Journal of the Chinese Institute of Engineers
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tcie20

A two-user approximation-based transmit beamforming for physical-layer multicasting in mobile cellular downlink systems

Chang-Lung Hsiao[a], Jiann-Ching Guey[b], Wern-Ho Sheen[c] & Ren-Jr Chen[a]
[a] ICL ITRI, 310 Hsinchu, Taiwan
[b] MediaTek, 300 Hsinchu, Taiwan
[c] Department of Communications Engineering, National Chung Cheng University, 621 Chiayi, Taiwan

Published online: 16 Mar 2015.

To cite this article: Chang-Lung Hsiao, Jiann-Ching Guey, Wern-Ho Sheen & Ren-Jr Chen (2015) A two-user approximation-based transmit beamforming for physical-layer multicasting in mobile cellular downlink systems, Journal of the Chinese Institute of Engineers, 38:6, 742-750, DOI: 10.1080/02533839.2015.1016879

To link to this article: http://dx.doi.org/10.1080/02533839.2015.1016879

Journal of the Chinese Institute of Engineers, 2015
Vol. 38, No. 6, 742–750, http://dx.doi.org/10.1080/02533839.2015.1016879

A two-user approximation-based transmit beamforming for physical-layer multicasting in mobile cellular downlink systems

Chang-Lung Hsiao[a]*, Jiann-Ching Guey[b], Wern-Ho Sheen[c] and Ren-Jr Chen[a]

[a]*ICL ITRI, 310 Hsinchu, Taiwan;* [b]*MediaTek, 300 Hsinchu, Taiwan;* [c]*Department of Communications Engineering, National Chung Cheng University, 621 Chiayi, Taiwan*

(*Received 3 December 2013; accepted 25 February 2014*)

Multicast has been known as an efficient transmission technique for group-oriented applications such as multi-party video conferencing, video streaming for paid users, online gaming, and social networking. In this paper, we investigate physical-layer multicasting in mobile cellular downlink systems, where the antennas at base station are employed to transmit common signals to multiple users simultaneously. A central design problem of downlink physical-layer multicasting is the search for the optimal beamforming vector that maximizes the multicast rate. Traditionally, the problem has been formulated as a quadratically constrained quadratic programming problem and shown to be NP-hard in general. In this paper, starting from examining the Karush–Kuhn–Tucker stationary conditions, a new method based on two-user approximation is proposed for the search for the optimal beamforming vector. The method is able to achieve a much higher multicast rate than the existing methods and provides an attractive trade-off between performance and complexity, especially for the case of using a large number of antennas. Using a large number of antennas at base station, also known as the large-scale multiple-input and multiple-output technique, has been regarded widely as one of the most promising technologies to increase system capacity, coverage, and user throughput for future generations of mobile cellular systems.

Keywords: physical-layer multicasting; mobile cellular downlink systems; multiple antennas; transmit beamforming

1. Introduction

Multicast is an efficient transmission technique for group-oriented applications such as multi-party video conferencing, video streaming for paid users, online gaming, and social networking. Multicast has been widely used in today's wired and wireless networks and will continue to be an important design in future generations of systems (3GPP 2013). In this paper, we investigate physical-layer multicasting for downlink cellular systems, where the antennas at base station are employed to transmit common signals to multiple users simultaneously. Physical-layer multicasting can be viewed as a form of multi-user multiple-input and multiple-output (MIMO) technique with the transmitted signals from base station intended for all involved users.

A central design problem of downlink physical-layer multicasting is the search for the optimal beamforming vector that maximizes the multicast rate (achievable rate by all users), with the help of channel station information at transmitter (CSIT) (Lopez 2002; Sun and Liu 2004; Sidiropoulos, Davidson, and Luo 2006; Hunger et al. 2007; Lozano 2007; Kim, Love, and Park 2011; Khojastepour, Salehi-Golsefidi, and Rangarajan 2011). Without CSIT, the common practice is to multicast the common signals isotropically to all users and use a transmit diversity technique such as STBCs (space-time block codes) to improve receiver performance (Sun and Liu 2004). In this paper, we focus on the beamforming design under CSIT which can be made available either by users' feedback in FDD systems or by exploiting the channel reciprocity in TDD systems.

In physical-layer multicasting, the multicast rate is limited by the user with the worst channel condition. Therefore, the common design goal was to maximize the minimum signal-to-noise ratio (SNR) among users while maintaining the same transmit power, known as the max–min fair problem, or equivalently to minimize the total transmit power under the constraint that the users' SNRs are larger than a given value, known as the quality-of-service (QoS) constraint problem (Sidiropoulos, Davidson, and Luo 2006).

One popular approach to solve the above optimization problem is to formulate it as a quadratically constrained quadratic programming (QCQP) problem. However, the constraints of the QCQP problem are non-convex, and the problem has been shown to be NP-hard in general (Sidiropoulos, Davidson, and Luo 2006). To avoid such an intractability, the semidefinite relaxation (SDR) technique was proposed in the work of Sidiropoulos, Davidson, and Luo (2006); by dropping the non-convex rank-one constraint, the original problem can be relaxed

*Corresponding author. Email: clhsiao@itri.org.tw

to a tractable semidefinite programming problem which can be solved by standard optimization software packages, such as SeDuMi (Sturm 1999). The obtained solution then can be used as an approximate solution to the original problem. Empirical evidence and theoretical analysis were provided by Sidiropoulos, Davidson, and Luo (2006); Luo et al. (2007); and Jindal and Luo (2006) to show the viability of the SDR-based techniques. Also, the SDR-based techniques have been applied to many different types of beamforming problems (Gershman et al. 2010) and have motivated the study of multicast methods with non-rank-one nature (Schad, Law, and Pesavento 2012; Wen et al. 2012; Wu, So, and Ma 2012; Wu, Ma, and So 2013).

In the works of Lozano (2007); Hunger et al. (2007); Abdelkader, Gershman, and Sidiropoulos (2010); and Kim, Love, and Park (2011), successive and/or iterative algorithms were proposed in search of the optimal beamforming vector, without using the above-mentioned high-complexity optimization software packages. The key idea behind these algorithms is to steer the beamforming vector in a direction to increase the received signal power of the user with the lowest SNR. The direction can be determined by a gradient method (Lozano 2007) or by a method of channel orthogonalization (Hunger et al. 2007; Abdelkader, Gershman, and Sidiropoulos 2010; Kim, Love, and Park 2011). By avoiding the use of optimization software packages, these algorithms can operate with a relatively low complexity and offer an attractive trade-off between performance and complexity.

