POWER-MINIMIZING AND RATE-MAXIMIZING TRANSCEIVERS WITH INTEGER BIT ALLOCATION: A DUALITY

Chien-Chang Li and Yuan-Pei Lin

Department of Electrical Engineering,
National Chiao Tung University,
1001 Ta Hsueh Road, Hsinchu, Taiwan 300, ROC.

ABSTRACT

In this paper, we consider the bit rate maximizing problem and power minimizing problem with integer bit allocation. We will derive the duality between these two problems. We will show that if a transceiver is optimal for the power-minimizing problem, it is also optimal for the rate maximizing problem, and the converse is true. For the power minimization problem with integer bit allocation, the solution has been found in the literature. However, there is no solution yet for the rate maximization problem with integer bit allocation. The duality allows us to develop an algorithm for finding the rate-maximizing transceiver with integer bit allocation using the solution of power-minimizing system. In the simulations, we will compute the optimal solution for the rate-maximizing problem using the proposed algorithm.

1. INTRODUCTION

MIMO channels arise in applications such as wireless communication systems that use multiple antennas, multicarrier communication systems, and also telephone cables that consist of many twisted pairs. Many criteria have been considered in the transceiver designs for MIMO channels, e.g., [1]-[7]. Optimal transceivers that maximize the mutual information are proposed in [1, 2]. Bit error rate (BER) minimizing transceivers are derived in [3, 4]. Optimal power minimizing transceivers are given in [5, 6]. A unified framework for designing MIMO systems with a power constraint is proposed in [7]. A number of useful objective functions can be considered in this framework. For example, the optimal MMSE transceivers that maximize the bit rate and mutual information can be designed using this unified approach. In [8]-[12], bit allocation is also incorporated in the design of MIMO systems. Optimal transceivers with bit allocation that minimize the transmit power are proposed in [8, 9]. Bit rate maximization systems are developed in [10, 11]. However, the bit allocation obtained in these designs are not integers in general. For the power minimization problem with integer bit allocation, an exhaustive search has been proposed in [12] to find the optimal solution. The transceiver design for maximizing bit rate with integer bit constraint has not been solved.

In the literature, bit rate and power are two commonly used optimality criteria for the transceiver design. When there is no integer constraint on the bit allocation, the duality between these two problems was discussed in [13]. In this paper, we will consider the connection between these two problems when bit allocation is integer constrained. We will show that these two are actually dual problems; the optimal solution obtained in either one problem is also optimal for the other. This conclusion is very similar to the case without integer bit constraint, but the proof is more involved. The duality will be derived without using any existing optimal solution. As a result, the duality can be obtained even for the rate maximizing problem with integer bit allocation, which has not been solved yet before. Furthermore, the duality result can be applied to develop an algorithm to find the optimal solution of the rate maximization problem with integer bit constraint using the solution of the power-minimizing problem. In the simulations, the optimal solutions for the rate-maximizing problem will be computed using the proposed algorithm.

2. SYSTEM MODEL

A generic MIMO communication system is shown in Fig. 1. The MIMO channel is modeled by a $P \times N$ memoryless matrix $\mathbf{H}$. The $P \times 1$ channel noise $\mathbf{q}$ is additive white Gaussian noise with variance N_0. The transmitter matrix $\mathbf{F}$ is of size $N \times M$ with $M \leq \min(P, N)$. The receiver matrix $\mathbf{G}$ is of size $M \times P$. The input of the transmitter is $\mathbf{s}$, an $M \times 1$ vector of modulation symbols. The symbols are assumed to

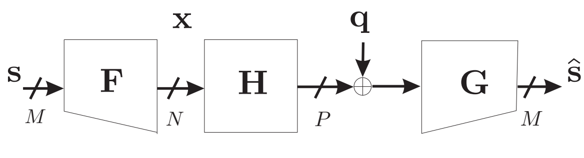

Fig. 1. MIMO communication system.

