A Dual Decomposition Approach to the Sum Power Gaussian Vector Multiple Access Channel Sum Capacity Problem

Wei Yu
Edward S. Rogers Sr. Department of Electrical and Computer Engineering
University of Toronto
10 King's College Road, Toronto, Ontario M5S 3G4, Canada
Email: weiyu@comm.utoronto.ca

Abstract — **The Gaussian vector multiple access channel with a sum-power constraint across all users, rather than the usual individual power constraint on each user, has recently been shown to be the dual of a Gaussian vector broadcast channel [1] [2]. Further, a numerical algorithm for the sum capacity under the sum power constraint has been proposed in [3]. This paper proposes a different algorithm for this problem based on a dual decomposition approach. The proposed algorithm works in the Lagrangian dual domain; it is based on a modified iterative water-filling algorithm for the multiple access channel; and it is guaranteed to converge to the sum capacity in all cases. This spectrum optimization problem for the sum-power multiple access channel is also applicable to the optimal power allocation problem for an OFDM system with correlated noise.**

I. Introduction

Consider a Gaussian multiple access channel with vector inputs and a vector output:

$$\mathbf{Y} = \sum_{k=1}^{K} H_k \mathbf{X_k} + Z, \tag{1}$$

where $\mathbf{X_k}$'s are $n \times 1$ vectors, H_k are $m \times n$ channel matrices, $\mathbf{Y}$ is a $m \times 1$ vector, and $\mathbf{Z}$ is a $m \times 1$ additive Gaussian random vector. The inputs $\mathbf{X_k}$ are assumed to be independent. In a conventional multiple-access channel, a separate input constraint applies to each input $\mathbf{X_k}$, i.e.

$$\mathbf{E X_k X_k^T} \leq P_k, \quad k = 1, \cdots, K. \tag{2}$$

This paper, however, deals with a different situation. While retaining the assumption that $\mathbf{X_k}$ are independent, this paper considers a Gaussian multiple access channel with a sum-power constraint applied to all $\mathbf{X_k}$'s at the same time:

$$\sum_{k=1}^{K} \mathbf{E X_k X_k^T} \leq P. \tag{3}$$

The objective is to find an efficient numerical algorithm to evaluate the sum capacity of the multiple access channel under the sum power constraint.

The capacity region of a multiple access channel under a fixed input distribution $p(\mathbf{x_1}) \cdots p(\mathbf{x_K})$ is a well-known pentagon region. For example, for a two-user multiple access channel, the capacity region can be expressed as follows:

$$\begin{aligned} R_1 &\leq I(\mathbf{X_1}; \mathbf{Y} | \mathbf{X_2}) \\ R_2 &\leq I(\mathbf{X_2}; \mathbf{Y} | \mathbf{X_1}) \\ R_1 + R_2 &\leq I(\mathbf{X_1}, \mathbf{X_2}; \mathbf{Y}). \end{aligned} \tag{4}$$

Thus, the sum capacity of the multiple access channel is the solution to the following mutual information maximization problem:

$$C = \max_{\mathrm{co}\{p(\mathbf{x_1}) \cdots p(\mathbf{x_K})\}} I(\mathbf{X_1} \cdots \mathbf{X_K}; \mathbf{Y}), \tag{5}$$

where the maximization is over the convex hull of all input distributions that satisfy the input constraint. This maximization problem is not easy to solve in general, because the input distribution must be a convex combination of independent distributions, and the constraint that the inputs must be independent is not a convex constraint.

However, for Gaussian vector multiple access channels, the mutual information maximization problem can be cast as a convex optimization problem. This is because for Gaussian channels, Gaussian inputs are optimal and the optimization over the input distributions becomes an optimization over the covariance matrices of $\mathbf{X_k}$. More precisely, let $S_k = \mathbf{E X_k X_k^T}$ be the input covariance matrix for the user k. The mutual information maximization problem under separate power constraints (2) becomes:

$$\begin{array}{ll} \text{maximize} & \dfrac{1}{2} \log \dfrac{|\sum_{k=1}^{K} H_k S_k H_k^T + S_z|}{|S_z|} \\ \text{subject to} & \mathrm{tr}(S_k) \leq P_k, \quad k = 1, \cdots, K, \\ & S_k \geq 0, \qquad k = 1, \cdots, K. \end{array} \tag{6}$$

