Minimum Power Multicast Beamforming With Superposition Coding for Multiresolution Broadcast and Application to NOMA Systems

Jinho Choi, *Senior Member, IEEE*

***Abstract*—Multicast beamforming with superposition coding (SC) is studied for multiresolution broadcast where both data streams of high priority (HP) and low priority (LP) are to be transmitted for a user close to a base station (BS), while only data stream of LP is to be transmitted to a user not close to the BS (e.g., a cell-edge user). Using SC, a minimum total transmission power beamforming problem has been formulated to find beamforming vectors and powers for both users. For given normalized beamforming vectors, a closed-form expression for the optimal power allocation is derived from which an iterative algorithm is considered to find beamforming vectors. The proposed multicast beamforming with SC is applied to nonorthogonal multiple access (NOMA) systems to support multiple users as a two-stage beamforming method.**

***Index Terms*—Superposition coding, multiresolution broadcast, multicast beamforming.**

I. Introduction

SUPERPOSITION coding (SC) for broadcast channels was considered to achieve the capacity of broadcast channels [1]. Since then, SC has been extensively studied in the area of information theory, while it has been recently considered in wireless systems to improve the spectral efficiency [2], [3]. In wireless communications, hierarchical modulation can be seen as an example of SC, which is used in various systems, e.g. broadcast systems [4]–[6], cooperative transmissions [7], and two-way relay networks [8]. In multiresolution broadcast such as in [4], [5], multiple signals can be transmitted using SC, where a user near to a transmitter can decode more signals, while a user far from the transmitter can only decode a few signals. Thus, the received signal quality can be gradually degraded as the distance between a user and the transmitter increases. That is, a user near to the transmitter can have a high resolution signal (or a signal of good quality), while a user far from the transmitter can only have a low resolution signal (or a signal of poor quality).

For broadcast/multicast[1] services in wireless communications, beamforming can also be used for effective transmissions. In [9], [10], multicast beamforming is studied where a beam can be used to a group of users who need to a common signal (i.e., a broadcast signal). In [10], multiple user groups are considered with multiple beams where each group is to receive different signals, while a single beam is considered in [9] as all users are to receive a common signal.

In this paper, we consider multicast beamforming with SC for multiresolution broadcast where two different broadcast signal streams exist. Among two users (one user is close to a base station (BS) and the other is not), the user close to the BS can receive two different data streams, while the other user can only receive one data stream. A minimum power multicast beamforming problem with SC is formulated. Unfortunately, since the problem is not a standard optimization problem, we cannot obtain optimal beamforming vectors as a closed-form solution. Thus, we first obtain a closed-form expression for the optimal power allocation for given beamforming vectors. Then, an iterative algorithm is derived to find beamforming vectors for two users, which has a performance that depends on initial beamforming vectors. It is noteworthy that the problem in this paper differs from that in [11]. In [11], SC and beamforming are studied for broadcast channels to find the capacity (or rate-region). On the other hand, the novelty of this paper is to study the optimal beamforming vectors that minimize the total transmission power for given target rates in a broadcast system with multiresolution transmissions where the interference is mitigated at a user (receiver) using successive interference cancellation (SIC).

Nonorthogonal multiple access (NOMA) is studied in [12], [13] where SC is employed to increase the spectral efficiency and SIC is used at a receiver to remove interfering signals. In NOMA, a pair of users of strong and weak channel gains is usually considered to share a common access channel. In [14], beamforming is employed with SC for NOMA in a multiuser system, where a beam is shared by a pair of users. Since two users sharing a beam have strong and weak channel gains, it is possible to apply it to multiresolution multicast services, where two users have different qualities of signals. In this paper, we also apply the proposed multicast beamforming with SC to

Manuscript received May 9, 2014; revised July 30, 2014, September 28, 2014, and December 14, 2014; accepted January 15, 2015. Date of publication January 20, 2015; date of current version March 13, 2015. This work was supported by National Research Foundation of Korea (grant number: NRF-2014R1A1A2A16051720). The associate editor coordinating the review of this paper and approving it for publication was G. Abreu.

The author is with the School of Information and Communications, Gwangju Institute of Science and Technology (GIST), Gwangju 500-712, Korea (e-mail: jchoi0114@gist.ac.kr).

Digital Object Identifier 10.1109/TCOMM.2015.2394393

[1]Both broadcast and multicast modes are point-to-multipoint transmissions. Multicast mode requires user subscriptions, but broadcast mode does not [5], [6]. However, in this paper, we do not distinguish them. That is, they are interchangeable.

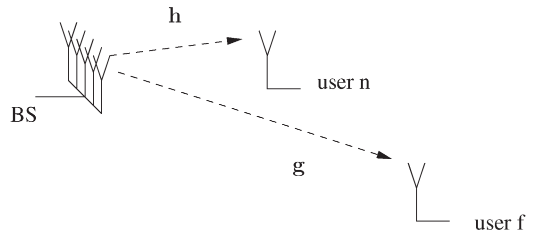


Fig. 1. An illustration of a system of multicast beamforming for multiresolution broadcast with two users.

NOMA for multiresolution multicast services with a two-stage beamforming method. The proposed two-stage beamforming method is considered for the same setup as in [14]. However, as each user in a group has a different dedicated beam, the proposed two-stage beamforming method can result in a better performance than that of the approach in [14] where no joint beam optimization problem is considered (only a beam is used to a group of users to suppress inter-group interference).

The rest of the paper is organized as follows. In Section II, we present a system model for multicast beamforming with SC and formulate an optimization problem to minimize the total transmission power for given target rates. To solve the optimization problem, an iterative algorithm is proposed in Section III. We apply the proposed multicast beamforming with SC to NOMA systems as a two-stage beamforming method in Section IV. Simulation results are presented in Section V and the paper is concluded with some remarks in Section VI.

Notation: Matrices and vectors are denoted by upper- and lower-case boldface letters, respectively. The superscripts T and H denote the transpose and Hermitian transpose, respectively. $\mathbb{E}[\cdot]$ denotes the statistical expectation. $\mathcal{CN}(\mathbf{a},\mathbf{R})$ represents the distribution of circularly-symmetric complex Gaussian (CSCG) random vectors with mean vector $\mathbf{a}$ and covariance matrix $\mathbf{R}$.

II. System Model and Problem Formulation

In this section, we present a system model for multicast beamforming with SC, formulate an optimization problem for optimal beamforming, and explain relations to existing multicast beamforming approaches.

A. Multicast Beamforming With SC

We consider downlink transmissions from a BS to two users per channel as illustrated in Fig. 1. The BS is equipped with an antenna array of L elements for downlink transmissions. One user is close to the BS, denoted by user n (i.e., near user), and the other user is not, denoted by user f (this user would be a far user or cell-edge user). Each user has a single receive antenna. Let s_{HP} and s_{LP} represent the signals of high-priority (HP) and low-priority (LP), respectively, as in multiresolution digital video broadcasting (DVB) services [4]. The HP and LP signals are transmitted using multicast beamforming with SC. It is assumed that user n can decode both the HP and LP data streams, while user f can only decode the HP data stream.

Let $\mathbf{w}$ and $\mathbf{v}$ denote the beamforming vectors for signals s_{LP} and s_{HP}, respectively. Then, using SC, the signal vector to be transmitted by the BS is given by

$$\mathbf{z} = \mathbf{w}s_{\mathrm{LP}} + \mathbf{v}s_{\mathrm{HP}}.$$

Let $\mathbf{h}$ and $\mathbf{g}$ denote the channel vectors of size $L \times 1$ from the BS to user n and user f (in general, it is expected that $\|\mathbf{h}\|^2 \gg \|\mathbf{g}\|^2$), respectively, which are assumed to be available at the BS for beamforming and power allocation. Throughout the paper, frequency-flat block fading channels are assumed. The channel vectors remain unchanged over the duration of a codeword transmission. Then, the received signals at the two users are

$$\begin{aligned} r_{\mathrm{n}} &= \mathbf{h}^{\mathrm{H}}\mathbf{z} + n_{\mathrm{n}} = \mathbf{h}^{\mathrm{H}}(\mathbf{w}s_{\mathrm{LP}} + \mathbf{v}s_{\mathrm{HP}}) + n_{\mathrm{n}}; \\ r_{\mathrm{f}} &= \mathbf{g}^{\mathrm{H}}\mathbf{z} + n_{\mathrm{f}} = \mathbf{g}^{\mathrm{H}}(\mathbf{w}s_{\mathrm{LP}} + \mathbf{v}s_{\mathrm{HP}}) + n_{\mathrm{f}}, \end{aligned} \tag{1}$$

where r_i and n_i represents the received signal and noise at user i, $i \in \{\mathrm{n},\mathrm{f}\}$, respectively. We assume that $\mathbb{E}[|s_j|^2] = 1$, $j \in \{\mathrm{LP},\mathrm{HP}\}$, and $n_i \sim \mathcal{CN}(0,\sigma_i^2)$, $i \in \{\mathrm{n},\mathrm{f}\}$, is independent.

