by d.c.	4.52	5.92	7.61	7.97	10.23	11.06
Step 2 by d.c.	2.11	2.91	4.02	4.81	6.77	8.05
Step 3 by SOCP	12.00	12.10	12.46	13.28	13.74	13.66
Step 3 by d.c.	6.78	6.94	8.06	9.12	10.83	11.98

Table II

THE AVERAGED ITERATION NUMBERS FOR NON-ORTHOGONAL TRANSMISSIONS: $P_n = 20$ dB, $P_T = 31.46$ dB, $N_R = 20$

M		5	7	9	11	13	15
Step 1	$(N = 25)$	15.10	16.32	19.67	23.09	24.14	27.92
	$(N = 30)$	18.54	20.77	21.94	25.01	27.35	29.01
Step 2	$(N = 25)$	3.82	4.02	4.92	5.91	7.88	9.45
	$(N = 30)$	5.04	6.23	7.51	8.44	10.09	12.61
Step 3	$(N = 25)$	18.25	20.40	22.99	24.12	25.81	27.27
	$(N = 30)$	19.41	22.35	26.02	29.40	31.31	34.87

VI. CONCLUSIONS

The joint optimization in beamforming and relay assignment is a hard mixed combinatoric program. It has been shown in this paper that such a problem can be solved very efficiently by d.c iterations (DCIs) of d.c. programming, which are of local search in nature and thus can converge quickly, but nevertheless yield globally optimized solutions. Extensive numerical results have demonstrated the viability of our DCIs as well as their superiority performance over other iterative procedures.

APPENDIX A: PROOF OF PROPOSITION 2

Suppose $\varphi(\mu)$ and (z^μ, x^μ) are the optimal value and optimal solution of the convex constrained program (7), so

$$\begin{aligned}\sup_{\mu\geq 0}\varphi(\mu) &= \sup_{\mu\geq 0}\min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}}\mathcal{L}(\boldsymbol{z},\boldsymbol{x},\mu)\\ &\leq \min_{(\boldsymbol{z},\boldsymbol{x})\in\mathcal{D}}\max_{\mu\geq 0}\mathcal{L}(\boldsymbol{z},\boldsymbol{x},\mu)\\ &= \min\ (5). \end{aligned} \tag{66}$$

Note that $\sum_{n=1}^{N}\boldsymbol{x}_n - \sum_{n=1}^{N}\boldsymbol{x}_n^2 \geq 0\ \forall(\boldsymbol{z},\boldsymbol{x}) \in \mathcal{D}$ so function $\mathcal{L}(\boldsymbol{z},\boldsymbol{x},\mu)$ is increasing in μ for $(\boldsymbol{z},\boldsymbol{x}) \in \mathcal{D}$. This means $\varphi(\mu)$ is increasing in μ and bounded by the optimal value of program (5). If $\sum_{n=1}^{N} x_n^{\mu_0} - \sum_{n=1}^{N}(x_n^{\mu_0})^2 = 0$ for some $0 \leq \mu_0 < +\infty$ then (z^{μ_0}, x^{μ_0}) is feasible to (5), so

$$\varphi(\mu_0) = \mathcal{L}(z^{\mu_0}, x^{\mu_0}, \mu_0) = f(z^{\mu_0}) - g(z^{\mu_0}) \geq \min\ (5)$$

which together with (66) imply (6) and moreover $\varphi(\mu_0) = \sup_{\mu\geq 0}\varphi(\mu)$ as well as

$$\varphi(\mu) = f(z^{\mu_0}) - g(z^{\mu_0}) = \min\ (5)\quad \forall\ \mu \geq \mu_0, \tag{67}$$

proving the second statement of the Proposition.