be zero mean and unit variance, i.e., $\mathrm{E}[s_k] = 0$ and $\sigma^2_{s_k} = 1$ for $k = 0, 1, \cdots, M-1$. The autocorrelation matrix of $\mathbf{s}$

is assumed to be $\mathrm{E}[\mathbf{s}\mathbf{s}^\dagger] = \mathbf{I}_M$, where $\dagger$ denotes the transpose conjugate and the notation $\mathbf{I}_M$ is used to represent the $M \times M$ identity matrix. Hence the total transmit power P is $P = \mathrm{E}\{\mathbf{x}^\dagger \mathbf{x}\} = \sum_{k=0}^{M-1}[\mathbf{F}^\dagger\mathbf{F}]_{kk}$, where $\mathbf{x}$ is the transmitter output indicated in Fig. 1 and the notation $[\mathbf{A}]_{kl}$ denotes the (k,l)-th element of matrix $\mathbf{A}$. The output of the receiver is given by $\hat{\mathbf{s}} = \mathbf{GHFs} + \mathbf{Gq}$. The error vector $\mathbf{e}$ is defined as $\mathbf{e} = \mathbf{s} - \hat{\mathbf{s}}$. The MMSE receiver is given by [14],

$$\mathbf{G} = \mathbf{F}^\dagger\mathbf{H}^\dagger[\mathbf{HFF}^\dagger\mathbf{H}^\dagger + N_0\mathbf{I}_P]^{-1}. \tag{1}$$

The mean-squared error (MSE) matrix $\mathbf{E} = \mathrm{E}[\mathbf{e}\mathbf{e}^\dagger]$ is given by [14] $\mathbf{E} = [N_0^{-1}\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{HF} + \mathbf{I}_M]^{-1}$. The k-th subchannel error variance is $\sigma_{e_k}^2 = [\mathbf{E}]_{kk}$. For QAM modulation, the symbol error rate ϵ_k of the k-th subchannel is well approximated by [15]

$$\epsilon_k \approx 4\left(1 - \frac{1}{2^{b_k/2}}\right)Q\left(\sqrt{\frac{3\beta_k}{(2^{b_k}-1)}}\right), \tag{2}$$

where b_k is the number of bits loaded on the k-th subchannel, and β_k is the signal to interference-plus-noise ratio (SINR) [16]. $\beta_k = 1/\sigma_{e_k}^2 - 1$. The function $Q(x)$ is the area under a Gaussian tail, i.e., $Q(x) = (1/\sqrt{2\pi})\int_x^\infty e^{-u^2/2}du$. The total number of bits that can be transmitted in one block is $B = \sum_{k=0}^{M-1} b_k$.

3. DUALITY OF THE TRANSCEIVER DESIGNS

We consider the power-minimizing and rate-maximizing problems with integer bit allocation. The power-minimizing problem with integer bit allocation is formulated as

$$(\mathscr{A}_{pow,int}) \quad \begin{array}{ll} \underset{\mathbf{F},\{b_k\}}{\text{minimize}} & P = \sum_{k=0}^{M-1}[\mathbf{F}^\dagger\mathbf{F}]_{kk} \\ \text{subject to} & \left\{\begin{array}{l} \sum_{k=0}^{M-1} b_k = B_0, \\ \epsilon_k \le \epsilon,\ 0 \le k \le M-1, \\ b_k \in Z^+,\ 0 \le k \le M-1, \end{array}\right. \end{array} \tag{3}$$

where Z^+ denotes the set of nonnegative integers. The rate-maximizing problem with integer bit allocation is formulated as

$$(\mathscr{A}_{rate,int}) \quad \begin{array}{ll} \underset{\mathbf{F},\{b_k\}}{\text{maximize}} & B = \sum_{k=0}^{M-1} b_k \\ \text{subject to} & \left\{\begin{array}{l} \sum_{k=0}^{M-1}[\mathbf{F}^\dagger\mathbf{F}]_{kk} \le P_0, \\ \epsilon_k \le \epsilon,\ 0 \le k \le M-1, \\ b_k \in Z^+,\ 0 \le k \le M-1. \end{array}\right. \end{array} \tag{4}$$