Similarly, the maximization problem under the sum power constraint (3) becomes:

$$\begin{array}{ll} \text{maximize} & \dfrac{1}{2} \log \dfrac{|\sum_{k=1}^{K} H_k S_k H_k^T + S_z|}{|S_z|} \\ \text{subject to} & \displaystyle\sum_{k=1}^{K} \mathrm{tr}(S_k) \leq P \\ & S_k \geq 0, \qquad k = 1, \cdots, K. \end{array} \tag{7}$$

where "tr" denotes the matrix trace operation, and S_z is the covariance matrix of $\mathbf{Z}$. Because $\log|\cdot|$ is a concave function in the set of positive semi-definite matrices, and constraints are linear, both problems belong to the class of convex optimization problems for which numerical solutions are in principle easy to obtain.

In fact, with a separate power constraint on each $\mathbf{X_k}$, the sum capacity problem (6) has an efficient solution called "iterative water-filling". The key observation is that the maximization problem can be solved by updating one S_k at a time while keeping all others fixed. Because the constraints are separable, each update can be done independently, and the iterative process converges to the sum capacity. However, this iterative procedure is not directly applicable to the sum-power problem (3) in which a coupled power constraint applies to all $\mathbf{X_k}$'s at the same time. The main idea of this paper is that via a technique called dual decomposition, the sum-power constraint may be de-coupled.

The numerical solution for the sum-power constrained multiple access channel was first considered in [3], in which an algorithm based on iterative water-filling was proposed for the sum-power problem. However, [3] did not fully establish the convergence of the proposed algorithm. This paper is motivated by the results in [3], but it uses a dual decomposition approach in network optimization, most recently used in the work of Xiao and Boyd [4]. The proposed algorithm is based on a modified iterative water-filling. It is efficient, and it converges in all cases.

The rest of the paper is organized as follows. The motivation for solving the sum-power sum-capacity multiple access channel is presented in section 2. Section 3 proposes a numerical solution to the sum-capacity problem based on a dual decomposition method, and proves its convergence. Section 4 presents simulation results. Section 5 contains concluding remarks.

II. Motivation

The sum-power multiple access channel capacity problem arises in two different applications. The primary motivation for studying the sum-power constraint is that a sum-power constrained Gaussian vector multiple access channel has the same sum capacity as a dual Gaussian vector broadcast channel [1] [2]. The broadcast channel problem is of great interest as it models a downlink transmission environment in a cellular wireless system, and the sum-capacity problem has been unsolved until very recently. The sum-power constrained multiple access channel also appears in a spectrum optimization problem for OFDM systems with correlated noise. These two motivating examples are presented in more details in the following sections.

A. Gaussian Vector Broadcast Channel Sum Capacity

Consider a Gaussian vector broadcast channel:

$$\mathbf{Y} = H\mathbf{X} + \mathbf{Z}, \tag{8}$$

where $\mathbf{X}$ is a vector-valued transmit signal, the vector $\mathbf{Y} = [\mathbf{Y_1} \cdots \mathbf{Y_K}]^T$ represents receive signals for users $1 \cdots K$, and $\mathbf{Z} = [\mathbf{Z_1} \cdots \mathbf{Z_K}]^T$ is the additive i.i.d. Gaussian noise with unit variance. Independent information is to be transmitted from $\mathbf{X}$ to each component of $\mathbf{Y}$.

The Gaussian vector broadcast channel is not degraded, and its capacity region is still an unsolved problem. Recently, several authors have tackled the problem using a precoding approach and successfully solved the Gaussian vector broadcast channel sum-capacity problem [5] [6] [2] [1]. In [6], the author showed that the optimal precoder is a decision-feedback equalizer, and the sum capacity is equal to the following minimax mutual information:

$$\max_{S_x} \min_{S_z} I(\mathbf{X}; \mathbf{Y}), \tag{9}$$

where the maximization is over the input constraint, and the minimization is over all possible noise covariance matrices S_z such that the diagonal terms are fixed, and the off-diagonal terms are allowed to vary. In other words, the minimization is over all joint distributions $\mathbf{Z_1}, \cdots \mathbf{Z_K}$, while keeping the marginal distributions fixed. The above optimization problem is a convex minimax problem, so it is in principle tractable.