In this paper, we focus on the design of the downlink physical-layer multicast beamforming; starting from examining the Karush–Kuhn–Tucker (KKT) stationary conditions, a new method based on two-user approximation is proposed for the search for the optimal beamforming vector. Our numerical results show that the proposed method can achieve a much higher multicast rate than the existing methods and provide an attractive performance-to-complexity trade-off, especially for the case of using a large number of antennas. Using a large number of antennas at base station, also known as the large-scale MIMO technique, has been regarded widely as one of the most promising technologies to increase system capacity, coverage, and user throughput for future generations of mobile cellular networks (Rusek et al. 2013).

The rest of the paper is organized as follows. System model and problem formulations are given in Section 2, followed by optimality analysis and the proposed method in Section 3 and Section 4, respectively. Simulation results are presented in Section 5, and a conclusion is given in Section 6.

Notation: Vectors and matrices are denoted by bold-faced lower and upper case letters, respectively. The operators $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^*$ stand for transposition, Hermitian transposition, and complex conjugate, respectively. $\angle(\cdot)$ denotes the argument of a complex scalar, and $E[\cdot]$ denotes the expectation operator.

2. System model and problem formulation

Consider a mobile cellular downlink system consisting of a base station with N antennas and M single-antenna users. A common signal s with average power of σ_s^2 is to be transmitted to all users from the base station. Let the $1 \times N$ complex vector $\boldsymbol{h}_i^H = \left[h_{i,1}^*, \cdots, h_{i,N}^*\right]$ denote the channel response (frequency non-selective) between the base station and user i, and let $\boldsymbol{p} = [p_1, \cdots, p_N]^T$ denote the beamforming vector. Then, the received signal at user i is given by

$$r_i = \boldsymbol{h}_i^H \boldsymbol{p} s + n_i,\ i = 1, \cdots, M, \tag{1}$$

where $n_i \sim \mathcal{CN}\left(0, \sigma_n^2\right)$ is the additive white Gaussian noise (AWGN). The design goal was to optimize the beamforming vector that maximizes the multicast rate, under a total power constraint at the base station. The multicast rate is defined as the maximum rate that is achievable by all users; in other words, there will be some users not able to be served if the transmission rate is higher than the multicast rate. Clearly, the multicast rate is limited by the user with the worst channel condition.

With CSIT, the Gaussian capacity of the ith user is given by

$$R_i = \log\left(1 + \frac{\sigma_s^2}{\sigma_n^2}\boldsymbol{p}^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{p}\right). \tag{2}$$

Since logarithm is a monotonically increasing function, the problem of the design of the optimal beamforming vector becomes a need to solve the optimization problem,

$$\begin{aligned} \mathcal{P}1 : \hat{\boldsymbol{p}} = \arg\max_{\boldsymbol{p}} & \left(\min_{i \in \{1,\cdots,M\}} \boldsymbol{p}^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{p}\right) \\ \text{subject to} & \, ||\boldsymbol{p}||_2^2 \le p_0, \end{aligned} \tag{3}$$

where p_0 is the maximum total transmitting power. It is straightforward to show that the power constraint in Equation (3) should be met with equality at the optimal solution. Therefore, without loss of generality, we assume $p_0 = 1$, and the power constraint is rewritten as follows: $||\boldsymbol{p}||_2^2 = 1$.

Alternatively, the beamforming vector $\boldsymbol{p}$ can be chosen to minimize the total transmit power while satisfying a minimum SNR requirement, $\gamma_{\min}$, for all users to provide a required QoS. In this case, the optimization problem becomes

$$\mathcal{P}2: \quad \hat{\boldsymbol{p}} = \arg\min_{\boldsymbol{p}}\left(||\boldsymbol{p}||_2^2\right) \\ \text{subject to } \boldsymbol{p}^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{p} \geq \gamma_{\min}\sigma_n^2/\sigma_s^2, \\ i \in \{1, \cdots, M\} \tag{4}$$

In the work of Sidiropoulos, Davidson, and Luo (2006), it was shown that $\mathcal{P}1$ and $\mathcal{P}2$ are NP-hard QCQP problems and equivalent to each other up to a scaling. In what follows, we just focus on the problem $\mathcal{P}1$.

3. The optimality

The max–min fair problem $\mathcal{P}1$ can be converted to the following standard form (Luenberger and Ye 2008)

$$\begin{aligned} \mathcal{P}1': \text{minimize} \quad & f(\boldsymbol{p}, x) = -x \\ \text{subject to} \quad & h(\boldsymbol{p}, x) = ||\boldsymbol{p}||_2^2 - 1 = 0 \\ & \boldsymbol{b}(\boldsymbol{p}, x) = \begin{bmatrix} b_1(\boldsymbol{p}, x) \\ \vdots \\ b_M(\boldsymbol{p}, x) \end{bmatrix} \leq 0 \\ & x \geq 0, \end{aligned} \tag{5}$$

where $b_i(\boldsymbol{p}, x) = x - \boldsymbol{p}^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{p}, i \in \{1, \cdots, M\}$. And, the KKT necessary conditions for the optimality at a feasible point are as follows:

$$\nabla f(\boldsymbol{p}, x) + \lambda \nabla h(\boldsymbol{p}, x) + \boldsymbol{\mu}^T \nabla \boldsymbol{b}(\boldsymbol{p}, x) = 0, \tag{6}$$

and

$$\boldsymbol{\mu}^T \boldsymbol{b}(\boldsymbol{p}, x) = \boldsymbol{0}, \tag{7}$$

where λ and $\boldsymbol{\mu}^T = [\mu_1, \cdots, \mu_M] \geq \boldsymbol{0}$ are the real-valued Lagrangian multipliers. From Equation (6), it can be shown that

$$\left(\sum_{i=1}^{M} \mu_i \boldsymbol{h}_i \boldsymbol{h}_i^H\right) \boldsymbol{p} = \lambda \boldsymbol{p}, \tag{8}$$

and $\sum_{i=1}^{M} \mu_i = 1$. In addition, μ_i is nonzero only when the constraint $b_i(\boldsymbol{p}, x) \leq 0$ is active; therefore, from Equation (7), we have

$$x - \boldsymbol{p}^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{p} = 0, i \in \{m | \mu_m \neq 0\}. \tag{9}$$

Multiplying both sides of Equation (8) with $\boldsymbol{p}^H$ leads to

$$\boldsymbol{p}^H \left(\sum_{i=1}^{M} \mu_i \boldsymbol{h}_i \boldsymbol{h}_i^H\right) \boldsymbol{p} = \lambda \boldsymbol{p}^H \boldsymbol{p}, \tag{10}$$

or equivalently

$$\sum_{i=1}^{M} \mu_i \boldsymbol{p}^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{p} = \lambda. \tag{11}$$

It is clear that when there is an optimal solution satisfying Equations (8)–(11), the received signal power of each active user $\boldsymbol{p}^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{p}, i \in \{m | \mu_m \neq 0\}$ is equal to the optimal value $x_{\max}$ and hence equal to the corresponding eigenvalue $\lambda_{\max}$. In other words, the optimal beamforming vector $\boldsymbol{p}$ is a eigenvector corresponding to the eigenvalue $\lambda_{\max}$ of the matrix $\left(\sum_{i=1}^{M} \mu_i \boldsymbol{h}_i \boldsymbol{h}_i^H\right)$.