Theorem 1. *Consider the power-minimizing problem $\mathscr{A}_{pow,int}$ with a target transmission rate B_0 and symbol error rate constraint ϵ. Suppose $(\mathbf{F}^*, \{b_k^*\})$ is optimal for $\mathscr{A}_{pow,int}$, and in this case the minimized power is P^*. Now for the problem $\mathscr{A}_{rate,int}$ with transmit power constraint $P_0 = P^*$ and error rate constraint ϵ, the same $(\mathbf{F}^*, \{b_k^*\})$ also maximizes the transmission rate and the maximized rate is equal to B_0.*

Proof: As $(\mathbf{F}^*, \{b_k^*\})$ is optimal for the problem $\mathscr{A}_{pow,int}$, The bit rate is $B^* = \sum_{k=0}^{M-1} b_k^* = B_0$, and all the symbol error rates satisfy $\epsilon_k^* \le \epsilon$. Now, let us consider the problem $\mathscr{A}_{rate,int}$ with power constraint $P_0 = P^*$ and error rate constraint ϵ. Suppose $(\tilde{\mathbf{F}}, \{\tilde{b}_k\})$ is optimal for the problem $\mathscr{A}_{rate,int}$ and the maximal bit rate is $\tilde{B} = \sum_{k=0}^{M-1}\tilde{b}_k$. All the corresponding error rates $\tilde{\epsilon}_k$ satisfy $\tilde{\epsilon}_k \le \epsilon$ and the transmit power $\tilde{P}$ satisfies the power constraint, i.e., $\tilde{P} \le P^*$. Since we already know the solution of $\mathscr{A}_{pow,int}$ can achieve bit rate B_0 with power P^*, the maximal bit rate $\tilde{B}$ in $\mathscr{A}_{rate,int}$ must be larger than or equal to B_0, i.e., $\tilde{B} \ge B_0$. We will prove the theorem by showing (i) the maximized rate $\tilde{B}$ is in fact equal to B_0, and (ii) the transmit power $\tilde{P}$ is equal exactly to P^*

(i) $\tilde{B} = B_0$: If $\tilde{B} = B_0$, we get the desired result that $(\mathbf{F}^*, \{b_k^*\})$ is optimal for $\mathscr{A}_{rate,int}$. Suppose $\tilde{B} > B_0$ and $(\tilde{\mathbf{F}}, \{\tilde{b}_k\})$ is optimal for $\mathscr{A}_{rate,int}$. First, we will show that there exists a system $(\hat{\mathbf{F}}, \{\hat{b}_k\})$ that achieves a transmit power $\hat{P} < P^*$ with bit rate $\hat{B} = \tilde{B} - 1$ and error rate $\hat{\epsilon}_k \le \epsilon$ for all k. Let $\tilde{\epsilon}_k$ be the error rate on the k-th subchannel of the optimal system. Then $\tilde{\epsilon}_k$ is given by

$$\tilde{\epsilon}_k = 4\left(1 - \frac{1}{2^{\tilde{b}_k/2}}\right)Q\left(\sqrt{\frac{3\tilde{\beta}_k}{(2^{\tilde{b}_k}-1)}}\right), \tag{5}$$

where $\tilde{\beta}_k = 1/\tilde{\sigma}_{e_k}^2 - 1$. The minimized power $\tilde{P}$ is given by $\tilde{P} = \sum_{k=0}^{M-1}[\tilde{\mathbf{F}}^\dagger\tilde{\mathbf{F}}]_{kk} \le P^*$. The bit rate $\tilde{B}$ is $\tilde{B} = \sum_{k=0}^{M-1}\tilde{b}_k$. Suppose $\tilde{b}_{k_0} > 0$ for some k_0-th subchannel. Consider a new system with the bit allocation changed to