Using a completely different technique, the authors of [2] and [1] showed that the sum capacity of a Gaussian vector broadcast channel is equal to the sum capacity of a dual Gaussian vector multiple access channel with $\mathbf{X'} = \mathbf{Y}$ as the transmitters, $H' = H^T$ as the channel matrix, $\mathbf{Y'} = \mathbf{X}$ as the receiver, and $\mathbf{Z'}$ as the unit variance additive white Gaussian noise:

$$\mathbf{Y'} = H'\mathbf{X'} + \mathbf{Z'}. \tag{10}$$

In addition, the dual vector multiple access channel has a *sum* power constraint across all transmitters rather than individual power constraints. Thus, solving for the sum capacity of a sum-power constrained Gaussian vector multiple access channel is equivalent to solving for the Gaussian vector broadcast channel sum capacity, thus motivating the problem proposed in this paper.

In fact, it can be shown that solving for the sum-power multiple access channel capacity (10) is equivalent to solving the minimax problem (9) [2]. However, such an equivalence can be established only if the input constraint on the original broadcast channel is a linear constraint. For Gaussian broadcast channels with arbitrary convex input constraints, the minimax expression appears to be a more general result. (See [7] for a discussion.)

B. OFDM with Correlated Noise

Consider a Gaussian channel with inter-symbol interference (ISI):

$$y_i = \sum_{l=0}^{\nu} \alpha_l x_{i-l} + n_i, \tag{11}$$

where $(\alpha_0, \cdots, \alpha_\nu)$ is the set of ISI coefficients, and n_i is the additive Gaussian noise. One way to deal with ISI is to consider a block transmission of N symbols, insert ν guard symbols between the blocks, and use orthogonal frequency division multiplex (OFDM) modulation. In this case, the equivalent vector channel

$$\begin{bmatrix} y_N \\ \vdots \\ y_1 \end{bmatrix} = H_C \begin{bmatrix} x_N \\ \vdots \\ x_1 \end{bmatrix} + \begin{bmatrix} n_N \\ \vdots \\ n_1 \end{bmatrix} \tag{12}$$

becomes circulant, and a pair of FFT and IFFT matrices can be used to diagonalize the channel. With an IFFT matrix as the modulator, and an FFT matrix as the demodulator, the vector channel is transformed into a set of parallel independent scalar channels, onto which independent signaling may be performed.

The traditional model for OFDM as described above works well when the noise vector is uncorrelated. This is because an uncorrelated noise vector remains uncorrelated after an FFT operation, and the resulting scalar channels will then be independent. In fact, as the block size goes to infinity, OFDM with optimal power and bit allocation achieves the capacity of an ISI channel. However, OFDM modulation is not optimal if the noise vector is correlated. In this case, the resulting parallel channel are not independent, and independent signaling is not optimal. However, in many practical applications, OFDM modulation (and thus independent signaling) is used regardless of whether noise correlation exists. It is not difficult to see that in these circumstances the capacity of the OFDM channel is the solution to the following maximization problem:

$$\begin{array}{ll} \text{maximize} & \frac{1}{2}\log\frac{|DS_xD^T+S_z|}{|S_z|} \\ \text{subject to} & S_x \text{ is diagonal} \\ & \mathrm{tr}(S_x)\le P \\ & S_x \ge 0, \end{array} \tag{13}$$

where D is the diagonal channel matrix after the FFT operation, S_z is the noise covariance matrix which is not necessarily diagonal. The optimization is over all diagonal transmit matrices S_x subject to a power constraint. The above problem is equivalent to the sum-power multiple access channel sum-capacity problem with each of the diagonal entries of S_x as a separate user. This provides another motivation for the proposed problem.

III. Dual Decomposition

Consider the optimization problem (7):

$$\begin{array}{ll} \text{maximize} & \frac{1}{2}\log\frac{|\sum_{k=1}^K H_kS_kH_k^T+S_z|}{|S_z|} \\ \text{subject to} & \sum_{k=1}^K \mathrm{tr}(S_k)\le P \\ & S_k\ge 0, \qquad k=1,\cdots,K, \end{array} \tag{14}$$