B. Problem Formulation

Denote by ζ_{HP} and ζ_{LP} the signal-to-interference-plus-noise ratio (SINR) to decode s_{HP}, and the signal-to-noise ratio (SNR) to decode s_{LP} after removing s_{HP} using SIC (provided that decoding of s_{HP} is successful), respectively, at user n. Then, we have

$$\zeta_{\mathrm{HP}} = \frac{|\mathbf{h}^{\mathrm{H}}\mathbf{v}|^2}{|\mathbf{h}^{\mathrm{H}}\mathbf{w}|^2 + \sigma_{\mathrm{n}}^2}; \quad \zeta_{\mathrm{LP}} = \frac{|\mathbf{h}^{\mathrm{H}}\mathbf{w}|^2}{\sigma_{\mathrm{n}}^2}. \tag{2}$$

In general, it is expected that the signal power of s_{HP} (i.e., $\|\mathbf{v}\|^2$) is stronger than that of s_{LP} (i.e., $\|\mathbf{w}\|^2$) so that the decoding of s_{HP} can be successful and SIC becomes possible to decode s_{LP}.

At user f, s_{LP} is considered as an interferer and it may not be decodable as the signal strength is weak. In this case, the SINR to decode s_{HP} is given by

$$\gamma_{\mathrm{HP}} = \frac{|\mathbf{g}^{\mathrm{H}}\mathbf{v}|^2}{|\mathbf{g}^{\mathrm{H}}\mathbf{w}|^2 + \sigma_{\mathrm{f}}^2}. \tag{3}$$

Denote by R_{LP} and R_{HP} the rates of LP and HP, respectively. If capacity-achieving codes are employed for encoding s_{HP} and s_{LP}, the rate can be given by $C(x) = \log_2(1+x)$, where x denotes the SINR or SNR. For convenience, let

$$\Gamma_j = 2^{R_j} - 1, \quad j \in \{\mathrm{LP},\mathrm{HP}\}.$$

Since the HP data stream, s_{HP}, has to be decodable at user f and user n, R_{HP} is upper-bounded as

$$R_{\mathrm{HP}} \le \min\{C(\gamma_{\mathrm{HP}}), C(\zeta_{\mathrm{HP}})\}, \tag{4}$$

which results in the SINR constraint for s_{HP}, $\min\{\gamma_{\mathrm{HP}}, \zeta_{\mathrm{HP}}\} \ge \Gamma_{\mathrm{HP}}$. In particular, $\zeta_{\mathrm{HP}} \ge \Gamma_{\mathrm{HP}}$ is necessary at user n so that s_{HP} is decodable and canceled for SIC in decoding s_{LP}. Since $\gamma_{\mathrm{HP}} \ge \Gamma_{\mathrm{HP}}$ is required at user f, we have the constraint in (4). For successful decoding of s_{LP} after SIC at user n, we also need $\zeta_{\mathrm{LP}} \ge \Gamma_{\mathrm{LP}}$. Consequently, for given $\{R_j\}$ or $\{\Gamma_j\}$,

$j \in \{\mathrm{LP}, \mathrm{HP}\}$, an optimal multicast beamforming with SC problem to minimize the total power for multiresolution broadcast can be formulated as follows:

$$\min \|\mathbf{w}\|^2 + \|\mathbf{v}\|^2$$
$$\text{subject to } \begin{cases} \zeta_{\mathrm{LP}} & \geq \Gamma_{\mathrm{LP}}; \\ \min\{\gamma_{\mathrm{HP}}, \zeta_{\mathrm{HP}}\} & \geq \Gamma_{\mathrm{HP}}. \end{cases} \tag{5}$$

The resulting beamforming is referred to as the minimum power multicast beamforming with SC.

Note that the total rate to user n is $R_{\mathrm{LP}} + R_{\mathrm{HP}}$, while the rate to user f is R_{HP} (thus, the sum rate becomes $R_{\mathrm{LP}} + 2R_{\mathrm{HP}}$). If SC is not used, multiuser downlink beamforming [15] can be used, which leads to the following minimum power beamforming problem:

$$\min \|\mathbf{w}\|^2 + \|\mathbf{v}\|^2$$
$$\text{subject to } \begin{cases} \frac{|\mathbf{h}^{\mathrm{H}}\mathbf{w}|^2}{|\mathbf{h}^{\mathrm{H}}\mathbf{v}|^2 + \sigma_{\mathrm{n}}^2} & \geq \Gamma_{\mathrm{HP+LP}} = 2^{R_{\mathrm{LP}} + R_{\mathrm{HP}}} - 1; \\ \frac{|\mathbf{g}^{\mathrm{H}}\mathbf{v}|^2}{|\mathbf{g}^{\mathrm{H}}\mathbf{w}|^2 + \sigma_{\mathrm{f}}^2} & \geq \Gamma_{\mathrm{HP}} = 2^{R_{\mathrm{HP}}} - 1. \end{cases} \tag{6}$$

However, this approach will result in a higher total transmission power than the multicast beamforming with SC in (5).

In this paper, as shown in (5), we only consider the total power minimization problem with SNR/SINR constraints (as quality of service (QoS) constraints). There could be different problem formulations. For example, the sum rate maximization with power constraints can also be considered, which may be studied as a further research issue in the future.

C. Relation to Existing Multicast Beamforming Methods

There are closely related works to multicast beamforming with SC in discussed Section II-A. In [9], multicast beamforming is studied to transmit a single data stream to multiple receivers without SC. That is, a single signal is to be transmitted from the BS to multiple receivers using a beamforming vector that is designed to guarantee a certain SNR at all receivers. The optimal beamforming problem for this multicast beamforming is NP-hard problem and much more difficult than conventional multiuser beamforming problem where each user is to receive a dedicated signal from a BS [9]. The problem in Section II-B reduces to the multicast beamforming problem with two receivers if there is no LP transmission (i.e., $s_{\mathrm{LP}} = 0$ or $\mathbf{w} = 0$). In this case, we have

$$\min \|\mathbf{v}\|^2$$
$$\text{subject to } \min\left\{ \frac{|\mathbf{h}^{\mathrm{H}}\mathbf{v}|^2}{\sigma_{\mathrm{n}}^2}, \frac{|\mathbf{g}^{\mathrm{H}}\mathbf{v}|^2}{\sigma_{\mathrm{f}}^2} \right\} \geq \Gamma_{\mathrm{LP}}.$$

Clearly, this problem differs from the problem in (5) in Section II-B due to the presence of the SINR constraint (i.e., $\min\{\zeta_{\mathrm{HP}}, \gamma_{\mathrm{HP}}\} \geq \Gamma_{\mathrm{HP}}$) and LP transmissions. Furthermore, the tools derived in [9] cannot be applied to solve (5).

In [10], a generalization of multicast beamforming is studied with multiple multicast groups. Multiple users are divided into multiple multicast groups, and a dedicated beam is to be formed for each multicast group. The problem in (5) can be reformulated with two multicast groups if user n can belong to both multicast groups and user f belongs to only one multicast group. However, since the approach in [10] allows only disjoint multicast groups, the derived approach in this work cannot be applied to (5). Note that no beam optimization problems are considered in [14], although SC is employed with multicast beamforming.

Consequently, the problem in (5) in Section II-B differs from existing multicast beamforming and there are no existing approaches to solve this problem yet. Note that it would be desirable to support more than two users for a better spectral efficiency. This problem would be studied in the future as a further research topic.

III. Iterative Algorithm for Multicast Beamforming With SC

In this section, we focus on (5). Since a closed-form expression for the solution to (5) cannot be found, we consider an iterative algorithm based on a derived closed-form expression for the optimal transmission powers for given beamforming vectors, $\mathbf{w}$ and $\mathbf{v}$. Unfortunately, since the cost function is highly nonlinear and not convex, it is not guaranteed that the iterative algorithm for joint power and beamforming can converge to optimal solutions.

A. Optimal Power Allocation

Let

$$\mathbf{w} = \sqrt{P_w}\bar{\mathbf{w}} \quad \text{and} \quad \mathbf{v} = \sqrt{P_v}\bar{\mathbf{v}},$$

where $P_w = \|\mathbf{w}\|^2$ and $P_v = \|\mathbf{v}\|^2$ are the transmission powers for s_{LP} and s_{HP}, respectively, and $\bar{\mathbf{w}}$ and $\bar{\mathbf{v}}$ are normalized beamforming vectors (i.e., $\|\bar{\mathbf{w}}\| = \|\bar{\mathbf{v}}\| = 1$). For convenience, define

$$\tau_{1,1} = |\mathbf{h}^{\mathrm{H}}\bar{\mathbf{w}}|^2, \quad \tau_{1,2} = |\mathbf{h}^{\mathrm{H}}\bar{\mathbf{v}}|^2;$$
$$\tau_{2,1} = |\mathbf{g}^{\mathrm{H}}\bar{\mathbf{w}}|^2, \quad \tau_{2,2} = |\mathbf{g}^{\mathrm{H}}\bar{\mathbf{v}}|^2.$$