Now, suppose that $\sum_{n=1}^{N} x_n^{(\mu)} - \sum_{n=1}^{N}(x_n^{(\mu)})^2 > 0$ for all $\mu > 0$. The sequence $\{(z^\mu, x^\mu)\}$ is bounded and by taking a subsequence if necessary one can assume $(z^\mu, x^{(\mu)}) \to$

$(z^{\infty}, x^{(\infty)}) \in \mathcal{D}$ with $\sum_{n=1}^{N} x_n^{(\infty)} - \sum_{n=1}^{N} (x_n^{(\infty)})^2 = 0$, because otherwise $\varphi(\mu) = f(z^{(\mu)}) - g(z^{(\mu)}) + \mu(\sum_{n=1}^{N} x_n^{(\mu)} - \sum_{n=1}^{N} (x_n^{(\mu)})^2) \to +\infty$, a contradiction. This means (z^{∞}, x^{∞}) is feasible to (5) and

$$\sup_{\mu \geq 0} \varphi(\mu) = f(z^{\infty}, x^{\infty}) - g(z^{\infty}, x^{\infty}) \geq \min\ (5),$$

which together with (66) yield (6). Furthermore, if the supremum of the right hand side of (6) attains at μ_0 then the second statement of the Proposition also follows by noticing that $f(z^{\infty}, x^{\infty}) - g(z^{\infty}, x^{\infty}) = \varphi(\mu_0)$ so (z^{∞}, x^{∞}) is the optimal solution of (7).

REFERENCES

[1] R. Pabst et al, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Communication Magazine*, vol. 42, pp. 80–89, Sept. 2004.

[2] Y. Liu, R. Hoshyar, X. Yang, and R. Tafazolli, "Integrated radio resource allocation for multihop cellular networks with fixed relay stations," *IEEE J. on Selected Areas in Communications*, vol. 24, pp. 2137–2146, Nov. 2006.

[3] L. Le and E. Hossain, "Multihop cellular networks: potential gains, research challenges, and a resource allocation framework," *IEEE Communication Magazine*, vol. 45, pp. 66–73, Sept. 2007.

[4] Y. Jing and H. Jafarkhani, "Single and multiple relay selection schemes and their achievable diversity orders," *IEEE Trans. Wireless Communications*, vol. 8, pp. 1414–1423, Mar. 2009.

[5] Y. Zhao, R. S. Adve, and T. J. Lim, "Improving amplify-and-forward relay networks: optimal power allocation versus selection," *IEEE Trans. Wireless Communications*, vol. 6, pp. 3114–3123, Aug 2007.

[6] M. Ju, H.-K. Song, and I.-M. Kim, "Joint relay-and-antenna selection in multi-antenna relay networks," *IEEE Trans. Communications*, vol. 58, pp. 3417–3422, Dec. 2010.

[7] Y. Li, Q. Yin, W. Su, and H.-M. Wang, "On the design of relay selection strategies in regenerative cooperative networks with outdated CSI," *IEEE Trans. Wireless Communications*, vol. 10, pp. 3086–3097, Sep. 2011.

[8] A. Ikhlef, D. S. Michalopoulos, and R. Schober, "Max-max relay selection for relays with buffers," *IEEE Trans. Wireless Communications*, vol. 11, pp. 1124–1135, Mar. 2012.

[9] D. S. Michalopoulos, H. A. Suraweera, G. K. Karagiannidis, and R. Schober, "Amplify-and-foward relay selection with outdated channel estimates," *IEEE Trans. Communications*, vol. 60, pp. 1278–1290, May 2012.

[10] S. S. Soliman and N. C. Beaulieu, "Exact analysis of dual-hop AF maximum end-to-end SNR relay selection," *IEEE Trans. Communications*, vol. 60, pp. 2135–2145, Aug. 2012.

[11] J. Laneman, D. Tse, and G. Wornel, "Cooperative diversity in wireless networks: effcient protocols and outage behavior," *IEEE Trans. Information Theory*, vol. 50, pp. 3062–3080, Dec 2004.

[12] X. Deng and A. M. Haimovich, "Power allocation for cooperative relaying in wireless network," *IEEE Communication Letters*, vol. 9, pp. 994–996, Nov. 2005.

[13] Y. Liang and V. Veeravalli, "Gaussian orthogonal relay channel: optimal resourse allocation and capacity," *IEEE Trans. Information Theory*, vol. 51, pp. 3284–3289, Sept 2005.

