A LOW COMPLEXITY OPTIMIZATION ALGORITHM FOR ZERO-FORCING PRECODING UNDER PER-ANTENNA POWER CONSTRAINTS

B. Li, H. H. Dam, K. L. Teo*

Curtin University
GPO Box U1987, Perth, WA 6845, Australia

A. Cantoni

The University of Western Australia
Crawley, WA, 6009, Australia

ABSTRACT

Zero-forcing beamforming (ZFBF) is a popular pre-coding scheme for MIMO systems. Most of the studies in the literature are under total power constraints. However, the per-antenna power constraints (PAPC) are more realistic. The state-of-the-art method is interior point method which is expensive to realize in practice due to the high computational complexity. Hence, a low complexity zero-forcing precoding scheme under the per-antenna power constraints is proposed in this paper. This is achieved by introducing a regularized dual method. Simulations are carried out to show the effectiveness of the proposed method, which has a low computational complexity. In addition, the algorithm can be implemented in parallel to further reduce the computational complexity.

***Index Terms*—** MIMO, zero-forcing beamforming (ZFBF), per-antenna power constraint (PAPC), dual method, parallel computation

1. INTRODUCTION

Transmitter design for the MU-MIMO systems has been studied intensively in the literature (see, for example, [1]-[3]). The dirty paper coding (DPC) [1] is known as the capacity-achieving scheme. However, it is difficult to be implemented in practical systems due to high computational complexity. Zero-forcing beamforming (ZFBF) is a popular linear precoding method because it provides a good trade-off between the complexity and the performance. Traditionally, ZFBF is studied under the total power constraint. However, for real world applications, each antenna of the transmitter has its own amplifier. Thus, the per-antenna power constraints (PAPC) (see, for example, [4, 5, 6, 7, 8, 9, 10]) are imposed.

ZFBF under the PAPC is a nontrivial problem. Although this problem can be solved by the interior point method [4, 5, 7, 8, 9], it is expensive to apply in practice. This is due to the fact that complex computations are involved, such as a solving nonlinear equation to obtain the Newton step as it is shown in [11]. Thus, it is expensive to realize this method with hardware in practice. For this, we focus on the development of a low complexity algorithm. A new regularized dual method is proposed in this paper. The idea is to solve the original problem by solving its dual problem. This is achieved by introducing the Lagrangian of the primal problem. However, the resulted dual problem is non-differentiable. For this, a Tikhonov regularization is applied to smooth the dual objective function by appending two prox-functions. The regularized dual objective becomes differentiable. Moreover, the optimal solution of the Lagrangian can be written analytically in a closed form. The convergence of the regularization is proved and the effective of regularization is analyzed. The computational complexity analysis shows that the complexity of the proposed method is much lower than that of state-of-the-art method. In addition, this algorithm can be implemented in a parallel manner. Simulations are carried out to show the effectiveness and efficiency of the proposed method.

2. PROBLEM FORMULATION

Consider the standard MISO multiuser broadcast channel

$$y_m = \mathbf{h}_m^H \mathbf{x} + n_m,\ m = 1, 2, \ldots, M, \tag{1}$$

where y_m is the received signal of the m^{th} user, $\mathbf{h}_m$ is the channel vector of length N of the m^{th} user, $\mathbf{x}$ is the transmitted vector of length N, and n_m is the complex Gaussian noise with mean 0 and variance σ^2. It can be written in a compact form as given below:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2}$$

where $\mathbf{y} = [y_1, y_2, \ldots, y_M]^\top$, $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_M]^H$, $\mathbf{n} = [n_1, n_2, \ldots, n_M]^\top$, $(\cdot)^\top$ denotes the transpose, and $(\cdot)^H$ denotes the conjugate transpose.

Here, the linear Zero-Forcing pre-coding transmitter is applied, i.e., $\mathbf{x} = \mathbf{W}\mathbf{s}$, $\mathbf{H}\mathbf{W} = \sqrt{\mathbf{\Lambda}}$ where $\mathbf{s}$ is the information vector of length M such that $E\left\{\mathbf{s}\mathbf{s}^H\right\} = \mathbf{I}$, $\mathbf{I}$ denotes the identity matrix of appropriate dimension, $\mathbf{W}$ is an $N \times M$ complex matrix, and $\mathbf{\Lambda}$ denotes a real and positive diagonal matrix.