Another observation is that the Lagrangian multipliers $\{\mu_i\}_{i=1}^{M}$ not only indicate the set of active inequality constraints but also play the roles of weighting the channel covariance matrix of users. According to the derived KKT condition of Equation (8), given a set of Lagrangian multipliers $\{\mu_i\}_{i=1}^{M}$, the beamforming vector $\boldsymbol{p}$ is the eigenvector of the matrix $\sum_{i=1}^{M} \mu_i \boldsymbol{h}_i \boldsymbol{h}_i^H$ and the channel gains with beamforming of each user are $\left\{\boldsymbol{p}^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{p}\right\}_{i=1}^{M}$, which may not satisfy the second derived KKT condition of Equation (9). It should be noted that increasing a specific Lagrangian multiplier, for example μ_m, results in increasing the weight of $\boldsymbol{h}_m \boldsymbol{h}_m^H$ in the matrix $\sum_{i=1}^{M} \mu_i \boldsymbol{h}_i \boldsymbol{h}_i^H$ and steering the beamforming vector $\boldsymbol{p}$ toward $\boldsymbol{h}_m$, and hence enhances the SNR of user m. This observation provides the physical meaning of the proposed iterative searching algorithm in Section 4.

4. The proposed two-user approximation method

Since finding the optimal beamforming vector in Equation (5) is an NP-hard problem, in this section, a new method based on two-user approximation is proposed. As is shown in Section 5, the proposed method can achieve a much higher multicast rate than the existing methods and provides an attractive tradeoff between performance and complexity, especially for the case of using a large number of antennas. The basic idea of the method is as follows. First, the optimal beamforming vector is derived in closed form for the two-user case. Second, an iterative algorithm, called two-user approximation algorithm (TUAA), is used to search for the optimal beamforming vector for the case of more than two users. In essence, in a new iteration, the user with the lowest SNR (using the beamforming vector obtained in previous iteration) is singled out, and all other users are treated collectively as a composite user. Then, the problem can be solved by the optimal two-user method, aiming to increase the lowest SNR. The iterations then continue until the maximum number of iterations is reached.

4.1. *Optimal solution for two users*

For the case of two users, it follows from Equations (8)–(9) that the optimal $\boldsymbol{p}$ satisfies the equations

$$\left(\mu_1 \boldsymbol{h}_1 \boldsymbol{h}_1^H + \mu_2 \boldsymbol{h}_2 \boldsymbol{h}_2^H\right) \boldsymbol{p} = \lambda \boldsymbol{p}, \tag{12}$$

and

$$\boldsymbol{p}^H \boldsymbol{h}_1 \boldsymbol{h}_1^H \boldsymbol{p} = \boldsymbol{p}^H \boldsymbol{h}_2 \boldsymbol{h}_2^H \boldsymbol{p} = \lambda, \tag{13}$$

with $\mu_1 + \mu_2 = 1$ and $\mu_1, \mu_2 > 0$. In addition, from Equation (13), $\boldsymbol{h}_1^H\boldsymbol{p} = \sqrt{\lambda}e^{j\theta_1}$, and $\boldsymbol{h}_2^H\boldsymbol{p} = \sqrt{\lambda}e^{j\theta_2}$ for some θ_1 and θ_2. Defining $\boldsymbol{H} = [\boldsymbol{h}_1\boldsymbol{h}_2]$ and $\boldsymbol{\mu} = [\mu_1\mu_2]^T$, Equation (12) becomes

$$\boldsymbol{H}\begin{bmatrix} \sqrt{\lambda}e^{j\theta_1} & 0 \\ 0 & \sqrt{\lambda}e^{j\theta_2} \end{bmatrix}\boldsymbol{\mu} = \lambda\boldsymbol{p}. \tag{14}$$

In a real environment, $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ can be considered as linearly independent, otherwise the two users will have the same channel response up to a scaling. In this case, by multiplying both sides of Equation (14) with $\left(\boldsymbol{H}^H\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H$, we have

$$\begin{bmatrix} \sqrt{\lambda}e^{j\theta_1} & 0 \\ 0 & \sqrt{\lambda}e^{j\theta_2} \end{bmatrix}\boldsymbol{\mu} = \lambda\left(\boldsymbol{H}^H\boldsymbol{H}\right)^{-1}\boldsymbol{H}^H\boldsymbol{p}. \tag{15}$$

Furthermore, using the notation $\boldsymbol{H}^H\boldsymbol{H} \triangleq \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{12}^* & \alpha_{22} \end{bmatrix}$, it follows that

$$\boldsymbol{\mu} = \frac{\lambda}{\alpha_{11}\alpha_{22} - |\alpha_{12}|^2}\begin{bmatrix} \alpha_{22} - \alpha_{12}e^{j(\theta_2-\theta_1)} \\ \alpha_{11} - \alpha_{12}^*e^{j(\theta_1-\theta_2)} \end{bmatrix}. \tag{16}$$