$$\hat{b}_k = \left\{\begin{array}{ll} \tilde{b}_{k_0} - 1, & k = k_0, \\ \tilde{b}_k, & \text{otherwise}, \end{array}\right. \tag{6}$$

and the transmitter changed to $\hat{\mathbf{F}} = \tilde{\mathbf{F}}\mathbf{D}$, where $\mathbf{D}$ is a diagonal matrix. $[\mathbf{D}]_{ll} = 1$ for $l \ne k_0$ and $[\mathbf{D}]_{ll} = \mu$ for $l = k_0$. $0 < \mu < 1$ is a positive real number to be chosen later. The bit rate of the new system is $\hat{B} = \tilde{B} - 1 \ge B_0$. The transmit power $\hat{P}$ of the new system is smaller than P^* because $0 < \mu < 1$. Next, we will show that with appropriate choice of μ, the error rate $\hat{\epsilon}_k$ of the new system still satisfies $\hat{\epsilon}_k \le \epsilon$. Using (2), $\hat{\epsilon}_k$ can be expressed as

$$\hat{\epsilon}_k = 4\left(1 - \frac{1}{2^{\hat{b}_k/2}}\right)Q\left(\sqrt{\frac{3\hat{\beta}_k}{(2^{\hat{b}_k}-1)}}\right) \tag{7}$$

$$\le 4\left(1 - \frac{1}{2^{\tilde{b}_k/2}}\right)Q\left(\sqrt{\frac{3\hat{\beta}_k}{(2^{\hat{b}_k}-1)}}\right), \tag{8}$$

where $\hat{\beta}_k = 1/\hat{\sigma}_{e_k}^2 - 1$. Observe that the symbol error rate $\hat{\epsilon}_k$ of the new system will be smaller than $\tilde{\epsilon}_k$ if the quantity in the Q function of (8) is larger than or equal to that in the Q function of (5), i.e.,

$$\frac{1}{2^{\hat{b}_k}-1}\left(\frac{1}{\hat{\sigma}_{e_k}^2} - 1\right) \ge \frac{1}{2^{\tilde{b}_k}-1}\left(\frac{1}{\tilde{\sigma}_{e_k}^2} - 1\right), \forall k. \tag{9}$$

Since $\mu < 1$, we have $\hat{\sigma}^2_{e_k} \leq \tilde{\sigma}^2_{e_k}$, for $k \neq k_0$ (see Lemma 1 in Appendix). This implies $\hat{\epsilon}_k \leq \tilde{\epsilon}_k \leq \epsilon$ for $k \neq k_0$. For $k = k_0$, we can always find $\mu < 1$ such that (9) is satisfied. For example, we can choose

$$\mu = \sqrt{\frac{1}{\tilde{\beta}_{k_0}+1}\left(1+\frac{2^{\hat{b}_{k_0}}-1}{2^{\tilde{b}_{k_0}}-1}\tilde{\beta}_{k_0}\right)}. \tag{10}$$

It can be verified that $1-\mu^2 = \frac{\tilde{\beta}_{k_0}}{\tilde{\beta}_{k_0}+1}\left(1-\frac{2^{\hat{b}_{k_0}}-1}{2^{\tilde{b}_{k_0}}-1}\right) > 0$, and thus $\mu < 1$. In this case, we have

$$\frac{1}{\hat{\sigma}^2_{e_{k_0}}} = \frac{\mu^2}{[\mathbf{B}]_{k_0k_0}} = \frac{\tilde{\sigma}^2_{e_{k_0}}}{[\mathbf{B}]_{k_0k_0}}\left(1+\frac{2^{\hat{b}_{k_0}}-1}{2^{\tilde{b}_{k_0}}-1}\tilde{\beta}_{k_0}\right) \tag{11}$$