Then, for given $\tau_{p,q}$'s (or $\bar{\mathbf{w}}_1$ and $\bar{\mathbf{v}}_1$), the problem in (5) is to minimize the total power as follows:

$$\min_{P_w, P_v \geq 0} P_w + P_v$$
$$\text{subject to } \begin{cases} P_w\tau_{1,1} \geq \Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2; \\ P_v\tau_{1,2} \geq \Gamma_{\mathrm{HP}}\left(P_w\tau_{1,1} + \sigma_{\mathrm{n}}^2\right); \\ P_v\tau_{2,2} \geq \Gamma_{\mathrm{HP}}\left(P_w\tau_{2,1} + \sigma_{\mathrm{f}}^2\right). \end{cases} \tag{7}$$

The problem in (7) is a linear programming, which can be rewritten as given by

$$\min_{P_w, P_v \geq 0} P_w + P_v$$
$$\text{subject to } \begin{bmatrix} \tau_{1,1} & 0 \\ -\Gamma_{\mathrm{HP}}\tau_{1,1} & \tau_{1,2} \\ -\Gamma_{\mathrm{HP}}\tau_{2,1} & \tau_{2,2} \end{bmatrix} \begin{bmatrix} P_w \\ P_v \end{bmatrix} \geq \begin{bmatrix} \Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2 \\ \Gamma_{\mathrm{HP}}\sigma_{\mathrm{n}}^2 \\ \Gamma_{\mathrm{HP}}\sigma_{\mathrm{f}}^2 \end{bmatrix}. \tag{8}$$

Lemma 1: For given $\tau_{p,q}$ (or normalized beamforming vectors, $\bar{\mathbf{w}}$ and $\bar{\mathbf{v}}$), the solution to the problem in (8) is

$$P_w^* = \frac{\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2}{\tau_{1,1}}; \quad (9)$$

$$P_v^* = \max\left\{\frac{(1+\Gamma_{\mathrm{LP}})\Gamma_{\mathrm{HP}}\sigma_{\mathrm{n}}^2}{\tau_{1,2}}, \frac{\Gamma_{\mathrm{HP}}\left(\frac{\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2\tau_{2,1}}{\tau_{1,1}}+\sigma_{\mathrm{f}}^2\right)}{\tau_{2,2}}\right\}. \quad (10)$$

Proof: From the first constraint in (8), it follows that

$$P_w \geq \frac{\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2}{\tau_{1,1}}.$$

Since we have

$$P_v \geq \frac{1}{\tau_{1,2}}\left(P_w\Gamma_{\mathrm{HP}}\tau_{1,1}+\Gamma_{\mathrm{HP}}\sigma_{\mathrm{n}}^2\right)$$

from the second constraint, the minimum of $P_w + P_v$ can be achieved if $P_w = \frac{\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2}{\tau_{1,1}}$ and $P_v = \frac{(1+\Gamma_{\mathrm{LP}})\Gamma_{\mathrm{HP}}\sigma_{\mathrm{n}}^2}{\tau_{1,2}}$. With the third constraint, we can also find the solution, which results in (9) and (10). This completes the proof. ■

It can be shown that the total power, $P_w^* + P_v^*$, becomes a quadratic function of Γ if $\Gamma = \Gamma_{\mathrm{HP}} = \Gamma_{\mathrm{LP}}$ from (9) and (10). This indicates that the total transmission can grow quadratically with the target SINR, Γ, if Γ is sufficiently high.

B. Iterative Beamforming

From Lemma 1, we are able to find a closed-form expression for the minimum total transmission power for given normalized beamforming vectors. From this, the joint beamforming optimization problem can be formulated as

$$\{\bar{\mathbf{w}}_*, \bar{\mathbf{v}}_*\} = \underset{\|\bar{\mathbf{w}}\|=1, \|\bar{\mathbf{v}}\|=1}{\arg\min}\ D(\bar{\mathbf{w}}, \bar{\mathbf{v}}), \quad (11)$$

where

$$D(\bar{\mathbf{w}}, \bar{\mathbf{v}}) = P_w^*(\bar{\mathbf{w}}) + P_v^*(\bar{\mathbf{w}}, \bar{\mathbf{v}}). \quad (12)$$

Here, since P_w^* in (9) is a function of $\tau_{1,1}$, which means that it is a function of $\bar{\mathbf{w}}$, it is denoted by $P_w^*(\bar{\mathbf{w}})$. Similarly, since P_v^* in (10) is a function of $\bar{\mathbf{w}}$ and $\bar{\mathbf{v}}$, it is denoted by $P_v^*(\bar{\mathbf{w}}, \bar{\mathbf{v}})$. Unfortunately, it is not easy to find the solution to (11) since $D(\bar{\mathbf{w}}, \bar{\mathbf{v}})$ is highly nonlinear in terms of $\bar{\mathbf{w}}$ and $\bar{\mathbf{v}}$. Thus, we may consider the following iterative algorithm based on the nonlinear Gauss-Seidel (GS) algorithm [16]:

$$\bar{\mathbf{v}}^{(t)} = \underset{\|\bar{\mathbf{v}}\|=1}{\arg\min} D\left(\bar{\mathbf{w}}^{(t-1)}, \bar{\mathbf{v}}\right); \quad (13)$$

$$\bar{\mathbf{w}}^{(t)} = \underset{\|\bar{\mathbf{w}}\|=1}{\arg\min} D\left(\bar{\mathbf{w}}, \bar{\mathbf{v}}^{(t)}\right), \quad (14)$$

where $\bar{\mathbf{w}}^{(t)}$ and $\bar{\mathbf{v}}^{(t)}$ denote $\bar{\mathbf{w}}$ and $\bar{\mathbf{v}}$ at the tth iteration. The convergence of the nonlinear GS algorithm depends on the convexity of $D(\mathbf{w}, \mathbf{v})$ [16]. Unfortunately, $D(\mathbf{w}, \mathbf{v})$ is not convex as $P_v^*(\bar{\mathbf{w}}, \bar{\mathbf{v}})$ is not a convex function of $\bar{\mathbf{w}}$. Note that $P_v^*(\bar{\mathbf{w}}, \bar{\mathbf{v}})$ is a pointwise maximum of $\frac{(1+\Gamma_{\mathrm{LP}})\Gamma_{\mathrm{HP}}\sigma_{\mathrm{n}}^2}{\tau_{1,2}}$ and $\left(\Gamma_{\mathrm{HP}}\left(\frac{\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2\tau_{2,1}}{\tau_{1,1}}+\sigma_{\mathrm{f}}^2\right)\right)\left(\frac{1}{\tau_{2,2}}\right)$. If both the two functions are convex, $P_v^*(\bar{\mathbf{w}}, \bar{\mathbf{v}})$ is convex. However, the second one is not convex as $\frac{\tau_{2,1}}{\tau_{1,1}} = \frac{|\mathbf{g}^{\mathrm{H}}\bar{\mathbf{w}}|^2}{|\mathbf{h}^{\mathrm{H}}\bar{\mathbf{w}}|^2}$ is not a convex function of $\bar{\mathbf{w}}$. This implies that the iterative algorithm in (13) and (14) may not achieve a global[2] minimum (in simulations, we use randomly chosen initial beamforming vectors for $\mathbf{v}^{(0)}$ and $\mathbf{w}^{(0)}$).

We first consider (13), i.e., finding $\bar{\mathbf{v}}$ for given $\bar{\mathbf{w}}$. To this end, we need the following result.

Lemma 2: For given $\mathbf{a}$ and $\mathbf{b}$, we have[3]

$$\hat{\mathbf{u}} = \underset{\|\mathbf{u}\|^2=1}{\arg\max}\min\left\{|\mathbf{a}^{\mathrm{H}}\mathbf{u}|^2, |\mathbf{b}^{\mathrm{H}}\mathbf{u}|^2\right\} = \frac{\mathbf{a}+\hat{\kappa}\mathbf{b}}{\|\mathbf{a}+\hat{\kappa}\mathbf{b}\|}, \quad (15)$$

where $\hat{\kappa}$ is defined in the following proof.