Since μ_1 and μ_2 are real-valued and $\mu_1 + \mu_2 = 1$, the solutions to Equation (16) are given by

$$\boldsymbol{\mu}^+ \triangleq \begin{bmatrix} \mu_{1,+} \\ \mu_{2,+} \end{bmatrix} = \frac{1}{\alpha_{11} + \alpha_{22} - 2|\alpha_{12}|}\begin{bmatrix} \alpha_{22} - |\alpha_{12}| \\ \alpha_{11} - |\alpha_{12}| \end{bmatrix}, \tag{17}$$

with

$$\lambda^+ = \frac{\alpha_{11}\alpha_{22} - |\alpha_{12}|^2}{\alpha_{11} + \alpha_{22} - 2|\alpha_{12}|}, \tag{18}$$

and $\theta_2 - \theta_1 = -\angle(\alpha_{12}) + 2n\pi, n \in \mathbb{Z}$, and

$$\boldsymbol{\mu}^- \triangleq \begin{bmatrix} \mu_{1,-} \\ \mu_{2,-} \end{bmatrix} = \frac{1}{\alpha_{11} + \alpha_{22} + 2|\alpha_{12}|}\begin{bmatrix} \alpha_{22} + |\alpha_{12}| \\ \alpha_{11} + |\alpha_{12}| \end{bmatrix}, \tag{19}$$

with

$$\lambda^- = \frac{\alpha_{11}\alpha_{22} - |\alpha_{12}|^2}{\alpha_{11} + \alpha_{22} + 2|\alpha_{12}|}, \tag{20}$$

and $\theta_2 - \theta_1 = -\angle(\alpha_{12}) + (2n+1)\pi, n \in \mathbb{Z}$. Since $\lambda^+ \geq \lambda^-$, from Equation (11) and the discussions that follow, the optimal beamforming vector is obtained with $\boldsymbol{\mu} = \boldsymbol{\mu}^+$ and $\lambda = \lambda^+$, that is,

$$\begin{aligned} \boldsymbol{p} &= \frac{1}{\lambda^+}\boldsymbol{H}\begin{bmatrix} \sqrt{\lambda^+}e^{j\theta_1} & 0 \\ 0 & \sqrt{\lambda^+}e^{j\theta_2} \end{bmatrix}\boldsymbol{\mu}^+ \\ &= \frac{1}{\sqrt{\lambda^+}}\left(\mu_{1,+}\boldsymbol{h}_1 + \mu_{2,+}\boldsymbol{h}_2e^{-j\angle(\alpha_{12})}\right)e^{j\theta_1}, \end{aligned} \tag{21}$$

where θ_1 can be set to zero without loss of any generality. If $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ are linearly dependent, on the other hand, it is clear that the optimal beamforming vector $\boldsymbol{p}$ is the channel response that is normalized to satisfy the power constraint.

4.2. *Two-user approximation algorithm*

The proposed TUAA consists of two steps: finding a composite channel for all other users if the user with the lowest SNR (using the beamforming vector obtained in previous iteration) is singled out, followed by applying the proposed optimal two-user method to find a new beamforming vector to increase the lowest SNR.

From Equation (8), for given Lagrangian multipliers $\{\mu_i\}_{i=1}^M$, the value λ and beamforming vector $\boldsymbol{p}$ satisfy the equation

$$\left(\mu_1\boldsymbol{h}_1\boldsymbol{h}_1^H + \mu_2\boldsymbol{h}_2\boldsymbol{h}_2^H + \cdots + \mu_M\boldsymbol{h}_M\boldsymbol{h}_M^H\right)\boldsymbol{p} = \lambda\boldsymbol{p}. \tag{22}$$

Suppose user m is the one with the lowest SNR (using $\boldsymbol{p}$), Equation (22) is rewritten as follows:

$$\left(\mu_m\boldsymbol{h}_m^H\boldsymbol{p}\right)\boldsymbol{h}_m + \sum_{i\in\{1,\cdots,M\}\backslash m}\left(\mu_i\boldsymbol{h}_i^H\boldsymbol{p}\right)\boldsymbol{h}_i = \lambda\boldsymbol{p}, \tag{23}$$

where $\left\{\boldsymbol{h}_i^H\boldsymbol{p}\right\}_{i=1}^M$ are complex numbers. For the proposed two-user approximation, the composite channel $\boldsymbol{g}$ shall be found so that

$$\left(\mu_m\boldsymbol{h}_m\boldsymbol{h}_m^H + (1-\mu_m)\boldsymbol{g}\boldsymbol{g}^H\right)\boldsymbol{p} = \lambda\boldsymbol{p}, \tag{24}$$

that is

$$\sum_{i\in\{1,\cdots,M\}\backslash m}\left(\mu_i\boldsymbol{h}_i^H\boldsymbol{p}\right)\boldsymbol{h}_i = \left((1-\mu_m)\boldsymbol{g}^H\boldsymbol{p}\right)\boldsymbol{g}, \tag{25}$$

with the notation

$$\boldsymbol{\Phi}_m \triangleq \sum_{i\in\{1,\cdots,M\}\backslash m}\mu_i\boldsymbol{h}_i\boldsymbol{h}_i^H, \tag{26}$$

Equation (23) becomes

$$\boldsymbol{\Phi}_m\boldsymbol{p} = \left((1-\mu_m)\boldsymbol{g}^H\boldsymbol{p}\right)\boldsymbol{g}, \tag{27}$$

where the right-hand side is the vector $\boldsymbol{g}$ multiplied by a complex scalar, implying that $\boldsymbol{g}$ is a vector parallel to the vector $\boldsymbol{\Phi}_m\boldsymbol{p}$. Denoting $\boldsymbol{\xi} \triangleq \boldsymbol{\Phi}_m\boldsymbol{p}$ and $\boldsymbol{g} = \rho\boldsymbol{\xi}$, from Equation (27), we obtain

$$\boldsymbol{\xi} = (1-\mu_m)(\rho\boldsymbol{\xi})^H\boldsymbol{p}(\rho\boldsymbol{\xi}) = (1-\mu_m)\rho^2\boldsymbol{p}^H\boldsymbol{\Phi}_m\boldsymbol{p}\boldsymbol{\xi}, \tag{28}$$

$$\rho = \frac{1}{\sqrt{(1-\mu_m)\boldsymbol{p}^H\boldsymbol{\Phi}_m\boldsymbol{p}}}, \tag{29}$$

and

$$\boldsymbol{g} = \frac{\boldsymbol{\Phi}_m\boldsymbol{p}}{\sqrt{(1-\mu_m)\boldsymbol{p}^H\boldsymbol{\Phi}_m\boldsymbol{p}}}. \tag{30}$$

Also, Equation (22) can be rewritten as $\boldsymbol{\Phi}_m\boldsymbol{p} + \mu_m\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{p} = \lambda\boldsymbol{p}$, and that leads to

$\boldsymbol{p}^H \boldsymbol{\Phi}_{\boldsymbol{m}} \boldsymbol{p} = \lambda - \mu_m \boldsymbol{p}^H \boldsymbol{h}_{\boldsymbol{m}} \boldsymbol{h}_{\boldsymbol{m}}^H \boldsymbol{p}$. Therefore, the composite channel is given by

$$\boldsymbol{g} = \rho\left(\lambda \boldsymbol{p} - \mu_m \boldsymbol{h}_{\boldsymbol{m}} \boldsymbol{h}_{\boldsymbol{m}}^H \boldsymbol{p}\right), \tag{31}$$

where

$$\rho = \frac{1}{\sqrt{(1-\mu_m)\left(\lambda - \mu_m |\boldsymbol{h}_{\boldsymbol{m}}^H \boldsymbol{p}|^2\right)}}. \tag{32}$$

Note from Equations (26) and (30) that the composite channel is a weighted combination of $\boldsymbol{h}_{\boldsymbol{i}}$, $i \in \{1, \cdots, M\} \backslash m$.