where the optimization variables are semi-definite matrices $\{S_k\}$. The optimization variable can also be thought of as $S = diag\{S_1,\cdots,S_K\}$ with the constraint $\mathrm{tr}(S)\le P$. Without the constraint that S must be diagonal, the problem is equivalent to a conventional Gaussian vector channel for which the well-known water-filling solution applies. But water-filling does not necessarily give a diagonal transmit covariance matrix. On the other hand, if S is kept as diagonal, but individual power constraints are applied to each of S_k rather than the sum power constraint, a numerical algorithm called "iterative water-filling" can be used to find the sum capacity efficiently [8]. As mentioned before, the key idea of iterative water-filling is that each S_k may be optimized individually while keeping all other S_k's fixed. The fixed point of the iterative algorithm is the global optimum. The primary difficulty in solving (14) is that while the transmit signals must be independent, the constraint on their covariance matrices $\{S_1,\cdots S_K\}$ is coupled. Recently, Jindal, Jafar, Vishwanath and Goldsmith [3] proposed an algorithm to solve the sum power problem based on an iterative water-filling approach. However, [3] did not establish the convergence of the algorithm in full generality.

The central idea in dual decomposition is to introduce a dual variable to decouple the coupled constraint. This approach was most recently used by Xiao and Boyd [4] in their study of joint optimization of routing and resource allocation in large-scale networks. This paper is inspired by [4]. The dual decomposition method works as follows: first, introduce new variables $\{P_1\cdots P_K\}$, and re-write the optimization problem into the following form in which only a single constraint is coupled:

$$\begin{array}{lll} \text{maximize} & \frac{1}{2}\log\frac{|\sum_{k=1}^K H_kS_kH_k^T+S_z|}{|S_z|} & \\ \text{subject to} & \mathrm{tr}(S_k)\le P_k & k=1,\cdots,K, \\ & S_k\ge 0, & k=1,\cdots,K, \\ & \sum_{k=1}^K P_k\le P. & \end{array} \tag{15}$$

Form the Lagrangian of the optimization problem with respect to the coupled constraint $\sum_{k=1}^K P_k\le P$ only:

$$\begin{aligned} & L(S_1,\cdots,S_K,P_1,\cdots,P_K,\lambda) \\ &= \frac{1}{2}\log\frac{|\sum_{k=1}^K H_kS_kH_k^T+S_z|}{|S_z|} - \lambda\left(\sum_{k=1}^K P_k - P\right). \end{aligned} \tag{16}$$

Let the dual objective be

$$g(\lambda) = \max_{S_1,\cdots,S_K,P_1,\cdots,P_K} L(S_1,\cdots,S_K,P_1,\cdots,P_K,\lambda), \tag{17}$$

where the maximization is under the constraints $\mathrm{tr}(S_k)\le P_k$ and $S_k\ge 0$. Because the original optimization problem is convex, the dual objective reaches a minimum at the optimal value of the primal problem. Thus, the sum-power multiple access channel sum capacity problem reduces to:

$$\begin{array}{ll} \text{minimize} & g(\lambda) \\ \text{subject to} & \lambda\ge 0. \end{array} \tag{18}$$

The key observation is that $g(\lambda)$ is easy to compute, and the above minimization problem can be solved much more efficiently than the original problem. Consider first the evaluation of $g(\lambda)$. By definition, $g(\lambda)$ is the solution to the following optimization problem:

$$\begin{array}{ll} \text{maximize} & \frac{1}{2}\log\frac{|\sum_{k=1}^K H_kS_kH_k^T+S_z|}{|S_z|} \\ & -\lambda\left(\sum_{k=1}^K P_k - P\right) \\ \text{subject to} & \mathrm{tr}(S_k)-P_k\le 0, \quad k=1,\cdots,K \\ & S_k\ge 0, \qquad k=1,\cdots,K \end{array} \tag{19}$$

Notice that the above maximization problem has de-coupled constraints. Therefore, an iterative water-filling like algorithm can be used to solve the problem efficiently. The iterative algorithm works as follows: in each step, maximize the objective over one pair of (S_k,P_k), while keeping all other (S_k,P_k)'s fixed. Since the objective is non-decreasing with each iteration, it must converge to a fixed point. At the fixed point, the set of (S_k,P_k) satisfies the KKT condition of the optimization problem (19). Thus, the fixed point is precisely the optimal solution.[1]

In fact, each step of the iterative algorithm is just a trivial evaluation of water-filling with a fixed water level λ. Without loss of generality, consider the optimization over (S_1,P_1) while

[1] A similar argument has been used in [8].