Proof: The problem in (15) can be reformulated as

$$\hat{\mathbf{u}} = \underset{\|\mathbf{u}\|^2=1, |\mathbf{a}^{\mathrm{H}}\mathbf{u}|^2=|\mathbf{b}^{\mathrm{H}}\mathbf{u}|^2}{\arg\max} |\mathbf{a}^{\mathrm{H}}\mathbf{u}|^2. \quad (16)$$

Let us consider a linear combination of $\mathbf{a}$ and $\mathbf{b}$ as $\mathbf{x} = t_1\mathbf{a} + t_2\mathbf{b}$, where t_1 and t_2 are decided to satisfy $|\mathbf{a}^{\mathrm{H}}\mathbf{x}|^2 = |\mathbf{b}^{\mathrm{H}}\mathbf{x}|^2$. Then, t_1 and t_2 are related as follows:

$$t_1\left(\|\mathbf{a}\|^2 - e^{j\phi}\psi^*\right) = t_2\left(e^{j\phi}\|\mathbf{b}\|^2 - \psi\right),$$

where $\psi = \mathbf{a}^{\mathrm{H}}\mathbf{b}$ and $\phi \in [0, 2\pi)$ is a phase to be decided. Define

$$\kappa(\phi) = \frac{\|\mathbf{a}\|^2 - e^{j\phi}\psi^*}{e^{j\phi}\|\mathbf{b}\|^2 - \psi}.$$

Then, $\kappa(\phi) = \frac{t_2}{t_1}$. Letting $t_1 = 1$, the maximization of $|\mathbf{a}^{\mathrm{H}}\mathbf{x}|^2$ with respect to ϕ is given by

$$\hat{\phi} = \arg\max_{\phi} |\mathbf{a}^{\mathrm{H}}\mathbf{x}|^2 = \arg\max_{\phi} \left|\|\mathbf{a}\|^2 + \psi\kappa(\phi)\right|^2.$$

It follows that

$$\tilde{\mathbf{u}} = \arg\max_{\mathbf{u}} |\mathbf{a}^{\mathrm{H}}\mathbf{u}|^2$$
$$\text{subject to}\begin{cases}\|\mathbf{u}\|^2 = 1; \\ |\mathbf{a}^{\mathrm{H}}\mathbf{u}|^2 = |\mathbf{b}^{\mathrm{H}}\mathbf{u}|^2; \\ \mathbf{u} \in \mathrm{Span}(\mathbf{a}, \mathbf{b}).\end{cases} \quad (17)$$

Then, we can see that $\tilde{\mathbf{u}}$ is the solution to the problem in (15) with the additional constraint that $\mathbf{u} \in \mathrm{Span}(\mathbf{a}, \mathbf{b})$. The solution is given by

$$\tilde{\mathbf{u}} = \frac{\mathbf{a}+\kappa(\hat{\phi})\mathbf{b}}{\|\mathbf{a}+\kappa(\hat{\phi})\mathbf{b}\|}, \quad (18)$$

Let $\hat{\kappa} = \kappa(\hat{\phi})$.

[2]Although it is not guaranteed that the GS algorithm can achieve a global minimum, we have found that it works well from simulations. In particular, with different random initial vectors of $\bar{\mathbf{w}}$ and $\bar{\mathbf{v}}$, the GS algorithm provides the same solution in most cases (but not always).

[3]Note that $e^{j\theta}\hat{\mathbf{u}}$, $\theta \in [0, 2\pi)$, is also a solution. For convenience, we assume that the phase of the first element of $\mathbf{u}$ is the same as that of $\mathbf{a}$.

We now show that $\tilde{\mathbf{u}}$ is the optimal solution even if the constraint $\mathbf{u} \in \text{Span}(\mathbf{a}, \mathbf{b})$ is removed, i.e., $\tilde{\mathbf{u}}$ is the solution to the problem in (16). With a vector $\mathbf{e} \neq 0$ that is orthogonal to $\text{Span}(\mathbf{a}, \mathbf{b})$, consider $\mathbf{u_e} = \tilde{\mathbf{u}} + \mathbf{e}$. This vector satisfies the constraint $|\mathbf{a}^{\mathrm{H}}\mathbf{u_e}|^2 = |\mathbf{b}^{\mathrm{H}}\mathbf{u_e}|^2$. Let $\bar{\mathbf{u}}_\mathbf{e} = \frac{\mathbf{u_e}}{||\mathbf{u_e}||}$. Then, it can be shown that

$$|\mathbf{a}^{\mathrm{H}}\bar{\mathbf{u}}_\mathbf{e}|^2 = \frac{|\mathbf{a}^{\mathrm{H}}\mathbf{u_e}|^2}{||\mathbf{u_e}||^2} = \frac{|\mathbf{a}^{\mathrm{H}}\tilde{\mathbf{u}}|^2}{||\tilde{\mathbf{u}}||^2 + ||\mathbf{e}||^2} \leq |\mathbf{a}^{\mathrm{H}}\tilde{\mathbf{u}}|^2,$$

which implies that $\tilde{\mathbf{u}}$ is the solution to the problem in (15). ■

According to Lemma 2, for given $\bar{\mathbf{w}}^{(t-1)}$, we can find $\bar{\mathbf{v}}^{(t)}$ in (13). From (10) and (12), we have

$$\bar{\mathbf{v}}^{(t)} = \arg\min_{||\bar{\mathbf{v}}||^2=1} D\left(\bar{\mathbf{w}}^{(t-1)}, \bar{\mathbf{v}}\right) = \arg\min_{||\bar{\mathbf{v}}||^2=1} P_v^*\left(\bar{\mathbf{w}}^{(t-1)}, \bar{\mathbf{v}}\right)$$
$$= \arg\max_{||\bar{\mathbf{v}}||^2=1} \min\left\{\frac{|\mathbf{h}^{\mathrm{H}}\bar{\mathbf{v}}|^2}{(1+\Gamma_{\mathrm{LP}})\Gamma_{\mathrm{HP}}\sigma_{\mathrm{n}}^2}, \frac{|\mathbf{g}^{\mathrm{H}}\bar{\mathbf{v}}|^2}{\Gamma_{\mathrm{HP}}\left(\frac{\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2\tau_{2,1}}{\tau_{1,1}} + \sigma_{\mathrm{f}}^2\right)}\right\}, \quad (19)$$

where $\tau_{1,1} = |\mathbf{h}^{\mathrm{H}}\bar{\mathbf{w}}^{(t-1)}|^2$ and $\tau_{2,1} = |\mathbf{g}^{\mathrm{H}}\bar{\mathbf{w}}^{(t-1)}|^2$. Thus, letting

$$\mathbf{a} = \frac{1}{\sqrt{(1+\Gamma_{\mathrm{LP}})\Gamma_{\mathrm{HP}}\sigma_{\mathrm{n}}^2}}\mathbf{h}; \quad \mathbf{b} = \frac{1}{\sqrt{\Gamma_{\mathrm{HP}}\left(\frac{\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2\tau_{2,1}}{\tau_{1,1}} + \sigma_{\mathrm{f}}^2\right)}}\mathbf{g},$$

$\bar{\mathbf{v}}^{(t)}$ in (13) can be found from Lemma 2.

We now move on to (14). For given $\tau_{2,2}$ or $\bar{\mathbf{v}}$, we can consider the optimal beamforming for $\bar{\mathbf{w}}$ to minimize $D(\bar{\mathbf{w}}, \bar{\mathbf{v}})$ as follows:

$$\min_{||\bar{\mathbf{w}}||^2=1} J(\bar{\mathbf{w}}), \quad (20)$$

where

$$J(\bar{\mathbf{w}}) = \frac{\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2}{\tau_{1,1}} + \max\left\{T_4, K_4\frac{\tau_{2,1}}{\tau_{1,1}}\right\}.$$

Here, we have

$$T_4 = \frac{(1+\Gamma_{\mathrm{LP}})\Gamma_{\mathrm{HP}}\sigma_{\mathrm{n}}^2}{\tau_{1,2}} - \frac{\Gamma_{\mathrm{HP}}\sigma_{\mathrm{f}}^2}{\tau_{2,2}}; \quad K_4 = \frac{\Gamma_{\mathrm{HP}}\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2}{\tau_{2,2}}.$$

Note that $D(\bar{\mathbf{w}}, \bar{\mathbf{v}}) = J(\bar{\mathbf{w}}) + \frac{\Gamma_{\mathrm{HP}}\sigma_{\mathrm{f}}^2}{\tau_{2,2}}$ for a fixed $\bar{\mathbf{v}}$.

Define the following disjoint sets:

$$\mathcal{W}_I = \left\{\mathbf{w} \,\middle|\, \frac{|\mathbf{g}^{\mathrm{H}}\mathbf{w}|^2}{|\mathbf{h}^{\mathrm{H}}\mathbf{w}|^2} > \frac{T_4}{K_4}\right\};$$

$$\mathcal{W}_{II} = \left\{\mathbf{w} \,\middle|\, \frac{|\mathbf{g}^{\mathrm{H}}\mathbf{w}|^2}{|\mathbf{h}^{\mathrm{H}}\mathbf{w}|^2} < \frac{T_4}{K_4}\right\};$$

$$\mathcal{W}_{III} = \left\{\mathbf{w} \,\middle|\, \frac{|\mathbf{g}^{\mathrm{H}}\mathbf{w}|^2}{|\mathbf{h}^{\mathrm{H}}\mathbf{w}|^2} = \frac{T_4}{K_4}\right\}.$$

Note that $\mathcal{W}_I \cup \mathcal{W}_{II} \cup \mathcal{W}_{III} = \mathbb{C}^L$. Then, we can decompose the optimization problem in (20) into three problems as follows:

$$\min_{||\bar{\mathbf{w}}||^2=1} J(\bar{\mathbf{w}}) = \min\left\{J_I(\bar{\mathbf{w}}), J_{II}(\bar{\mathbf{w}}), J_{III}(\bar{\mathbf{w}})\right\}, \quad (21)$$

where

$$J_p(\bar{\mathbf{w}}) = \min_{||\bar{\mathbf{w}}||^2=1, \bar{\mathbf{w}}\in\mathcal{W}_p} J(\bar{\mathbf{w}}), \ p \in \{I, II, III\}.$$