With the composite channel $\boldsymbol{g}$, the optimal two-user method for optimal beamforming vector in Section 4.1. can be applied using $\boldsymbol{H} = [\boldsymbol{h}_{\boldsymbol{m}}\ \boldsymbol{g}]$, $\alpha_{11} = \boldsymbol{h}_{\boldsymbol{m}}^H \boldsymbol{h}_{\boldsymbol{m}}$, $\alpha_{22} = \rho^2\left[\lambda^2 - 2\lambda\mu_m |\boldsymbol{h}_{\boldsymbol{m}}^H \boldsymbol{p}|^2 + \mu_m^2 \alpha_{11} |\boldsymbol{h}_{\boldsymbol{m}}^H \boldsymbol{p}|^2\right]$, $|\alpha_{12}| = \rho|\lambda - \mu_m \alpha_{11}||\boldsymbol{h}_{\boldsymbol{m}}^H \boldsymbol{p}|$.

As mentioned previously, the idea behind the TUAA searching algorithm is to update the Lagrangian multipliers $\{\mu_i\}_{i=1}^M$ to increase the minimum SNR among users. In the kth iteration, the user with the lowest SNR under current beamforming vector $\boldsymbol{p}^{(k)}$ shall be singled out first, says user m with channel response $\boldsymbol{h}_{\boldsymbol{m}}$. The composite channel $\boldsymbol{g}^{(k)}$ relative to user m under $\boldsymbol{p}^{(k)}$ is obtained by Equation (31). To update the Lagrangian multiplier $\mu_m^{(k)}$ corresponding to user m, the optimal two-user solution in Section 4.1. then is applied with $\boldsymbol{h}_{\boldsymbol{m}}$ and $\boldsymbol{g}^{(k)}$ to obtain an updated $\mu_m^{(k+1)}$ corresponding to user m and a new beamforming vector $\boldsymbol{p}^{(k+1)}$. And finally, the weights of the other users are updated by

$$\mu_i^{(k+1)} = \frac{1-\mu_m^{(k+1)}}{1-\mu_m^{(k)}} \mu_i^{(k)}, \tag{33}$$

for $i = 1, \ldots, M$ and $i \neq m$. The detailed algorithm is described in Table 1, where $k_{\max}$ is the maximum number of iterations and the M channel responses of each user are taken as the initial points. Benefiting from approximately reducing the multiple users optimization problem to a two-user case and applying a close-form

Table 1. Two-user approximation algorithm (TUAA).

Input:	$\{\boldsymbol{h}_{\boldsymbol{i}}\}_{i=1}^M$, $k_{\max}$
Output:	$\boldsymbol{p}$
1:	for $c = 1 : M$ loop for each of the M users
2:	$\boldsymbol{\mu} = \boldsymbol{0}, \boldsymbol{\mu}(c) = 1$ set the weight of user c with one and zeros for the other users
3:	$\boldsymbol{p} = \boldsymbol{h}_{\boldsymbol{c}}/|\boldsymbol{h}_{\boldsymbol{c}}|$, $\lambda = \boldsymbol{p}^H \boldsymbol{h}_{\boldsymbol{c}} \boldsymbol{h}_{\boldsymbol{c}}^H \boldsymbol{p}$ set the initial beamforming vector and value λ
4:	$m = \arg\ \min_{i \in \{1,\cdots,M\}} \boldsymbol{p}^H \boldsymbol{h}_{\boldsymbol{i}} \boldsymbol{h}_{\boldsymbol{i}}^H \boldsymbol{p}$ target the user with minimum received signal power
5:	$k = 0$
6:	while $k < k_{\max} \&\& \boldsymbol{\mu}(m) < 1$ loop with $k_{\max}$ times at most
7:	$\rho = \frac{1}{\sqrt{(1-\boldsymbol{\mu}(m))\left(\lambda - \boldsymbol{\mu}(m)|\boldsymbol{h}_{\boldsymbol{m}}^H \boldsymbol{p}|^2\right)}}$, $\boldsymbol{g} = \rho\left(\lambda \boldsymbol{p} - \boldsymbol{\mu}(m) \boldsymbol{h}_{\boldsymbol{m}} \boldsymbol{h}_{\boldsymbol{m}}^H \boldsymbol{p}\right)$ obtain the composite channel relative to user m according to Equations (31) and (32)
8:	$\left[\boldsymbol{\mu}^+, \lambda^+, \boldsymbol{p}\right] = \mathrm{Opt2}(\boldsymbol{h}_{\boldsymbol{m}}, \boldsymbol{g})$ find the new weight for user m and new $\boldsymbol{p}$ by applying the close-form optimal solution to $\boldsymbol{h}_{\boldsymbol{m}}$ and $\boldsymbol{g}$
9:	$\boldsymbol{\mu} = \frac{1-\boldsymbol{\mu}^+(1)}{1-\boldsymbol{\mu}(m)} \boldsymbol{\mu}$, $\boldsymbol{\mu}(m) = \boldsymbol{\mu}^+(1)$ rescaling the $\boldsymbol{\mu}$
10:	$\lambda = \lambda^+$ update value λ
11:	$m = \arg\ \min_{i \in \{1,\cdots,M\}} \boldsymbol{p}^H \boldsymbol{h}_{\boldsymbol{i}} \boldsymbol{h}_{\boldsymbol{i}}^H \boldsymbol{p}$ find the user with minimum signal power under new $\boldsymbol{p}$.
13:	$k = k + 1$
14:	end while
15:	$\boldsymbol{p}_c = \boldsymbol{p}$, $\mathrm{gain}_c = \boldsymbol{p}^H \boldsymbol{h}_{\boldsymbol{m}} \boldsymbol{h}_{\boldsymbol{m}}^H \boldsymbol{p}$ save the cth candidate vector
16:	end for
17:	$\boldsymbol{p} = \boldsymbol{p}_{\arg\ \max_{i \in \{1,\cdots,M\}} \mathrm{gain}_i}$ Select the optimal vector from M candidates

Note: $k_{\max}$ can be configured for performance and complexity trade-off.

solution with only vector operations at each iteration, the complexity of TUAA is given by $O(Mk_{\max}(N + MN))$.