keeping all other (S_k, P_k) fixed. The Karush-Kuhn-Tucker condition for the optimization problem (19) is just:

$$\frac{1}{2}H_1^T\left(\sum_{k=1}^K H_k S_k H_k^T + S_z\right)^{-1} H_1 = \lambda I + \Phi_k, \tag{20}$$

where Φ_k is the dual variable associated with the constraint $S_k \geq 0$. Notice that this is exactly the water-filling condition for a Gaussian vector channel, except that in this case, λ is fixed. Thus, the usual water-filling procedure only needs to be modified slightly in order to find the optimal (S_k, P_k). The idea is to fill up transmit power to a fixed water level, rather than water-filling with a fixed total power. More precisely, let

$$\left(\sum_{k=2}^K H_k S_k H_k^T + S_z\right) = Q^T \Lambda Q \tag{21}$$

be an eigenvalue decomposition. The maximization problem (19) is equivalent to the maximization of

$$\frac{1}{2}\log\left|\Lambda^{-\frac{1}{2}} Q H_1 S_1 H_1^T Q^T \Lambda^{-\frac{1}{2}} + I\right| - \lambda P_1. \tag{22}$$

Let

$$\Lambda^{-\frac{1}{2}} Q H_1 = U \begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_K \end{bmatrix} V^T \tag{23}$$

be a singular-value decomposition. Then, the optimal S_1 is just:

$$S_1 = V \begin{bmatrix} \left(\frac{1}{2\lambda} - \frac{1}{s_1^2}\right)_+ & & \\ & \ddots & \\ & & \left(\frac{1}{2\lambda} - \frac{1}{s_K^2}\right)_+ \end{bmatrix} V^T. \tag{24}$$

The optimal P_1 is just:

$$P_1 = \sum_{k=1}^K \left(\frac{1}{2\lambda} - \frac{1}{s_K^2}\right)_+. \tag{25}$$

The next step is to use the same procedure to find the optimal (S_2, P_2), while keeping $(S_1, P_1), (S_3, P_3), \cdots, (S_K, P_K)$ fixed. The next step is to update (S_3, P_3) and (S_4, P_4), $\cdots$, then (S_1, P_1), $(S_2, P_2) \cdots$. The iterative procedure is very efficient. It is guaranteed to converge, and it converges to the optimal value of (19) and thus $g(\lambda)$.

It remains to minimize $g(\lambda)$ subject to the constraint $\lambda \geq 0$. $g(\lambda)$ is a concave function. Further, the constraint set is a one-dimensional interval. Thus, a standard search algorithm on λ yields satisfactory results. Unfortunately, $g(\lambda)$ is not necessarily differentiable, so it is not always possible to take its gradient. Nevertheless, the structure of $g(\lambda)$ (i.e. (19)) still reveals information on the possible search direction. In particular, it is possible to find a subgradient h such that for all λ',

$$g(\lambda') \geq g(\lambda) + h \cdot (\lambda' - \lambda). \tag{26}$$

Since the objective functions of $g(\lambda)$ and $g(\lambda')$ differ in only $(\lambda' - \lambda)(P - \sum_{k=1}^K P_k)$, the following choice of h

$$h = P - \sum_{k=1}^K P_k \tag{27}$$

satisfies the subgradient condition (26). The subgradient search suggests that

$$\begin{array}{lll} \text{increase } \lambda & \text{if} & \sum_{k=1}^K P_k > P \\ \text{decrease } \lambda & \text{if} & \sum_{k=1}^K P_k < P. \end{array}$$

The search direction is intuitively obvious as λ is the water-filling level, and it should be adjusted according to whether the total power constraint is exceeded. Because such adjustment occurs in a one-dimensional space, it can be done efficiently using a bisection method.