The first problem is equivalent to the following problem:

$$\min_{\bar{\mathbf{w}}\in\mathcal{W}_I, ||\bar{\mathbf{w}}||^2=1} \frac{\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2}{|\mathbf{h}^{\mathrm{H}}\bar{\mathbf{w}}|^2} + K_4\frac{|\mathbf{g}^{\mathrm{H}}\bar{\mathbf{w}}|^2}{|\mathbf{h}^{\mathrm{H}}\bar{\mathbf{w}}|^2}. \quad (22)$$

This problem can be solved using the maximum SNR (MSNR) beamforming approach [17] if the constraint that $\bar{\mathbf{w}} \in \mathcal{W}_I$ is not taken into account, which is given by

$$\bar{\mathbf{w}}_I = \arg\max_{||\bar{\mathbf{w}}||^2=1} \frac{\bar{\mathbf{w}}^{\mathrm{H}}\mathbf{h}\mathbf{h}^{\mathrm{H}}\bar{\mathbf{w}}}{\bar{\mathbf{w}}^{\mathrm{H}}\left(\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2\mathbf{I} + K_4\mathbf{g}\mathbf{g}^{\mathrm{H}}\right)\bar{\mathbf{w}}}$$
$$= \kappa_4\left(\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2\mathbf{I} + K_4\mathbf{g}\mathbf{g}^{\mathrm{H}}\right)^{-1}\mathbf{h}, \quad (23)$$

where κ_4 is a normalization factor. If $\bar{\mathbf{w}}_I \in \mathcal{W}_I$, this is the solution to the first problem. Otherwise, there is no solution (i.e., $J_I(\bar{\mathbf{w}}) = \infty$).

The second problem is equivalent to the following problem:

$$\min_{\bar{\mathbf{w}}\in\mathcal{W}_{II}, ||\bar{\mathbf{w}}||^2=1} \frac{\Gamma_{\mathrm{LP}}\sigma_{\mathrm{n}}^2}{|\mathbf{h}^{\mathrm{H}}\bar{\mathbf{w}}|^2}. \quad (24)$$

Without the constraint that $\bar{\mathbf{w}} \in \mathcal{W}_{II}$, the solution becomes $\bar{\mathbf{w}}_{II} = \frac{\mathbf{h}}{||\mathbf{h}||}$. If this solution lies in $\mathcal{W}_{II}$, this is the solution to the second problem. Otherwise, there is no solution.

The third problem is equivalent to the following problem:

$$\bar{\mathbf{w}}_{III} = \arg\max |\mathbf{h}^{\mathrm{H}}\bar{\mathbf{w}}|^2$$
$$\text{subject to } \begin{cases} |\mathbf{h}^{\mathrm{H}}\bar{\mathbf{w}}|^2 = \frac{K_4}{T_4}|\mathbf{g}^{\mathrm{H}}\bar{\mathbf{w}}|^2; \\ ||\bar{\mathbf{w}}||^2 = 1. \end{cases} \quad (25)$$

Letting $\mathbf{a} = \mathbf{h}$ and $\mathbf{b} = \sqrt{\frac{K_4}{T_4}}\mathbf{g}$, the third problem can be found using Lemma 2. Since this solution always exists and is finite, the existence of the solution to the problem in (21) is guaranteed. Finally, using (21), we can perform (14).

IV. Application to NOMA Systems

Most NOMA systems in [13], [14], [18], [19] are to support two users per channel using SC. For a pair of users of strong and weak channel gains, the user of strong channel gain can decode his/her desired signals (i.e., signals of LP) as well as the signals to the user of weak channel gain (i.e., signals of HP). Thus, NOMA is well-suited to multiresolution broadcast services, and it is possible to employ an existing NOMA approach with single-beam per pair of users, e.g., [14] for multiresolution broadcast services. In this section, we apply the proposed minimum power multicast beamforming with SC to NOMA systems for multiresolution broadcast and compare it with the approach in [14].

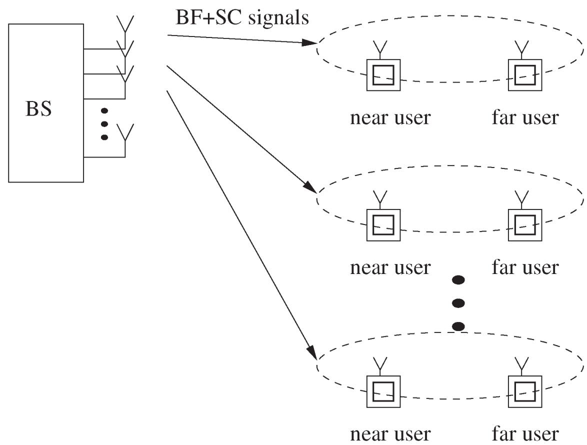


Fig. 2. An illustration of an NOMA system to support $2K$ users with beamforming and SC.

A. Two-Stage Beamforming for NOMA

For convenience, assume that there are K near users and K far users. There are K groups (clusters) of users and each group consists of one near user and one far user. A BS is to transmit $2K$ signals using SC and beamforming. An illustration of the NOMA system is shown in Fig. 2. We consider two-stage beamforming. To avoid the interference to the other groups, zero-forcing (ZF) beamforming is employed first. Within a group, the minimum power multicast beamforming with SC in the previous sections is considered. Certainly, the two-stage beamforming approach is suboptimal, but tractable, as we do not consider any *overall* optimization problem to find beamforming vectors with SC. For optimal beamforming with SC for NOMA systems of multiple users, we need further research by possibly generalizing the approaches in [9], [10] with SC.

Denote by $\mathbf{h}_k$ and $\mathbf{g}_j$ the channel vectors to the kth near and the jth far users, respectively. Furthermore, denote by $s_{\mathrm{n},k}$ and $s_{\mathrm{f},j}$ the signals to the kth near and the jth far users, respectively. Suppose that in the kth group, the channel vectors are $\{\mathbf{h}_k, \mathbf{g}_{j(k)}\}$, where $j(k)$ denotes the index for the far user who is paired with the kth near user. Note that $\{j(1), \ldots, j(K)\} = \{1, \ldots, K\}$. Using a certain user allocation or pairing algorithm [14], [20], we can find all the paired users, $\{(1, j(1)), \ldots, (K, j(K))\}$. In NOMA systems, different signals (not broadcast signals) are to be transmitted to near and far users. Thus, $R_{\mathrm{LP}} = R_{\mathrm{n}}$ (not $R_{\mathrm{LP}} + R_{\mathrm{HP}}$) becomes the data rate to a near user and $R_{\mathrm{HP}} = R_{\mathrm{f}}$ becomes the data rate to a far user. Furthermore, all the near (far) users have the same data rate, R_{n} (R_{f}, resp.) for convenience.

Let $\mathbf{H}_k = [\mathbf{h}_k \ \mathbf{g}_{j(k)}]$. In addition, define

$$\mathbf{H}_{-k} = [\mathbf{H}_1 \ \ldots \ \mathbf{H}_{k-1} \ \mathbf{H}_{k+1} \ \ldots \ \mathbf{H}_K].$$

Thus, the size of $\mathbf{H}_{-k}$ is $L \times 2(K-1)$. We assume that L is sufficiently large in NOMA systems such that $L \geq 2K$. To avoid the interference to the other groups, we consider an $L \times M_k$ matrix, $\mathbf{B}_k$, where $M_k = L - \mathrm{rank}(\mathbf{H}_{-k})$, such that

$$\mathbf{B}_k \perp \mathbf{H}_{-k}. \tag{26}$$

In addition, we assume $\mathbf{B}_k^{\mathrm{H}} \mathbf{B}_k = \mathbf{I}$ for normalization purposes. It is easy to see that

$$M_k \geq M \overset{\Delta}{=} \min_k \{M_1, \ldots, M_K\} \geq L - 2(K-1),$$

where M is referred to as the effective number of antennas for multicast beamforming with SC. Throughout the paper, we assume that $M \geq 2$. Using the singular value decomposition (SVD) of $\mathbf{H}_{-k}$, $\mathbf{B}_k$ can be found. Let

$$\mathbf{H}_{-k} = \bar{\mathbf{U}}_k \bar{\Sigma}_k \bar{\mathbf{V}}_k^{\mathrm{H}},$$

where $\bar{\mathbf{U}}_k$ is an $L \times L$ unitary matrix, $\bar{\Sigma}_k$ is an $L \times 2(K-1)$ diagonal matrix, and $\bar{\mathbf{V}}_k$ is a $2(K-1) \times 2(K-1)$ unitary matrix. Then, the column vectors of $\mathbf{B}_k$ are the left-singular vectors (the column vectors of $\bar{\mathbf{U}}_k$) that correspond to zero singular values.