5. Simulation results

In this section, the performance of the proposed method (denoted as TUAA) is evaluated by computer simulations and compared to the methods of GSdL (Gram–Schmidt orthogonalization with damped Lozano's local refinement) and QRdL (QR decomposition with damped Lozano's local refinement) in Abdelkader, Gershman, and Sidiropoulos (2010) and the method of dLLI (damped Lozano with Lopez Initialization) in Matskani et al. (2009). In Abdelkader, Gershman, and Sidiropoulos (2010), it was shown that QRdL provides substantially higher multicast rate than other beamforming methods under a similar complexity to the SDR-based approaches, and GSdL suffers from a slight loss in multicast rate but with a lower computational complexity, as compared with QRdL. dLLI is compared here mainly for its low computational complexity, and the multicast capacity (Jindal and Luo 2006) is also included for the purpose of benchmarking. As reported in Abdelkader, Gershman, and Sidiropoulos (2010), the complexity of QRdL and GSdL is $O(J(N^3 + MN))$ and $O(MN^3 + M^2N)$, respectively, where J is the number of initial points used for the search of the optimal beamforming vector in QRdL and is a design parameter used to trade-off between performance and complexity.

The channel vector between base station and user i is defined as follows: $\boldsymbol{h_i} = \left[c_{i,1}, \cdots c_{i,N}\right]^T$, where $\left\{c_{i,j}\right\}_{j=1}^{N}$, $\forall i$ are $\mathcal{CN}(0,1)$ independent random variables, and the performance is evaluated with $\frac{\sigma_s^2}{N\sigma_n^2}E\left[\boldsymbol{h_i}^H\boldsymbol{h_i}\right] = 0\,\text{dB}$. The multicast rate and computational complexity (in terms of MATLAB computation time) are averaged values over 1000 randomly generated channel realizations. In all simulations, unless specified otherwise, $k_{\max} = 100$ is used for TUAA and $J = 200$ for QRdL as suggested in Abdelkader, Gershman, and Sidiropoulos (2010).

Figures 1 and 2, the comparisons are made under different numbers of transmit antennas (N) and users (M). As we can see from Figure 1, TUAA performs between QRdL and GSdL in terms of multicast rate and computational complexity for cases of small numbers of transmit antennas $(N = 8)$. For cases of large numbers of antennas $(N = 64)$, however, TUAA significantly outperforms QRdL and GSdL in multicast rate but with a computational complexity comparable to or even lower

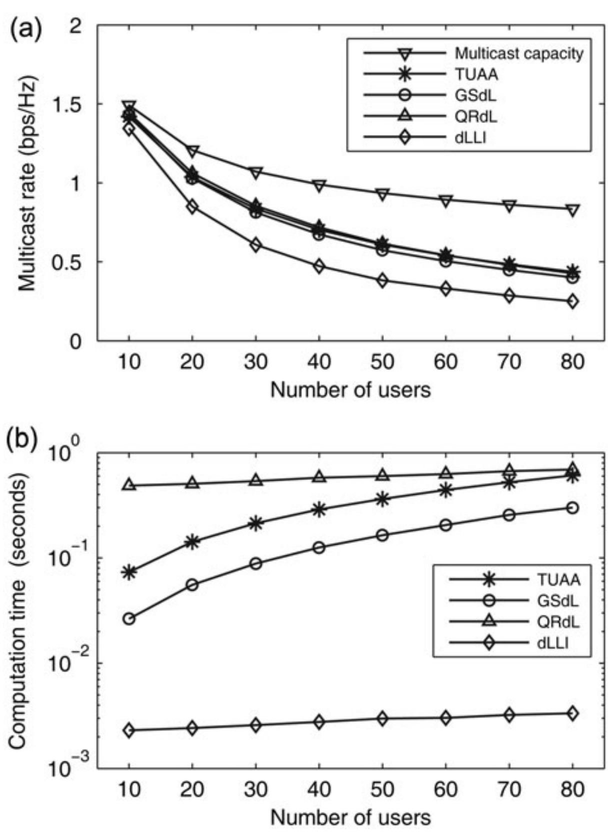

Figure 1. Comparison of different beamforming methods with $N = 8$: (a) multicast rate and (b) computational complexity.

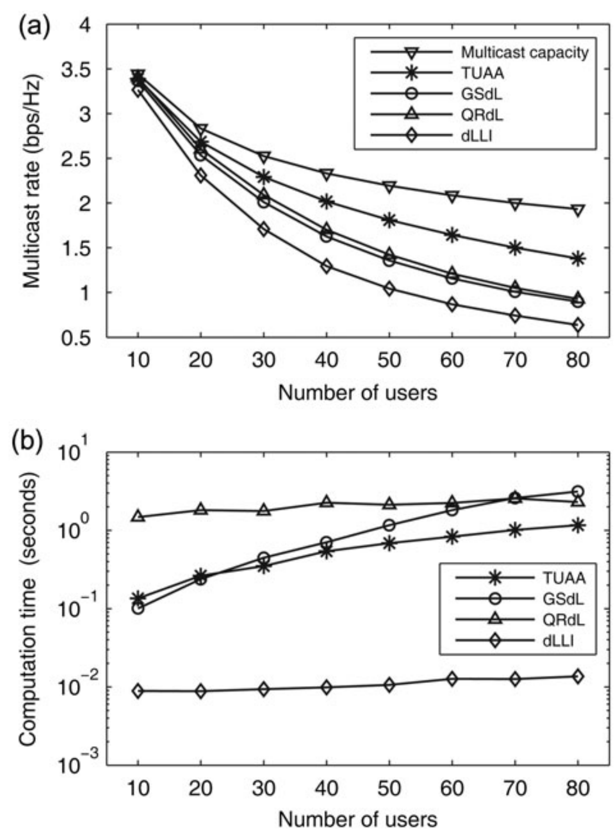

Figure 2. Comparison of different beamforming methods with $N = 64$: (a) multicast rate and (b) computational complexity.

than that of GSdL, as shown in Figure 2. For example, the improvements in multicast rate are about 9.7% for $(N = 64, M = 30)$ and 48.8% for $(N = 64, M = 80)$, respectively. Recall that the large-scale MIMO technique has been regarded widely as one of the most promising technologies to increase system capacity, coverage, and user throughput for future generations of mobile cellular systems. As expected, dLLI has the lowest complexity but with a multicast rate much lower than the others.