*The first author performed the work while at The University of Western Australia

The minimum information rate is taken as the performance measure, which is denoted by $r(m)$ for each user and is given by

$$\min_m r(m) = \min_m \log_2 (1 + SINR(m)),\ m = 1, 2, \ldots, M, \quad (3)$$

where $SINR(m)$ is the signal-to-interference-plus-noise ratio (SINR) for each user, which is given by $SINR(m) = \left|(\mathbf{HW})_{m,m}\right|^2 \Big/ \sigma^2$, $m = 1, 2, \ldots, M$, as the mutual interferences are 'zero-forced'. To limit the power on the amplifier of each antenna, the per-antenna power constraints are imposed as follows:

$$\sum_{m=1}^{M} |\mathbf{e}_n^T \mathbf{w}_m|^2 \le P,\ n = 1, 2, \ldots, N, \quad (4)$$

where $\mathbf{w}_m$ is the mth column vector of $\mathbf{W}$, $\mathbf{e}_n$ is a vector of length N with a 1 in the nth element while 0 in the other elements, and P is the maximum allowable power on each antenna.

The problem under consideration may now be formally stated below:

Problem 2.1

$$\begin{array}{lll} \max\limits_{\mathbf{w}_m, r_0} & r_0 & m = 1, 2, \ldots, M \\ s.\ t. & \log_2 \left(1 + \frac{|\mathbf{h}_m^H \mathbf{w}_m|^2}{\sigma^2}\right) \ge r_0, & m = 1, 2, \ldots, M, \\ & \sum\limits_{m=1}^{M} |\mathbf{e}_n^\top \mathbf{w}_m|^2 \le P, & n = 1, 2, \ldots, N \\ & \mathbf{h}_j^H \mathbf{w}_m = 0, & \forall j \ne m, \\ & & 1 \le j, m \le M. \end{array}$$

For convenience, we transform Problem 2.1 into a real form. From [5], it follows from letting

$$\mathbf{x} = [\mathbf{w}_{1\mathrm{Re}}{}^\top\ \mathbf{w}_{1\mathrm{Im}}{}^\top\ \ldots\ \mathbf{w}_{M\mathrm{Re}}{}^\top\ \mathbf{w}_{M\mathrm{Im}}{}^\top]^\top,$$

where $\mathbf{x} \in \mathbb{R}^{2NM}$, that Problem 2.1 can be written as the following optimization problem.

Problem 2.2

$$\min_{\mathbf{x}, t} \quad -t \quad (5)$$

$$s.\ t. \quad \mathbf{H}_1 \mathbf{x} \le -t\mathbf{1}, \quad (6)$$

$$\mathbf{x}^\top \mathbf{A}_n \mathbf{x} \le P, \quad n = 1, 2, \ldots, N, \quad (7)$$

$$\mathbf{H}_2 \mathbf{x} = \mathbf{0}, \quad (8)$$

where $\mathbf{1}$ is a vector of ones with appropriate dimension, $\mathbf{A}_n = \mathbf{diag}\{\mathbf{B}_n\ \mathbf{B}_n\ \cdots\ \mathbf{B}_n\} \in \mathbb{R}^{2NM \times 2NM}$, $\mathbf{B}_n \in \mathbb{R}^{2N \times 2N}$ is a diagonal matrix with 1 appearing in the (n, n)th and $(n + N, n + N)$th positions and 0 elsewhere, $\mathbf{H}_1 \in \mathbb{R}^{M \times 2NM}$, and $\mathbf{H}_2 \in \mathbb{R}^{(2M(M-1)+M) \times 2NM}$. For the structures of $\mathbf{H}_1$ and $\mathbf{H}_2$, we refer the readers to [7].