$$\geq \left(1+\frac{2^{\hat{b}_{k_0}}-1}{2^{\tilde{b}_{k_0}}-1}\tilde{\beta}_{k_0}\right), \tag{12}$$

where $\mathbf{B} = (N_0^{-1}\tilde{\mathbf{F}}^\dagger\mathbf{H}^\dagger\mathbf{H}\tilde{\mathbf{F}} + \mathbf{D}^{-2})^{-1}$. The last inequality comes from the fact that $\tilde{\sigma}^2_{e_{k_0}} \geq [\mathbf{B}]_{k_0k_0}$. Rearranging (12), we can see that (9) is satisfied for $k = k_0$. Therefore, we have $\hat{\epsilon}_k \leq \tilde{\epsilon}_k \leq \epsilon$ for all k. This means ($\hat{\mathbf{F}}$, $\{\hat{b}_k\}$) can achieve a transmit power $\hat{P} < P^*$ with bit rate $\hat{B} = \tilde{B} - 1$ and error rate $\hat{\epsilon}_k \leq \epsilon$ for all k. Using this technique, we can finally find a system that achieves bit rate $B^{'} = B_0$ with transmit power $P^{'} < P^*$ and error rate $\epsilon^{'}_k \leq \epsilon$. This contradicts the assumption that ($\mathbf{F}^*$, $\{b^*_k\}$) is optimal for $\mathscr{A}_{pow,int}$. Thus we have $\tilde{B} = B_0$.

(ii) $\tilde{P} = P^*$: Suppose $\tilde{P} < P^*$. This means $\tilde{\mathbf{F}}$ and $\{\tilde{b}_k\}$ can achieve a smaller transmit power and still satisfy all the constraints in $\mathscr{A}_{pow,int}$. This contradicts the assumption that $\mathbf{F}^*$ and $\{b^*_k\}$ are optimal for $\mathscr{A}_{pow,int}$. So we have $\tilde{P} = P^*$.

Therefore, we conclude that the maximized bit rate for the problem $\mathscr{A}_{rate,int}$ is B_0 and the power used is P^* and the solution ($\mathbf{F}^*$, $\{b^*_k\}$) of $\mathscr{A}_{pow,int}$ is also optimal for the problem $\mathscr{A}_{rate,int}$. $\triangle\triangle\triangle$

When the symbol error rate constraint ϵ is fixed, the maximal rate for $\mathscr{A}_{rate,int}$ is a function of the power constraint P_0. Similarly, for a fixed ϵ, the minimal power of $\mathscr{A}_{pow,int}$ is a function of target rate B_0. For convenience, we use $P^*_{int}(x)$ to denote the minimal transmit power for $\mathscr{A}_{pow,int}$ when the target bit rate x is given and $B^*_{int}(x)$ to denote the maximal bit rate for $\mathscr{A}_{rate,int}$ when the power constraint is x.

The functions $B^*_{int}(x)$ and $P^*_{int}(x)$. Using theorem 3, we will see that $B^*_{int}(x)$ is not continuous. It is a staircase-like function as shown in Fig. 2(a). This means a nonzero increase in the power constraint does not necessarily implies a nonzero increase in the maximized bit rate. To explain this, consider the problem $\mathscr{A}_{pow,int}$ with two target bit rates B_1 and $B_1 + 1$. Let $P_1 = P^*_{int}(B_1)$ and $P_2 = P^*_{int}(B_1 + 1)$. We can plot the minimal transmit power as a function of target bit rate as in Fig. 2(b). By Theorem 1, we know $B^*_{int}(P_1) = B_1$ and $B^*_{int}(P_2) = B_1 + 1$. Now suppose the power constraint P_0 for $\mathscr{A}_{rate,int}$ is such that $P_1 < P_0 < P_2$. Then the maximal bit rate $B^*_{int}(P_0)$ for $\mathscr{A}_{rate,int}$ is equal to B_1 as we will see next. Since we already know that the maximal bit rate is B_1 when the power constraint is P_1, we have $B^*_{int}(P_0) \geq B_1$. Suppose $B^*_{int}(P_0) > B_1$. This contradicts the fact that P_2 is the minimal power for $\mathscr{A}_{pow,int}$ when the target bit rate is $B_1 + 1$. Hence we have $B^*_{int}(P_0) = B_1$. This implies that for any power constraint P that satisfies $P_1 \leq P_0 < P_2$, the maximal bit rate is $B^*_{int}(P_0) = B_1$. When the power constraint $P_0 = P_2$, the maximal bit rate is increased to $B_1 + 1$. Therefore, $B^*_{int}(x)$ is the staircase-like function in Fig. 2(a).