Using the multicast beamforming with SC, the signal to be transmitted by the BS is given by

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{B}_k \left(\mathbf{w}_k s_{\mathrm{n},k} + \mathbf{v}_{j(k)} s_{\mathrm{f},j(k)} \right). \tag{27}$$

Denote by $r_{\mathrm{n},k}$ and $r_{\mathrm{f},j(k)}$ the received signals at the near and far users in the kth cluster, respectively. Then, we have

$$\begin{aligned} r_{\mathrm{n},k} &= \mathbf{h}_k^{\mathrm{H}} \mathbf{x} + n_{\mathrm{n},k} \\ &= \mathbf{h}_k^{\mathrm{H}} \mathbf{B}_k \left(\mathbf{w}_k s_{\mathrm{n},k} + \mathbf{v}_{j(k)} s_{\mathrm{f},j(k)} \right) + n_{\mathrm{n},k}; \\ r_{\mathrm{f},j(k)} &= \mathbf{g}_{j(k)}^{\mathrm{H}} \mathbf{x} + n_{\mathrm{f},j(k)} \\ &= \mathbf{g}_{j(k)}^{\mathrm{H}} \mathbf{B}_k \left(\mathbf{w}_k s_{\mathrm{n},k} + \mathbf{v}_{j(k)} s_{\mathrm{f},j(k)} \right) + n_{\mathrm{f},j(k)}, \end{aligned} \tag{28}$$

where $n_{\mathrm{n},k} \sim \mathcal{CN}(0, \sigma_{\mathrm{n},k}^2)$ and $n_{\mathrm{f},j(k)} \sim \mathcal{CN}(0, \sigma_{\mathrm{f},j(k)}^2)$ are the background noise terms at the near and far users, respectively.

Then, the minimum power multicast beamforming problem for the kth group can be considered as in (5) with the composite channel vectors that are given by

$$\{\mathbf{h}, \mathbf{g}\} = \left\{ \mathbf{B}_k^{\mathrm{H}} \mathbf{h}_k, \mathbf{B}_k^{\mathrm{H}} \mathbf{g}_{j(k)} \right\}$$

and the target rates $R_{\mathrm{n}} (= R_{\mathrm{LP}})$ and $R_{\mathrm{f}} (= R_{\mathrm{HP}})$ for the near and far users, respectively. Note that if the users in a cluster are to receive broadcast signals (or receive high resolution signals as in multicast beamforming with SC), the near user can have a data rate of $R_{\mathrm{n}} + R_{\mathrm{f}} = R_{\mathrm{LP}} + R_{\mathrm{HP}}$.

In this proposed two-stage beamforming with SC for NOMA systems, the user pairing is important to reduce the total transmission power. Let $P_w(k)$ and $P_v(j(k))$ denote the transmission powers to the near and far users in group k, respectively, that are found from (9) and (10) with optimal multicast beamforming vectors. Then, an optimal user pairing problem can be given by

$$\{j^*(1), \ldots, j^*(K)\} = \underset{\{j(1), \ldots, j(K)\}}{\arg\min} \ P_w(k) + P_v(j(k)). \tag{29}$$

If an exhaustive search is used to solve (29), the complexity is proportional to $K!$. To avoid a high computational complexity, a greedy algorithm can be used.

To derive a low-complexity user pairing method, we can also replace the cost function $P_w(k) + P_v(j(k))$ with another cost function that can be found with a much lower complexity, since the complexity to find multicast beamforming vectors can be high for each pair of $\mathbf{h}_k$ and $\mathbf{g}_{j(k)}$. To this end, we can consider the correlation of paired channel vectors, $\mathbf{h}_k$ and $\mathbf{g}_{j(k)}$. As an extreme case, suppose that the channel vectors in each group, $\mathbf{h}_k$ and $\mathbf{g}_{j(k)}$, are perfectly correlated, i.e., $\mathbf{h}_k \propto \mathbf{g}_{j(k)}$. Then, since the rank of $\mathbf{H}_{-k}$ is $K-1$, the number of column vectors of $\mathbf{B}_k$ becomes $M_k = L-(K-1)$ that is greater than $K-2(K-1)$. Thus, the effective number of antennas for multicast beamforming with SC can be higher, which can provide a better performance of multicast beamforming with SC. Based on this observation, we consider the correlation of $\mathbf{h}_k$ and $\mathbf{g}_{j(k)}$ as a cost function. Consequently, the proposed greedy algorithm to find the index for the far user in the kth group, $j(k)$, can be given by

$$j^*(k) = \underset{j \in \mathcal{K}_{\mathrm{f}}(k-1)}{\arg\max} \frac{|\mathbf{h}_k^{\mathrm{H}} \mathbf{g}_j|^2}{\|\mathbf{h}_k\|^2 \|\mathbf{g}_j\|^2}, \quad k = 1, 2, \rightarrow K, \tag{30}$$

where $\mathcal{K}_{\mathrm{f}}(k) = \mathcal{K}_{\mathrm{f}}(k-1) \setminus j^*(k)$ and $\mathcal{K}_{\mathrm{f}}(0) = \{1, \ldots, K\}$. The complexity of this greedy algorithm is $O(\sum_{k=1}^{K} k) = O(K^2)$.

B. An Existing Approach With Beamforming for NOMA

For comparison purposes, we can briefly present the beamforming approach for NOMA systems in [14]. In this approach, a single beam is shared by two users in a group, which is denoted by $\mathbf{c}_k$ for the kth user group. This is the ZF beamforming vector for the near user. That is,

$$[\mathbf{c}_1 \ldots \mathbf{c}_K] = \left([\mathbf{h}_1 \ldots \mathbf{h}_K]^{\mathrm{H}}\right)^{\dagger},$$

where † denotes the pseudo-inverse. Thus, for convenience, this approach is referred to as the single beam per group (SBPG) approach in this paper. In SBPG, SC is also considered, and the transmitted signal is given by

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{c}_k \left(\sqrt{\alpha_k P_k} s_{\mathrm{n},k} + \sqrt{(1-\alpha_k) P_k} s_{\mathrm{f},k}\right),$$

where α_k is the parameter that decide the power ratio between the near and far users within a user group and P_k is the nominal transmission power. Note that the actual transmission power to user group k is $\|\mathbf{c}_k\|^2 P_k$. Due to the ZF beamforming, the near users do not suffer from the interferences from the signals to the other user groups. However, the far users suffer from the interference. The resulting SNR/SINR can be given by

$$\gamma_{\mathrm{n},k} = \frac{\left|\mathbf{h}_k^{\mathrm{H}} \mathbf{c}_k\right|^2 \alpha_k P_k}{\sigma_{\mathrm{n},k}^2};$$

$$\gamma_{\mathrm{f},k} = \frac{\left|\mathbf{g}_{j(k)}^{\mathrm{H}} \mathbf{c}_k\right|^2 (1-\alpha_k) P_k}{\left|\mathbf{g}_{j(k)}^{\mathrm{H}} \mathbf{c}_k\right|^2 \alpha_k P_k + \sum_{q \neq k} \left|\mathbf{g}_{j(k)}^{\mathrm{H}} \mathbf{c}_q\right|^2 P_q + \sigma_{\mathrm{f},j(k)}^2}, \tag{31}$$

where $\gamma_{\mathrm{n},k}$ and $\gamma_{\mathrm{f},k}$ are the SNR for the near user and SINR for the far user in group k, respectively. Note that it is assumed that SIC is used to cancel the signal to the far user at the near user and this SIC is always successful in [14]. Based on this assumption, $\gamma_{\mathrm{n},k}$ in (31) does not have interference terms. This assumption is valid if

$$\bar{\gamma}_{\mathrm{f},k} = \frac{\left|\mathbf{h}_k^{\mathrm{H}} \mathbf{c}_k\right|^2 (1-\alpha_k) P_k}{\left|\mathbf{h}_k^{\mathrm{H}} \mathbf{c}_k\right|^2 \alpha_k P_k + \sigma_{\mathrm{n},j(k)}^2} \geq 2^{R_{\mathrm{f}}} - 1. \tag{32}$$

In general it is expected that $\bar{\gamma}_{\mathrm{f},k} > \gamma_{\mathrm{f},k}$, because $\bar{\gamma}_{\mathrm{f},k}$ does not have the inter-group interference. Thus, if $R_{\mathrm{f}} < \log_2(1+\gamma_{\mathrm{f},k})$ and $\mathbf{h}_k$ and $\mathbf{g}_{j(k)}$ are highly correlated, the inequality in (32) could be satisfied and SIC can be successfully carried out at the near user in SBPG.

In SBPG, if a near user needs a data rate comparable to that of the far user in the same user group, α_k has to be sufficiently close to 0 so that the SINR $\gamma_{\mathrm{f},k}$ is close to the SNR $\gamma_{\mathrm{n},k}$. This can result in a low total rate or a low spectral efficiency for this user group. We will discuss this problem and compare the performances of the proposed two-stage beamforming approach and the SBPG approach in [14] in Section V with simulation results.

V. Simulation Results

A. Multicast Beamforming for Multiresolution Broadcast

For simulations, we assume that $\mathbf{h}$ and $\mathbf{g}$ are independent and $\mathbf{h} \sim \mathcal{CN}(0, \frac{1}{L}\mathbf{I})$ and $\mathbf{g} \sim \mathcal{CN}(0, \frac{1}{d^\eta L}\mathbf{I})$, where d is the relative distance between the BS and user f (provided that the distance between the BS and user n is normalized) and η is the path loss exponent. Throughout simulations, η is set to 2, $d \geq 1$, and $R_{\mathrm{HP}} = R_{\mathrm{LP}}$ for convenience. Furthermore, we assume that $\sigma_{\mathrm{f}}^2 = \sigma_{\mathrm{n}}^2 = 1$.