3. A REGULARIZED DUAL METHOD

The idea of the proposed method is to solve Problem 2.2 via solving its dual problem. Towards this goal, we introduce the Lagrangian of Problem 2.2

$$L(t, \mathbf{x}, \lambda, \mathbf{v}, \mu) = -t + \lambda^\top (\mathbf{H}_1 \mathbf{x} + t\mathbf{1}) + \mathbf{v}^\top \mathbf{H}_2 \mathbf{x} + \sum_{n=1}^{N} \mu_n \left(\mathbf{x}^\top \mathbf{A}_n \mathbf{x} - P\right). \quad (9)$$

Let $d(\lambda, \mathbf{v}, \mu) = \min_{\mathbf{x}, t} L(t, \mathbf{x}, \lambda, \mathbf{v}, \mu)$ be the dual function. Then the dual problem of Problem 2.2 can be written as

Problem 3.1

$$\max_{\lambda, \mathbf{v}, \mu} \quad d(\lambda, \mathbf{v}, \mu) \quad (10)$$

$$s.\ t. \quad \lambda \succeq \mathbf{0} \quad (11)$$

$$\mu \succeq \mathbf{0}, \quad (12)$$

where $\mathbf{a} \succeq \mathbf{0}$ means each element of $\mathbf{a}$ is greater than or equal to 0.

Since Problem 2.2 is convex and the Slater's condition is satisfied, the strong duality holds [11]. Thus, Problem 2.2 can be solved through solving Problem 3.1.

3.1. Regularization

Note that the dual function is not differentiable. We apply Tikhonov regularization to smooth the dual function as that in [12]. To begin, we denote $d_t(t) = \rho t^2$ and $d_{\mathbf{x}}(\mathbf{x}) = \rho\|\mathbf{x}\|^2$ as two prox-functions, where $\rho = M\alpha/((N^2 + MN)P)$ is a smoothing parameter and $\alpha > 0$. By appending $d_t(t)$ and $d_{\mathbf{x}}(\mathbf{x})$ into Lagrangian (9), we obtain

$$L_\rho(t, \mathbf{x}, \lambda, \mathbf{v}, \mu) = L(t, \mathbf{x}, \lambda, \mathbf{v}, \mu) + d_t(t) + d_{\mathbf{x}}(\mathbf{x}) \quad (13)$$

The optimal solution of $\min\limits_{t, \mathbf{x}} L_\rho(t, \mathbf{x}, \lambda, \mathbf{v}, \mu)$ can be written in a closed form as

$$t^* = \frac{1}{2\rho}\left(1 - \sum_{m=1}^{M} \lambda_m\right) \quad (14)$$

and

$$\mathbf{x}^* = -\frac{1}{2}\mathbf{S}(\mu, \rho)\left(\mathbf{H}_1^\top \lambda + \mathbf{H}_2^\top \mathbf{v}\right) \quad (15)$$

where $\mathbf{S}(\mu, \rho) = \mathbf{diag}\,(\mathbf{C}\ \mathbf{C}\ \cdots\ \mathbf{C}) \in \mathbb{R}^{2MN \times 2MN}$ and

$$\mathbf{C} = \mathbf{diag}\left(\frac{1}{\rho + \mu_1}\ \frac{1}{\rho + \mu_2}\ \cdots\ \frac{1}{\rho + \mu_N}\right) \in \mathbb{R}^{N \times N}.$$

From (13), (14) and (15), it follows that

$$\begin{aligned} d_\rho(\lambda, \mathbf{v}, \mu) &= d(\lambda, \mathbf{v}, \mu) + d_t(t) + d_{\mathbf{x}}(\mathbf{x}) \\ &= -\frac{1}{4}\left(\mathbf{H}_1^\top \lambda + \mathbf{H}_2^\top \mathbf{v}\right)^\top \mathbf{S}(\mu, \rho)\left(\mathbf{H}_1^\top \lambda + \mathbf{H}_2^\top \mathbf{v}\right) \\ &\quad -\frac{1}{4\rho^2}\left(\sum_{m=1}^{M} \lambda_m - 1\right)^2 - P\sum_{n=1}^{N} \mu_n \end{aligned} \quad (16)$$

We refer the regularized dual problem as Problem 3.1(ρ) by replacing (10) in Problem 3.1 with (16). Then, we give the error bound that introduced by the regularization with the following theorem.