[14] A. H. Madsen and J. Zhang, "Capacity bounds and power allocation for cooperative strategies in Gaussian relay networks," *IEEE Trans. Information theory*, vol. 51, pp. 2020–2040, June 2005.

[15] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *IEEE Trans. Information Theory*, vol. 55, pp. 2499–2511, Jun. 2009.

[16] I. Ahmed, A. Nasari, D. S. Michalopoulos, R. Schober, and R. K. Mallik, "Relay subset selection and fair power allocation for best and partial relay selection in generic noise and interference," *IEEE Trans. Wireless Communications*, vol. 11, pp. 1828–1939, May 2012.

[17] V. I. Morgenshtern and H. Bolcskei, "Crystallization in large wireless networks," *IEEE Trans. Information Theory*, vol. 53, pp. 3319–3349, Oct. 2007.

[18] D. H. N. Nguyen, H. H. Nguyen, and H. D. Tuan, "Distributed beamforming in relay-assisted multiuser communications," in *Proc. of IEEE International Conference on Communications (ICC), Dresden, Germany*, June 2009.

[19] S. Fazeli-Dehkordy, S. Shahbazpanahi, and S. Fazor, "Multiple peer-to-peer communications using a network of relays," *IEEE Trans. Signal Processing*, vol. 57, pp. 3053–3062, Aug 2009.

[20] L. Dong, A. P. Petropulu, and H. V. Poor, "Weighted cross-layer cooperative beamforming for wireless networks," *IEEE Trans. Signal Processing*, vol. 57, pp. 3240–3252, Aug. 2009.

[21] A. Phan, H. D. Tuan, H. H. Kha, and H. H. Nguyen, "Nonsmooth optimization-based beamforming in multiuser wireless relay networks," in *Proc. of 4th International Conference on Signal Processing and Communication Systems (ICSPCS), Gold Coast, Australia*, Dec. 2010.

[22] A. H. Phan, H. D. Tuan, and H. H. Kha, "Space-time beamforming for multiuser wireless relay networks," in *Proc. of International Conference on Acoustics, Speech and Signal Processing(ICASSP), Prague, Czech*, May 2011.

[23] A. H. Phan, H. D. Tuan, and H. H. Kha, "Optimized solutions for beamforming problems in amplify-and-forward wireless relay networks," in *Proc. of Global Communication Conference (Globecom), Houston, TX, USA*, Dec. 2011.

[24] E. Koyuncu and H. Jafarkhani, "Distributed beamforming in wireless multiuser relay-interference networks with quantized feedback," *IEEE Trans. Information Theory*, vol. 58, pp. 4538–3576, Jul. 2012.

[25] B. Wang, Z. Han, and K. J. R. Liu, "Distributed relay selection and power control for multi-user cooperative communication networks using Stackelberg game," *IEEE Trans. Mobile Computing*, vol. 8, pp. 975–990, Jul 2009.

[26] K. T. Phan, D. H. N. Nguyen, and T. Le-Ngoc, "Joint power allocation and relay selection in cooperative networks," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM), Honolulu, Hawaii, USA*, Dec. 2009.

[27] G. Zheng, Y. Zhang, C. Li, and K.-K. Wong, "A stochastic optimization approach for joint relay assignment and power allocation in orthogonal amplify-and-forward cooperative wireless networks," *IEEE Trans. Wireless Communications*, vol. 12, pp. 4091–4099, Dec. 2011.

[28] S. Mallick, M. R. Rashid, and V. K. Bhargava, "Joint relay selection and power allocation for decode-and-forward cellular relay network with imperfect CSI," in *Proc. of IEEE Globecom*, 2011.

[29] H. Fang, X. Lin, and T. M. Lok, "Power allocation for multiuser cooperative communication networks under relay selection degree bound," *IEEE Trans. Vehicle Technology*, vol. 61, pp. 2991–3001, Sept. 2012.

[30] U. Rashid, H. D. Tuan, and H. H. Nguyen, "Relay beamforming in multi-user wireless relay networks based on throughput maximin optimization," *IEEE Trans. Commun.*, vol. 61, pp. 1739–1749, 2013.