Figure 3 shows the comparison results under different numbers of transmit antennas, with the user number fixed at $M = 30$. Again, as is seen, TUAA outperforms GSdL and QRdL in multicast rate for the cases of using larger numbers of transmit antennas, say $N > 16$. Also, it is noted that the improvement provided by TUAA increases with the number of transmit antennas, with the computational complexity increasing at a rate lower than all other methods.

Figure 4 shows the performance of TUAA under different maximum iteration numbers, with the number of transmit antennas fixed at $N = 32$. It is seen that the multicast rate of TUAA can be further enhanced at the expense of a larger computational complexity by increasing $k_{\max}$. Using QRdL, on the other hand, the improvement is quite marginal, even if J is increased from 200 to 800, in which case the computational complexity is increased by about 4 times. Also, it is worthy of noting that in this example with $k_{\max} = 50$, TUAA has a multicast rate comparable to QRdL and GSdL but has the lowest complexity.

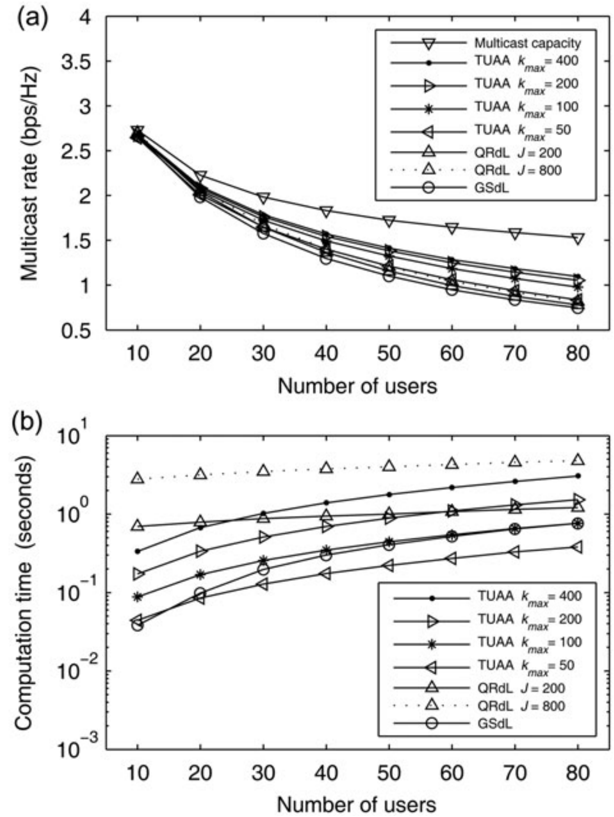

Figure 4. Performance of TUAA under different maximum numbers of iterations ($N = 32$): (a) multicast rate and (b) computational complexity.

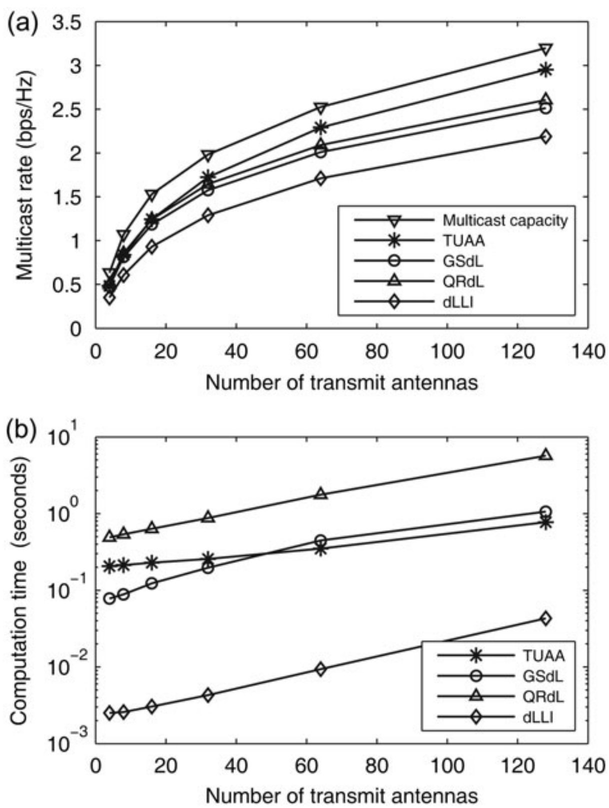

Figure 3. Comparison of different beamforming methods with $M = 30$: (a) multicast rate and (b) computational complexity.

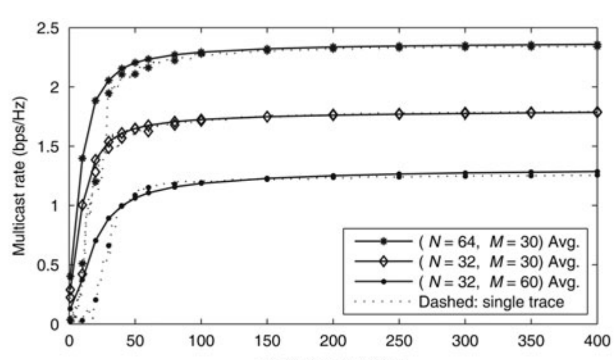

Figure 5. The convergence behavior of TUAA with $(N, M) = (64, 30)$, $(32, 30)$, and $(32, 60)$.

The convergence behavior of TUAA is investigated with computer simulations under different parameters. Figure 5 shows the convergence behavior averaged over 1000 realizations as well as that of a typical realization for the cases of $(N, M) = (64, 30)$, $(32, 30)$, and (32, 60). It is seen that the convergence of the proposed TUAA is quite insensitive to the number of transmit antennas and the number of users. Similar behaviors are observed for different system setups. It also shows that $k_{\max} = 50$ is essential to bring the TUAA into full play and using $k_{\max} > 50$ basically is a trade-off between performance and complexity. Unfortunately, it would be very difficult, if not impossible, to select $k_{\max}$ in an analytical way. From our extensive simulation results, $k_{\max} = 100$ is selected to provide a good performance/complexity trade-off for the considered systems.