Theorem 3.1 *Suppose the Slater conditions hold and let* $\left(t_\rho^*, \mathbf{x}_\rho^*\right)$ *and* $(t^*, \mathbf{x}^*)$ *be optimal solutions of Problem 3.1(ρ) and Problem 2.2, respectively. Then, for any ρ,*

$$t^* - t_\rho^* \leq \alpha. \tag{17}$$

Proof Under the Slater condition, both Problem 2.2 and Problem 3.1(ρ) have solutions. We can easily prove that

$$t^* - t_\rho^* < \rho t^{*2} + \rho \|\mathbf{x}^*\|^2 \tag{18}$$

From (8), we have $\sum_{i=1}^{N} \mathbf{x}^\top \mathbf{A}_i \mathbf{x} \leq NP$. Note that $\sum_{i=1}^{N} \mathbf{A}_i = I_N$, where I_N is an $N \times N$ identity matrix. Then, it follows that

$$\|\mathbf{x}\|^2 \leq NP. \tag{19}$$

By adding all the rows of $\mathbf{H}_1$ and all the components of $-t\mathbf{1}$ in (6), we have $\mathbf{L}^\top \mathbf{x} \geq Mt$, where

$$\mathbf{L} = \left[\begin{array}{cccccc} \mathbf{h}_{1\mathrm{Re}}^\top & \mathbf{h}_{1\mathrm{Im}}^\top & \cdots & \cdots & \mathbf{h}_{M\mathrm{Re}}^\top & \mathbf{h}_{M\mathrm{Im}}^\top \end{array}\right].$$

Since $t^* > 0$, then it follows that

$$t^2 \leq \|\mathbf{L}\|^2 \|\mathbf{x}\|^2 / M^2. \tag{20}$$

By considering the definition of Euclidian norm, we know that $\|\mathbf{L}\|^2 = MN$. Thus, from (19) and (20), we obtain

$$t^2 \leq N^2 P / M. \tag{21}$$

Combining (18), (19) and (21) and knowing that $\rho = M\alpha/((N^2 + MN)P)$, we obtain the desired relation. This completes the proof.

Remark 3.1 *(19) and (21) provide a 'good' bound for this problem. In fact, this is achieved by taking the conjugate of channel as the weights of the beamformer under the total power constraint. This is known as matched filter in signal processing, which provides the best signal to noise ratio. We can refer to this as an 'idealized' scenario. This 'idealized' bound may not be achieved in this problem since the feasible set bounded by PAPC is a subset of that bounded by the total power constraint.*

From the definition of ρ, we know that ρ only depends on α for a specified scenario, where N, M and P are fixed. Thus, we can give a rule of choosing ρ according to Theorem 3.1. For this, we define $\alpha = \beta N \sqrt{P/M}$, where β is called error bound parameter. In fact, β is the portion of the bound of the regularization error to the performance of the 'idealized' scenario according to (21) and Remark 3.1. In addition, the definition of α shows that the error bound $t^* - t_\rho^*$ depends on N and square root of $1/M$ linearly.

3.2. Algorithm

In the kth iteration, the primal variables $\mathbf{x}^{(k)}$ and $t^{(k)}$ can be updated in parallel with (14)-(15). The dual variables λ, $\mathbf{v}$ and μ can be updated by

$$\left\{ \begin{array}{lcl} \lambda^{(k+1)} & = & \max\left\{0, \lambda^{(k)} + \frac{1}{L_\lambda}\frac{\partial d_\rho}{\partial \lambda}\right\} \\ \mathbf{v}^{(k+1)} & = & \mathbf{v}^{(k)} + \frac{1}{L_\mathbf{v}}\frac{\partial d_\rho}{\partial \mathbf{v}} \\ \mu^{(k+1)} & = & \max\left\{0, \mu^{(k)} + \frac{1}{L_\mu}\frac{\partial d_\rho}{\partial \mu}\right\} \end{array} \right. \tag{22}$$