The proposed algorithm is summarized as follows:

Algorithm 1 *Sum-power multiple access channel capacity via dual decomposition:*

1. *Initialize λ_{min} and λ_{max}.*
2. *Let $\lambda = (\lambda_{min} + \lambda_{max})/2$.*
3. *Solve for $(S_k, P_k)_{k=1}^K$ in the optimization problem (19) by iteratively optimizing each of (S_k, P_k) using (21)-(25) while keeping all other (S_k, P_k) fixed.*
4. *If $\sum_{k=1}^K P_k > P$, then set $\lambda_{min} = \lambda$, else set $\lambda_{max} = \lambda$.*
5. *If $|\lambda_{min} - \lambda_{max}| \leq \epsilon$, stop. Otherwise, goto step 2.*

The proposed algorithm can also be viewed as an iterative method to solve the KKT condition of the original optimization problem. The KKT condition of the sum-power multiple access channel sum capacity problem (7) consists of the stationarity condition:

$$\frac{1}{2}H_i^T\left(\sum_{k=1}^K H_k S_k H_k^T + S_z\right)^{-1} H_i = \lambda I + \Phi_i, \tag{28}$$

for $i = 1, \cdots, K$, $\lambda \geq 0$, $\Phi_i \geq 0$, and the power constraint:

$$\sum_{k=1}^K \mathrm{tr}(S_k) \leq P, \quad S_k \geq 0. \tag{29}$$

The dual decomposition method starts with a fixed λ, solves (28) for each $i = 1, \cdots, K$, then adjusts λ according to the search direction suggested by the power constraint.

IV. Numerical Results

A numerical example with a 50 transmitters each with a single antenna and a receiver with three antennas is presented. Each entry in the 50×3 channel matrix is an i.i.d. Gaussian random variable with mean 0 and variance 1. The total power constraint is set to be 5. The iterative algorithm is run with a guaranteed error gap less than 10^{-5}. The convergence behavior of the algorithm is plotted in Figure 1. Each iteration of the algorithm consists of a set of 50 water-fillings with a fixed water level. Horizontal segments of the curve represent the number of iterations for each fixed λ. The step shape of the curve is indicative of the bisection algorithm on λ.

The proposed algorithm can have a slower convergence rate as compared to the algorithm proposed by Jindal, Jafar, Vishwanath and Goldsmith in [3]. The main advantage of the current approach is that it guarantees convergence in all cases.

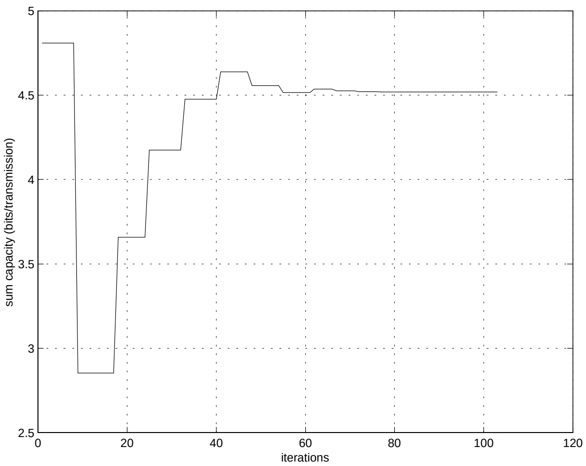

Fig. 1: Convergence of the dual decomposition method

V. Conclusion

This paper proposes a numerical solution to solve the sum capacity in a sum-power constrained multiple access channel. The sum-power constraint is decoupled in the dual domain using a dual decomposition method. The algorithm is based on a modified iterative water-filling method. It is guaranteed to converge to the sum capacity.

References

[1] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum-rate capacity of Gaussian MIMO broadcast channels," *submitted to IEEE Trans. Info. Theory*, August 2002.

[2] P. Viswanath and D. Tse, "Sum capacity of the multiple antenna Gaussian broadcast channel and uplink-downlink duality," *submitted to IEEE Trans. Info. Theory*, July 2002.

[3] N. Jindal, S. Jafar, S. Vishwanath, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," in *Asilomar Conf.*, 2002.

[4] L. Xiao, *Dual decomposition method in network routing and resource allocation*, Ph.D. thesis, Stanford University, 2003.

[5] G. Caire and S. Shamai, "On achievable rates in a multi-antenna broadcast downlink," in *Allerton Conf. on Comm., Control Comp.*, 2000.

[6] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *submitted to IEEE Trans. Info. Theory*, Nov. 2001, available at http://www.comm.utoronto.ca/∼weiyu/publications.

[7] W. Yu, "Uplink-downlink duality under convex covariance constraints," in *Canadian Workshop on Information Theory*, 2003.

[8] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for gaussian vector multiple access channels," *to appear in IEEE Trans. Info. Theory*, Also in Proc. Int. Symp. Info. Theory (ISIT), 2001.