[31] P. Apkarian and H. D. Tuan, "Robust control via concave optimization: local and global algorithms," *IEEE Trans. on Automatic Control*, vol. 45, pp. 299–305, Feb. 2000.

[32] P. Apkarian and H. D. Tuan, "Concave programming in control theory," *J. of Global Optimization*, vol. 15, pp. 243–270, Apr. 1999.

[33] H. D. Tuan, P. Apkarian, S. Hosoe, and H. Tuy, "D.c. optimization approach to robust controls: the feasibility problems," *International J. of Control*, vol. 73, pp. 89–104, Feb. 2000.

[34] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Fast global optimal power allocation in wireless networks by local d.c. programming," *IEEE Trans. Wireless Communications*, vol. 11, pp. 510–515, Feb. 2012.

[35] H. H. Kha, H. D. Tuan, H. H. Nguyen, and T. Pham, "Optimization of cooperative beamforming for SC-FDMA multi-user multi-relay networks by tractable d.c. programming," *IEEE Trans. Signal Processing*, vol. 61, pp. 467–479, 2012.

[36] H. Tuy, *Convex Analysis and Global Optimization*. Kluwer Academic, 1998.

[37] H. Konno, P. T. Thach, and H. Tuy, *Optimization on low rank nonconvex structure*. Kluwer Academic, 1997.

[38] H. Tuy, M. Minoux, and N. T. H. Phuong, "Discrete monotonic optimization with application to a discrete location problem," *SIAM Journal on Optimization*, vol. 17, no. 1, pp. 78–97, 2006.

[39] H. D. Tuan, T. T. Son, H. Tuy, and P. T. Khoa, "Monotonic optimization based decoding for linear codes," *J. of Global Optimization*, vol. 55, pp. 301–312, 2013.

[40] H. D. Tuan, T. T. Son, H. Tuy, and H. H. Nguyen, "Optimum multi-user detection by nonsmooth optimization," in *Proc. of IEEE International*

Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, May. 2011.

[41] M. R. Bhatnagar and Arti M. K., "Selection beamforming and combining in decode-and-forward MIMO relay networks," *IEEE Comm. Letters*, vol. 17, pp. 1556–1559, Nov. 2013.

[42] S. A. Jafar, *Interference alignment: a new look at signal dimensions in a communication network*. Now Publisher Inc, 2011.

[43] N. N. Tran, H. H. Nguyen, H. D. Tuan, and D. E. Dodds, "Training designs for amplify-and-forward relaying with spatially correlated antennas," *IEEE Trans. Vehicular Technology*, vol. 61, pp. 2864–2870, Jul. 2012.

[44] Arti M. K., R. Bose, M. R. Bhatnagar, and A. Hjorungnes, "Relay strategies for high rate space-time code in cooperative MIMO networks," in *Proc. IEEE international symposium on wireless communication systems (ISWCS)*, 2007.

[45] M. R. Bhatnagar, Arti M. K., A. H. R. Bose, and L. Song, "A time efficient multi- user relay strategy for cooperative MIMO networks," in *Proc. IEEE international symposium on wireless personal multimedia communications (WPMC)*, 2007.

[46] M. R. Bhatnagar, Arti M. K., A. Hjorungnes, R. Bose, and L. Song, "Multi-user relaying of high-rate space-time code in cooperative networks," *Wireless Personal Communications*, vol. 54, pp. 69–81, 2010.

[47] T. Pham-Dinh, N. Nguyen-Canh, and H. A. Le-Thi, "An efficient combined DCA and B&B using DC/SDP relaxation for globally solving binary quadratic programs," *J. of Global Optimization*, vol. 48, pp. 595–632, 2010.

[48] H. Tuy, "Robust solution of nonconvex global optimization problems," *J. of Global Optimization*, vol. 32, pp. 307–323, Feb 2005.

[49] H. Tuy and N. T. Hoai-Phuong, "A robust algorithm for quadratic optimization under quadratic constraints," *J. of Global Optimization*, vol. 37, pp. 557–596, Apr. 2007.