Fig. 3(a) shows the relationship between the sum rate $(2R_{\mathrm{HP}} + R_{\mathrm{LP}})$ and the total transmission power with $L=5$ and $d=2$, which confirms that the multicast beamforming with SC has a higher spectral efficiency than the conventional multiuser beamforming in (6). Note that the sum rate is $3R_{\mathrm{HP}}$ as it is assumed that $R_{\mathrm{HP}} = R_{\mathrm{LP}}$. In Fig. 3(b), we show the average number of iterations of the iterative algorithm to find beams in Section III when initial vectors are randomly chosen. It is shown that the number of iterations of the iterative algorithm in (13) and (14) is not large (at most 3 iterations) and the number of iterations slowly increases with the transmission power. Note that the number of iterations.

The impact of the relative distance, d, on the total transmission power is shown in Fig. 4 when $L=5$ and $R_{\mathrm{HP}} = R_{\mathrm{LP}} = 3$. It is shown that the proposed multicast beamforming with SC can provide about 2 dB transmission power gain compared with the conventional multiuser beamforming in (6) when $d=2$, while the power gain gap decreases as d decreases. This gain results from SC. Since the distance between user n and BS is fixed and the SNR is independent of the signal power of HP transmissions (see (2)), the signal power of LP transmissions, which is decodable only by user n, may not change (provided that signals of HP are decodable and can be removed) in the proposed multicast beamforming with SC. Thus, the increase of the total transmission power when d increases mainly results from the increase of signal power of HP transmissions for user f. On the other hand, in the conventional multiuser beamforming, the signal power to user n also increases when that to user f

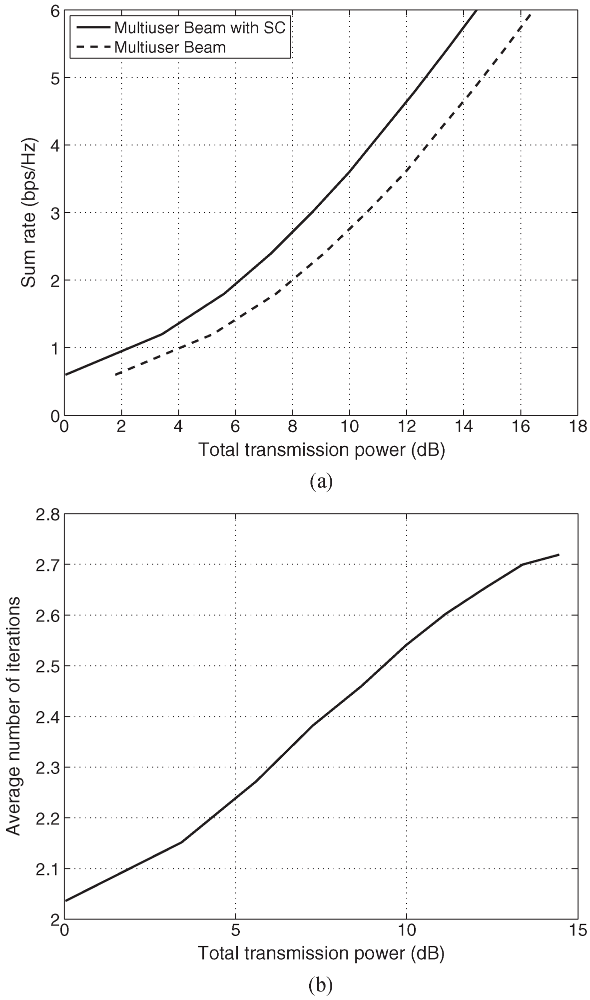


Fig. 3. Sum rate and average number of iterations ($L = 5$ and $d = 2$): (a) Sum rate versus total transmission power; (b) Average number of iterations for the iterative algorithm to find beams in Section III.

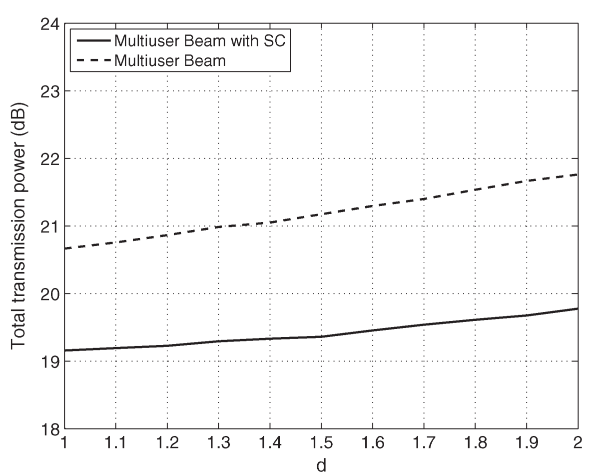


Fig. 4. Sum rate versus the relative distance, d, with $L = 5$ and $R_{\mathrm{HP}} = R_{\mathrm{LP}} = 3$.

increases due to the interference (see Eq. (6)). From this, we can see that the power gap increases with d.

Fig. 5 shows the impact of L on the total transmission power when $d = 2$ and $R_{\mathrm{HP}} = R_{\mathrm{LP}} = 3$. As L increases, better beamforming results are expected. Thus, the total transmission

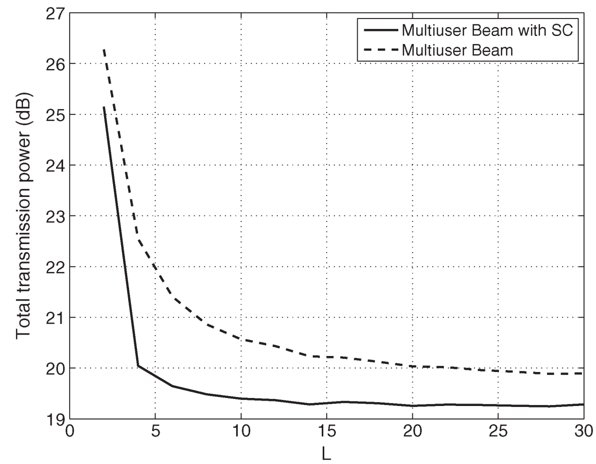


Fig. 5. Sum rate versus L with $d = 2$ and $R_{\mathrm{HP}} = R_{\mathrm{LP}} = 3$.

powers decrease with L in both the multicast beamforming with SC and the conventional multiuser beamforming, while the multicast beamforming with SC has a lower total transmission power than that without SC. It is also shown that the transmission power decreases rapidly with L when $L \leq 10$, while the increase of L does not result in a significant decrease of the transmission power when $L > 10$. With a sufficiently large L, the impact of the interference on the performance can be insignificant. As a result, the performance mainly depends on the noise variance and average channel gains in this case. That is, the performance may not be improved further by increasing L once L is sufficiently large as shown in Fig. 5.

B. Two-Stage Beamforming for NOMA

In this subsection, we present simulation results of the proposed two-stage beamforming method for NOMA systems. The channel vectors are independent and generated as $\mathbf{h}_k \sim \mathcal{CN}(0, \frac{1}{L}\mathbf{I})$ and $\mathbf{g}_k \sim \mathcal{CN}(0, \frac{1}{d^{\eta}L}\mathbf{I})$. We also assume that $\sigma^2_{\mathrm{f},k} = \sigma^2_{\mathrm{n},k} = 1$ for all k, $d = 2$, and $\eta = 2$ in this subsection.

To compare with the SBPG approach in [14], simulations are carried out with different target rates (for the proposed approach in this paper) or transmission powers (for the SBPG approach) with $K = 5$ pairs of users when $L = 20$. The SBPG approach uses the ZF beamforming and there are $K = 5$ beams (they are orthogonal between 5 near users). The results are shown in Fig. 6. For the SBPG approach in [14], we assume $P_k = P$ (i.e., equal total transmission power to all user groups) and $\alpha_k = \alpha \in (0, 1)$ for all k with two different values of α, 0.5 and 0.01 for power allocation. As α decreases (or approaches 0), we can see that the more transmission power is allocated to far users from (31). From Fig. 6, it seems that the rates to near and far users become comparable with $\alpha = 0.01$, while the rate to near users becomes much higher than that to far users if $\alpha = 0.5$. It is noteworthy that an exhaustive search is used for the SBPG approach to find the best pairs that can achieve the maximum sum rate for a given total transmission power per group.

For the proposed two-stage beamforming approach, we assume $R_{\mathrm{f}} = R_{\mathrm{n}}$. Based on simulation results, the proposed approach can perform much better than the SBPG approach in

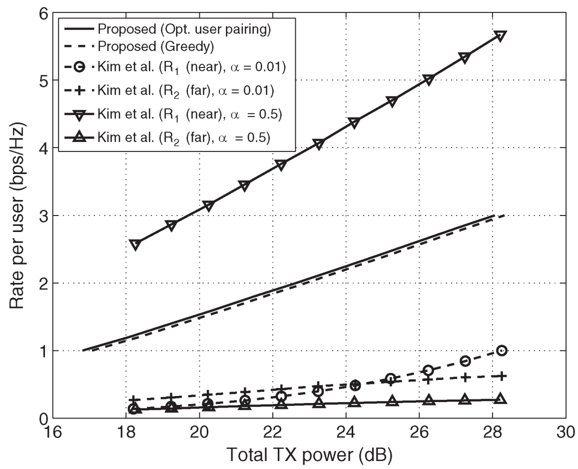


Fig. 6. Rate per user and total transmission powers with $L = 20$ and $K = 5$.