where $\mathbf{g}(\mathbf{x}) = [g_1(\mathbf{x})\ g_2(\mathbf{x})\ \cdots\ g_N(\mathbf{x})]^\top$ and $g_n(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}_n \mathbf{x} - P$, $n = 1, 2, \ldots, N$. L_λ, $L_\mathbf{v}$ and L_μ are Lipschitz constants and they can be obtained as in [13]. The gradients of (16) with respect to λ, $\mathbf{v}$ and μ are:

$$\frac{\partial d_\rho}{\partial \lambda} = \mathbf{H}_1 \mathbf{x} + t\mathbf{1}, \quad \frac{\partial d_\rho}{\partial \mathbf{v}} = \mathbf{H}_2 \mathbf{x},$$

$$\frac{\partial d_\rho}{\partial \mu_n} = \mathbf{x}^\top \mathbf{A}_n \mathbf{x} - P, \quad n = 1, 2, \ldots, N.$$

Particularly, according to (15) and the structure of $\mathbf{S}$, the primal variables $\mathbf{x}$ can be updated in parallel as in (23), where $\mathbf{v} = \left[\mathbf{v}_1^\top \mathbf{v}_2^\top \cdots \mathbf{v}_{2M+1}^\top\right]^\top \in \mathbb{R}^{2M(M-1)+M}$ is the dual variable and v_j^i is the ith element of $\mathbf{v}_j$. The stopping criterion is the duality gap, the difference between the function value of the primal problem and that of the dual problem, meets a tolerance.

3.3. Convergence Analysis

We present the convergence results in the following theorems.

Theorem 3.2 *Let* $\left(t_{\rho_k}^*, \mathbf{x}_{\rho_k}^*\right)$ *be an optimal solution of Problem 3.1(ρ_k), and* $\lim_{k \to +\infty} \rho_k = 0$. *Then, there exits a subsequence* $\left\{\left(t_{\rho_{\bar{k}}}^*, \mathbf{x}_{\rho_{\bar{k}}}^*\right)\right\}$ *of the sequence* $\left\{\left(t_{\rho_k}^*, \mathbf{x}_{\rho_k}^*\right)\right\}_{k=1}^{+\infty}$ *such that it converges to an optimal solution of Problem 2.2 as* $k \to +\infty$.

Proof Since $\left(t_{\rho_k}^*, \mathbf{x}_{\rho_k}^*\right)$ is bounded for all ρ_k, there exits a subsequence $\left\{\left(t_{\rho_{\bar{k}}}^*, \mathbf{x}_{\rho_{\bar{k}}}^*\right)\right\}$ such that $\left(t_{\rho_{\bar{k}}}^*, \mathbf{x}_{\rho_{\bar{k}}}^*\right) \to (\bar{t}, \bar{\mathbf{x}})$ as $k \to +\infty$. We will show that $(\bar{t}, \bar{\mathbf{x}})$ is an optimal solution of Problem 2.2. We can prove that $(\bar{t}, \bar{\mathbf{x}})$ is an optimal solution of Problem 2.2 with the fact that the limit of the objective function of the corresponding primal problem of Problem 3.1($\rho_{\bar{k}}$) is $-\bar{t}$. This completes the proof.

From Proposition 1.3.3 in [13], we can obtain that the convergence rate of the algorithm is $o\left(\frac{1}{k}\right)$.

4. NUMERICAL RESULTS AND COMPUTATIONAL COMPLEXITY ANALYSIS

The base station array considered in the numerical studies is a uniform planar circular array. It consists of N isotropic elements and the inter-element spacing is equal to 0.5λ. There

$$\begin{cases} \mathbf{x}_{2i-1}^k &= -\frac{1}{2}\mathbf{C}\left(-\lambda_i^k \mathbf{h}_{i\mathrm{Re}} + \sum\limits_{j=1,j\neq i}^{M} \left(v_j^i\right)^k \mathbf{h}_{j\mathrm{Re}} - \sum\limits_{j=M+1,j\neq M-i}^{2M} \left(v_j^i\right)^k \mathbf{h}_{j-M\mathrm{Im}} - \left(v_{2M+1}^i\right)^k \mathbf{h}_{i\mathrm{Im}}\right) \\ \mathbf{x}_{2i}^k &= -\frac{1}{2}\mathbf{C}\left(-\lambda_i^k \mathbf{h}_{i\mathrm{Im}} + \sum\limits_{j=1,j\neq i}^{M} \left(v_j^i\right)^k \mathbf{h}_{j\mathrm{Im}} + \sum\limits_{j=M+1,j\neq M-i}^{2M} \left(v_j^i\right)^k \mathbf{h}_{j-M\mathrm{Re}} + \left(v_{2M+1}^i\right)^k \mathbf{h}_{i\mathrm{Re}}\right) \end{cases} \tag{23}$$