[50] H. Tuy, "A new minimax theorem with applications," *J. of Global Optimization*, vol. 50, no. 3, pp. 371–378, 2011.

[51] H. A. Le-Thi and T. Pham-Dinh, "DC optimization algorithms for solving the trust region subproblem," *SIAM J. on Optimization*, vol. 8, pp. 476–505, Apr. 1998.

[52] A. H. Phan, H. D. Tuan, H. H. Kha, and D. T. Ngo, "Non-smooth optimization for efficient beamforming in cognitive radio multicast transmission," *IEEE Trans. Signal Processing*, vol. 60, pp. 2941–2951, Jun. 2012.

[53] A. H. Phan, H. D. Tuan, H. H. Kha, and H. H. Nguyen, "Beamforming optimization in multi-user amplify-and-forward wireless relay networks," *IEEE Trans. Wireless Communications*, vol. 11, pp. 1510–1520, Apr. 2012.

[54] H. Tuy and H. D. Tuan, "Generalized S-lemma and strong duality in nonconvex quadratic programming," *J. of Global Optimization*, vol. 56, pp. 1045–1072, 2013.

[55] J. C. Dunn, "Rates of convergence for conditional gradients algroritms," *SIAM J. on Control and Optimization*, vol. 17, pp. 187–121, Jan. 1979.

[56] P. Gahinet, A. Nemirovski, A. J. Laub, and M. Chilali, *LMI Control Toolbox User's Guide*. The Mathworks Partner Series, 1995.

[57] M. Chen, S. Serbetli, and A. Yener, "Distributed power allocation strategies for parallel relay networks," *IEEE Trans. Wireless Communications*, vol. 7, pp. 552 –561, Feb. 2008.

[58] T. S. Rappaport, *Wireless Communications: Principles and Practice.* Upper Saddle River, NJ. Prentice-Hall, 1996.

Enlong Che was born in Urumqi, China. He received the B.S. degree in electrical engineering from Zhejiang University, Hangzhou, China, in 2009. Currently, he is pursuing the Ph.D. degree at the University of Technology, Sydney, Australia. His research interests include convex optimization, wireless communication and signal processing. During his graduate studies, he has been working on optimized transmission strategies for wireless relay communication and multi-cell multi-user networks.

Hoang Duong Tuan (M'94) received the diploma (Hon.) and the PhD degrees, both in applied mathematics, from Odessa State University, Ukraine, in 1987 and 1991, respectively. He spent nine academic years in Japan as an Assistant Professor in the Department of Electronic-Mechanical Engineering, Nagoya University from 1994 to 1999, and then as an Associate Professor in the Department of Electrical and Computer Engineering, Toyota Technological Institute, Nagoya from 1999 to 2003. He has been a Professor in the School of Electrical Engineering and Telecommunications, from 2003 to 2011. He is currently a Professor of Centre for Health Technologies, University of Technology, Sydney. He has been involved in research with the areas of optimization, control, signal processing, wireless communication and bio-informatics for 20 years.

Ha H. Nguyen (M'01-SM'05) received the B.Eng. degree from the Hanoi University of Technology (HUT), Hanoi, Vietnam, in 1995, the M.Eng. degree from the Asian Institute of Technology (AIT), Bangkok, Thailand, in 1997, and the Ph.D. degree from the University of Manitoba, Winnipeg, MB, Canada, in 2001, all in electrical engineering. He joined the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK, Canada, in 2001, and became a full Professor in 2007. He holds adjunct appointments at the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada, and TRLabs, Saskatoon, SK, Canada, and was a Senior Visiting Fellow in the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, Australia during October 2007-June 2008. His research interests include spread spectrum systems, error-control coding and diversity techniques in wireless communications. Dr. Nguyen was an Associate Editor for the IEEE Transactions on Wireless Communications during 2007-2011. He currently serves as an Associate Editor for the IEEE Transactions on Vehicular Technology and the IEEE Wireless Communications Letters. He was a Co-chair for the Multiple Antenna Systems and Space-Time Processing Track, IEEE Vehicular Technology Conferences (Fall 2010, Ottawa, ON, Canada and Fall 2012, Quebec, QC, Canada).