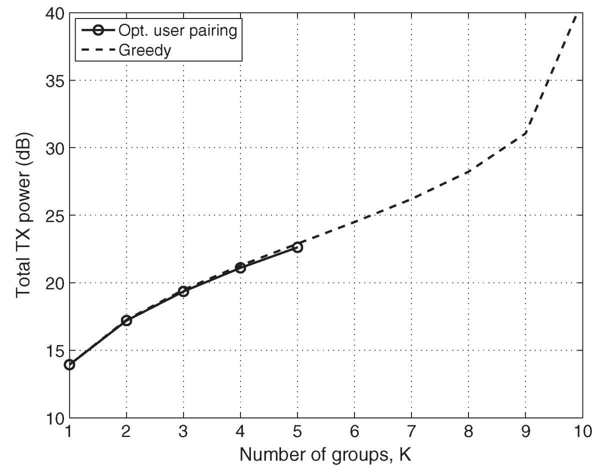


Fig. 8. Total transmission power versus K with $L = 20$ and $R_{\mathrm{f}} = R_{\mathrm{n}} = 2$.

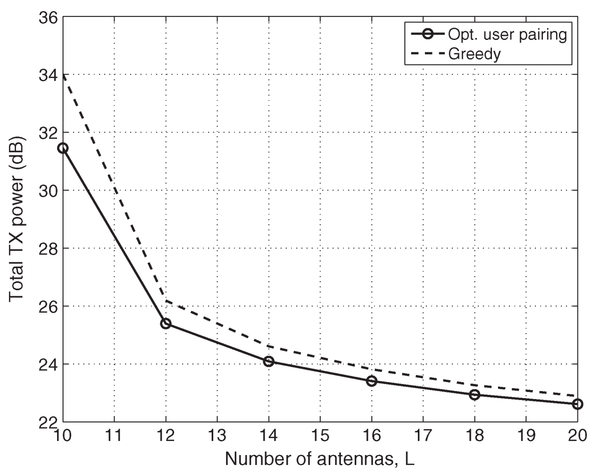


Fig. 7. Total transmission power versus L with $K = 5$ and $R_{\mathrm{f}} = R_{\mathrm{n}} = 2$.

[14] if the rates to near and far users are to be comparable. For example, with $\alpha = 0.01$, the rates to near and far users are identical (where the rate per user is about 0.5 bps/Hz) when the total transmission power is about 24.2 dB. At the same total transmission power, the proposed approach can provide a rate per user of about 2.2 bps/Hz, which is about 4 times higher than that of the SBPG approach. Since the proposed approach has two separate beams to near and far users with joint power allocation, it can outperform the SBPG approach when the rates to near and far users are to be comparable.

To see the impact of L on the performance of the proposed two-stage beamforming approach, simulations are carried out with various values of L with $K = 5$ and $R_{\mathrm{f}} = R_{\mathrm{n}} = 2$ and the results are shown in Fig. 7. As L increases, the total transmission power decreases, which indicates that narrow beams are also desirable to save the transmission power in NOMA systems.

It is noteworthy that the optimal user pairing and greedy methods are used for the proposed approach and their performances are shown in Figs. 6 and 7. The performance difference is negligible as shown in Fig. 6, and becomes smaller as L increases as shown in Fig. 7.

Finally, the total transmission power is shown for various values of K for a fixed L in Fig. 8 with $L = 20$ and $R_{\mathrm{f}} = R_{\mathrm{n}} = 2$. We can see that the total transmission power increases with K as more users are to be supported. An interesting observation is that as K approaches $L/2$, the total transmission power can increase significantly due to the lack of degrees of freedom (when $K = 10$, M becomes 2). In Fig. 8, the optimal user pairing is considered up to $K = 5$ as its complexity becomes prohibitively high as K increases although it can provide a better performance than the greedy algorithm.

VI. Concluding Remarks

Multicast beamforming with SC was considered for multiresolution broadcast with two users (one is close to the BS and the other is not). To find beamforming vectors, we formulated a minimum power beamforming problem with SC when the target rates (or SNR/SINR) are given. It was shown that the power optimization problem becomes a linear programming for given normalized beamforming vectors, which has a closed-form expression. Using this closed-form expression and based on the nonlinear GS algorithm, an iterative algorithm was derived to find beamforming vectors. Unfortunately, it was not possible to guarantee that this iterative algorithm converges to the optimal solution. Simulation results showed that the proposed multicast beamforming with SC has a lower total transmission power than that of the conventional multiuser beamforming without SC. Furthermore, the proposed multicast beamforming with SC was applied to NOMA systems to support multiple (more than two) users with different data streams as a two-stage beamforming method. It was shown that the proposed two-stage beamforming method can perform better than an existing approach that also uses beamforming with SC.

References

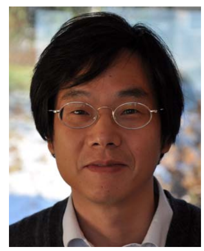
[1] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[2] R. Zhang and L. Hanzo, "A unified treatment of superposition coding aided communications: Theory and practice," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 3, pp. 503–520, 2011.

[3] S. Vanka *et al.*, "Superposition coding strategies: Design and experimental evaluation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2628–2639, Jul. 2012.

[4] "Digital video broadcasting (DVB); Framing structure, channel coding and modulation for satellite services to handheld devices (SH) below

3 GHz," Sophia-Antipolis Cedex, France, ETSI EN 302 583 V1.1.1 (2008-03), 2008.
[5] "LTE; Multimedia broadcast/multicast service (MBMS); Stage 1 (Version 11.1.0 Release 11)," Cedex, France, TS 22.146, 2013.
[6] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-Advanced for Mobile Broadband*, 2nd ed. New York, NY, USA: Academic, 2013.
[7] E. Larsson and B. Vojcic, "Cooperative transmit diversity based on superposition modulation," *IEEE Commun. Lett.*, vol. 9, no. 9, pp. 778–780, Sep. 2005.
[8] J. M. Park, S.-L. Kim, and J. Choi, "Hierarchically modulated network coding for asymmetric two-way relay systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2179–2184, Jun. 2010.
[9] N. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.
[10] M. Kaliszan, E. Pollakis, and S. Stanczak, "Multigroup multicast with application-layer coding: Beamforming for maximum weighted sum rate," in *Proc. IEEE WCNC*, Apr. 2012, pp. 2270–2275.
[11] D. Wajcer, S. Shamai, and A. Wiesel, "On superposition coding and beamforming for the multi-antenna Gaussian broadcast channels," in *Proc. Inf. Theory Appl. Workshop*, 2006, pp. 1–6, UCSD.
[12] P. Wang, J. Xiao, and L. Ping, "Comparison of orthogonal and non-orthogonal approaches to future wireless cellular systems," *IEEE Veh. Technol. Mag.*, vol. 1, no. 3, pp. 4–11, Sep. 2006.
[13] Y. Saito *et al.*, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE VTC Spring*, Jun. 2013, pp. 1–5.
[14] B. Kim *et al.*, "Non-orthogonal multiple access in a downlink multi-user beamforming system," in *Proc. IEEE MILCOM*, Nov. 2013, pp. 1278–1283.
[15] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraint," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.
[16] B. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989.
[17] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York, NY, USA: Wiley, 1980.
[18] T. Takeda and K. Higuchi, "Enhanced user fairness using non-orthogonal access with SIC in cellular uplink," in *Proc. IEEE VTC Fall*, 2011, pp. 1–5.
[19] J. Choi, "Non-orthogonal multiple access in downlink coordinated two-point systems," *IEEE Commun. Lett.*, vol. 18, no. 2, pp. 313–316, Feb. 2014.
[20] A. Hottinen and E. Viterbo, "Optimal user pairing in downlink MU-MIMO with transmit precoding," in *Proc. WiOPT*, Apr. 2008, pp. 97–99.



Jinho Choi (SM'02) was born in Seoul, Korea. He received B.E. (magna cum laude) degree in electronics engineering from Sogang University, Seoul, in 1989 and the M.S.E. and Ph.D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, in 1991 and 1994, respectively. He is with Gwangju Institute of Science and Technology (GIST) as a Professor. Prior to joining GIST in 2013, he was with the College of Engineering, Swansea University, United Kingdom, as a Professor/Chair in Wireless. His research interests include wireless communications and array/statistical signal processing. He authored two books published by Cambridge University Press in 2006 and 2010. Prof. Choi received the 1999 Best Paper Award for Signal Processing from EURASIP, 2009 Best Paper Award from WPMC (Conference), and is Senior Member of IEEE. Currently, he is an Associate Editor of IEEE COMMUNICATIONS LETTERS, Editor of JOURNAL OF COMMUNICATIONS AND NETWORKS (JCN) since 2005, and has served as Associate Editor of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from 2005 to 2007 and as Associate Editor of *ETRI Journal*.