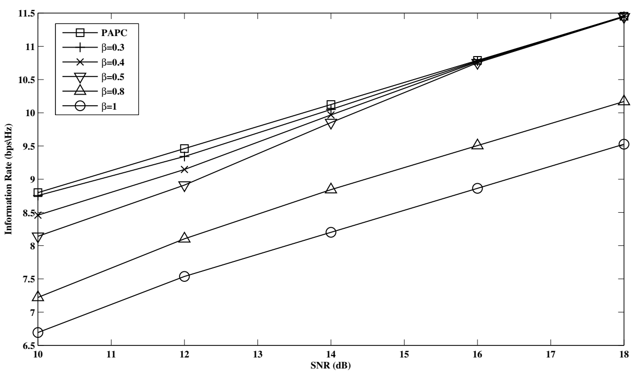

Fig. 1. ZFBF under PAPC as a function of N with different ρ.

Table 1. Computational Complexity Comparison

	[7, 8, 9]	Proposed
Iteration Complexity	$O\left(M^3N^3\right)$	$O\left(M^2N^2\right)$
Convergence Rate	$O\left(\ln(1/\epsilon)\right)$	$O\left(1/\epsilon\right)$

Table 2. Iteration Complexity Comparison

N,M	[7, 8, 9]	Proposed
$N = 12, M = 4$	1.11×10^5	2.30×10^3
$N = 12, M = 6$	3.73×10^5	5.18×10^3
$N = 12, M = 12$	2.99×10^6	2.07×10^4
$N = 24, M = 4$	8.85×10^5	9.22×10^3
$N = 24, M = 6$	2.99×10^6	2.07×10^4
$N = 24, M = 12$	2.39×10^7	8.29×10^4
$N = 24, M = 18$	8.06×10^7	1.87×10^5

are M users around the base station. Set $\sigma^2 = 0.001$. The simulation is implemented in the Matlab environment. We implement Algorithm 1 within a deterministic line-of-sight channel tested by The Commonwealth Scientific and Industrial Research Organization (CSIRO) in rural Australia [15]. We plot the performances with different values of the error bound parameter β as a function of SNR in Fig. 1. As it is shown in Fig. 1, the performance converges to the optimum, which is denoted as PAPC, as β decreases. We can see that the performance is already very closed to the optimal performance when $\beta = 0.3$. Recalling Theorem 3.1 and the definition of α, we know that, in this case, the bound of the regularization error is 30% of the 'idealized' scenario. We cannot tight the bound error further since the 'idealized' scenario only serves as a reference and could not be achieved as explained in Remark 3.1.

Then, we compare the computational complexity per iteration of the proposed method with that of state-of-the-art interior point method in [7, 8, 9]. The details of the comparison are shown in Table 1. From the comparison, we can see that there is a trade off between the complexity per iteration and the convergence rate. More specifically, the interior point method gains convergence rate at the expense of a much higher cost on the computational complexity per iteration, resulting in much higher expense on the hardware implementation. Moreover, it cannot be implemented in a parallel manner. Hence, the proposed algorithm is much more attractive from the practical point of view. Then, we compare the complexity per iteration between the interior point method and proposed method with different N and M. The results are shown in Table 2. From Table 2, we can see that the computational complexity per iteration of the proposed method is much lower than that of the interior point method. Then, we shall study how much computational complexity we can reduce by implementing the parallel computation. By comparing (15) with (23), we can see that the primal update in (15) can be decomposed into $2M$ streams. We compare the occupational complexity of the primal updates between (15) and (23) in Table 3. As it is shown in Table 3, we can further reduce the complexity by exploring the parallelism.