6. Conclusion

The important problem of physical-layer multicasting is investigated in this paper for downlink mobile cellular systems. A two-user approximation method is proposed for use in the search for the optimal beamforming vector that maximizes the multicast rate. The performance of the proposed method is evaluated through extensive computer simulations. As compared to the state-of-the-art methods in the literature, the proposed method can achieve a much higher multicast rate and provide an attractive tradeoff between performance and complexity, especially for cases using a large number of antennas. Using a large number of antennas is regarded as one of the key enabling technologies in future mobile cellular systems beyond the 4th generation.

Nomenclature

$\boldsymbol{g}$	the composite channel vector
$\boldsymbol{h}_i$	the channel response (frequency non-selective) between the base station and user i
J	the number of initial points used for the search of the optimal beamforming vector in QRdL
$k_{\max}$	the maximum number of iterations in TUAA
λ, μ_i	the real-valued Lagrangian multipliers
M	number of users
N	number of transmit antennas
n_i	the AWGN of user i
$\boldsymbol{p}$	the beamforming vector
r_i	the received signal at user i

Funding

This work was supported by the Ministry of Economic Affairs (MOEA), ROC [grant number 102-EC-17-A-03-01-0620].

References

Abdelkader, A., A. B. Gershman, and N. D. Sidiropoulos. 2010. "Multiple-antenna Multicasting Using Channel Orthogonalization and Local Refinement." *IEEE Transactions on Signal Processing* 58 (7): 3922–3927. doi:10.1109/TSP.2010.2046436.

Gershman, A. B., N. D. Sidiropoulos, S. ShahbazPanahi, M. Bengtsson, and B. Ottersten. 2010. "Convex Optimization-based Beamforming." *IEEE Signal Processing Magazine* 27 (3): 62–75. doi:10.1109/MSP.2010.936015.

3GPP (3rd Generation Partnership Project) 2013. *Evolved Universal Terrestrial Radio Access Network (E-UTRAN); General Aspects and Principles for Interfaces Supporting Multimedia Broadcast Multicast Service (MBMS) within E-UTRAN (Release 11)*. 3GPP TS 36.440 V11.2.0. Valbonne, France: 3rd Generation Partnership Project.

Hunger, R., D. A. Schmidt, M. Joham, A. Schwing, and W. Utschick. 2007. "Design of Single-group Multicasting-beamformers." In *Proceedings of IEEE International Conference on Communications*, Glasgow, Scotland, 24–28 June 2007: 2499–2505. Washington, DC: IEEE.

Jindal, N., and Z.-Q. Luo. 2006. "Capacity Limits of Multiple Antenna Multicast." In *Proceedings of IEEE International Symposium on Information Theory*, Seattle, WA, 9–14 July 2006: 1841–1845. Washington, DC: IEEE.

Khojastepour, M. A., A. Salehi-Golsefidi, and S. Rangarajan. 2011. "Towards an Optimal Beamforming Algorithm for Physical Layer Multicasting." In *Proceedings of IEEE Information Theory Workshop*, Paraty, Brazil, 16–20 October 2011: 395–399. Washington, DC: IEEE.

Kim, I. H., D. J. Love, and S. Y. Park. 2011. "Optimal and Successive Approaches to Signal Design for Multiple Antenna Physical Layer Multicasting." *IEEE Transactions on Communications* 59 (8): 2316–2327. doi:10.1109/TCOMM.2011.060911.080464.

Lopez, M. J. 2002. "Multiplexing, Scheduling, and Multicasting Strategies for Antenna Arrays in Wireless Networks." PhD dissertation, Massachusetts Institute of Technology.

Lozano, A. 2007. "Long-term Transmit Beamforming for Wireless Multicasting." In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, Honolulu, HI, 15–20 April 2007: 417–420, III. Washington, DC: IEEE.

Luenberger, D. G., and Y. Ye. 2008. *Linear and Nonlinear Programming*. New York: Springer.

Luo, Z.-Q., N. D. Sidiropoulos, P. Tseng, and S. Zhang. 2007. "Approximation Bounds for Quadratic Optimization with Homogeneous Quadratic Constraints." *SIAM Journal on Optimization* 18 (1): 1–28. doi:10.1137/050642691.

Matskani, E., N. D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas. 2009. "Efficient Batch and Adaptive Approximation Algorithms for Joint Multicast Beamforming and Admission Control." *IEEE Transactions on Signal Processing* 57 (12): 4882–4894. doi:10.1109/TSP.2009.2026621.

Rusek, F., D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson. 2013. "Scaling up MIMO: Opportunities and Challenges with Very Large Arrays." *IEEE Signal Processing Magazine* 30 (1): 40–60. doi:10.1109/MSP.2011.2178495.

Schad, A., K. L. Law, and M. Pesavento. 2012. "A Convex Inner Approximation Technique for Rank-two Beamforming in Multicasting Relay Networks." In *Proceedings of the 20th European Signal Processing Conference*, Bucharest, Romania, 27–31 August 2012: 1369–1373. Washington, DC: IEEE.

Sidiropoulos, N. D., T. N. Davidson, and Z.-Q. Luo. 2006. "Transmit Beamforming for Physical-layer Multicasting." *IEEE Transactions on Signal Processing* 54 (6): 2239–2251. doi:10.1109/TSP.2006.872578.

Sturm, J. F. 1999. "Using SeDuMi 1.02, a MATLAB Toolbox for Optimization over Symmetric Cones." *Optimization Methods and Software* 11 (1–4): 625–653. doi:10.1080/10556789908805766.

Sun, Y., and K. J. R. Liu. 2004. "Transmit Diversity Techniques for Multicasting over Wireless Networks." In *Proceedings of IEEE Wireless Communications and Networking Conference*, Atlanta, GA, 21–25 March 2004: 593–598, Vol. 1. Washington, DC: IEEE.

Wen, X., K. L. Law, S. J. Alabed, and M. Pesavento. 2012. "Rank-Two Beamforming for Single-group Multicasting Networks Using OSTBC." In *Proceedings of IEEE Sensor Array and Multichannel Signal Processing Workshop*, Hoboken, NJ, 17–20 June 2012: 69–72. Washington, DC: IEEE.

Wu, S. X., A. M.-C. So, and W.-K. Ma. 2012. "Rank-Two Transmit Beamformed Alamouti Space-time Coding for Physical-layer Multicasting." In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, 25–30 March 2012: 2793–2796. Washington, DC: IEEE.

Wu, S. X., W.-K. Ma, and A. M.-C. So. 2013. "Physical-layer Multicasting by Stochastic Transmit Beamforming and Alamouti Space-time Coding." *IEEE Transactions on Signal Processing* 61 (17): 4230–4245. doi:10.1109/TSP.2013.2263500.