5. CONCLUSIONS

A low complexity beamformer design is proposed by means of a regularized dual method. An $O\left(M^2N^2\right)$ iteration complexity is achieved by applying the proposed mehtod while for the state-of-the-art method it is $O\left(M^3N^3\right)$. Furthermore, the proposed method can further reduce the complexity to $O\left(N^2\right)$ by exploring the parallelism. Thus, the proposed algorithm is more attractive in practice. The smoothing parameter ρ can be chosen by setting the error bound parameter β. It shows that when β decreases the performance goes to the optimal performance. However, as it is shown, β cannot be set too small as well.

Table 3. Complexity Reduction by Exploring Parallelism

	(15)	(23)
Computational Complexity	$O\left(M^2N^2\right)$	$O\left(N^2\right)$

6. REFERENCES

[1] G. Caire and S. Shamai (Shitz), "On the achievable throughput of multiatenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory.*, vol. 49, no. 7, pp. 1691-1706, Jun. 2003.

[2] P. S. Udupa and J. S. Lehnert, "Optimizing zero-forcing precoders for MIMO broadcast systems," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1516-1524, Aug. 2007.

[3] A. Wiesel, Y. C. Eldar, and S. Shamai (Shitz), "Linear precoding via conic optimizaiton for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161-176, Jan. 2006.

[4] A. Wiesel, Y. C. Eldar, and S. Shamai (Shitz), "Zero-forcing precoding and generalized inverses," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4409-4418, Sept. 2008.

[5] K. Karakayali, R. Yates, G. Foschini, and R. Valenzuela, "Optimal zero-forcing beamforming with per-antenna power constraints," in *Proc. IEEE Int. Symp. Inf. Theory.*, 2007, pp. 101-105.

[6] S. R. Lee, J. S. Kim, S. H. Moon, H. B. Kong, and I. Lee, "Zero-forcing beamforming in multiuser MISO downlink systems under per-antenna power constraint and equal-rate metric," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 20-24, Jan. 2013.

[7] B. Li, H. H. Dam, A. Cantoni and K. L. Teo, "A primal-dual interior point method for optimal zero-forcing beamformer design under per-antenna power constraints," *Optim. Lett.*, vol. 8 no. 6, pp. 1829-1843, 2014.

[8] H. H. Dam and A. Cantoni, "Interior point method for optimum zero-forcing beamforming with per-antenna power constraints and optimal step size," *Signal Process.*, vol. 106, pp. 10-14, 2015.

[9] B. Li, H. H. Dam, A. Cantoni and K. L. Teo, "Some interesting properties for zero-forcing beamformer design under per-antenna power constraints in rural areas," *J. Glob. Optim.*, DOI 10.1007/s10898-014-0237-4.

[10] B. Li, H. H. Dam, A. Cantoni and K. L. Teo, "A first-order optimal zero-forcing beamformer design for multiuser MIMO systems via a regularized dual accelerated gradient method," *IEEE Commun. Lett.*, vol. 19 no. 2, pp. 195-198, Feb. 2015.

[11] S. Boyd and L. Vandenberghe, *Covex optimization*. Cambrige, UK: Cambrige University Press, 2004.

[12] J. Koshal, A. Nedic, and U. V. Shanbhag, "Multiuser optimization: distributied algorithms and error analysis," *SIAM J. Optim.*, vol. 21, no. 3, pp. 10461081, 2011.

[13] D. P. Bertsekas, *Nonlinear Programming*. Belmont, Massachusetts, USA: Athena Scientific, 1999.

[14] S. J. Benson, Y. Ye, and X. Zhang, "Solving large-scale sparese semidefinite programs for combinational optimization," *SIAM J. Optim.*, vol. 10, no. 2, pp. 443-461, 2000.

[15] H. Suzuki, D. Robertson, N. L. Ratnayake, and K. Ziri-Castro, "Prediction and measurement of multiuser MIMO-OFDM channel in rural Australia" in *IEEE 75th Vehicular Technology Conference*, 2012, pp. 1 - 5.