From the plot of $B^*_{int}(P_0)$ in Fig. 2(a) we can see that for $\mathscr{A}_{rate,int}$ there can be many solutions that achieve the same maximal bit rate, but with transmit power smaller than P_0. To establish the duality with $\mathscr{A}_{pow,int}$, we will consider the solution with the smallest transmit power among all possible solutions.

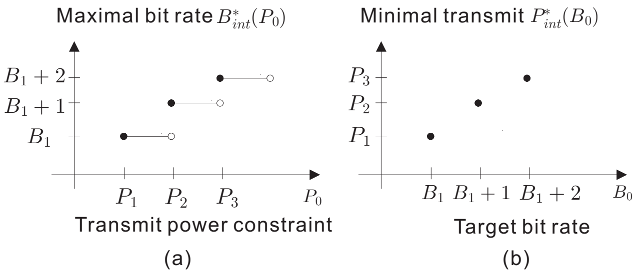

Fig. 2. (a) Maximal bit rate as a function of power constraint P_0 for $\mathscr{A}_{rate,int}$. (b) Minimal transmit power as a function of target bit rate B_0 for $\mathscr{A}_{pow,int}$.

Theorem 2. *Consider the problem $\mathscr{A}_{rate,int}$ with power constraint P_0 and symbol error rate constraint ϵ. Suppose the maximized rate is B^* and ($\mathbf{F}^*$, $\{b^*_k\}$) is the solution that achieves B^* with the smallest transmit power P^* among all possible solutions. Given target rate $B_0 = B^*$ and error rate constraint ϵ for the problem $\mathscr{A}_{pow,int}$, ($\mathbf{F}^*$, $\{b^*_k\}$) also minimizes the transmit power and the minimal power is P^*.*

Proof: As ($\mathbf{F}^*$, $\{b^*_k\}$) is optimal for $\mathscr{A}_{rate,int}$, the maximized rate is $B^* = \sum_{k=0}^{M-1} b^*_k$. The transmit power is $P^* \leq P_0$, and all the error rates satisfy $\epsilon^*_k \leq \epsilon$. Consider the power minimizing problem $\mathscr{A}_{pow,int}$ with target bit rate $B_0 = B^*$ and the same error rate constraint ϵ. Suppose ($\tilde{\mathbf{F}}$, $\{\tilde{b}_k\}$) is optimal for $\mathscr{A}_{pow,int}$ and the minimized power is $\tilde{P}$. The bit rate $\sum_{k=0}^{M-1} \tilde{b}_k$ is equal to the target bit rate B^*. Since we already know ($\mathbf{F}^*$, $\{b^*_k\}$) can achieve bit rate B^* with transmit power P^*, the minimal power $\tilde{P}$ must be smaller than or equal to P^*, i.e., $\tilde{P} \leq P^*$. If $\tilde{P}$ is equal to P^*, we get the desired result that ($\mathbf{F}^*$, $\{b^*_k\}$) is an optimal solution for $\mathscr{A}_{pow,int}$. Assume $\tilde{P}$ is smaller, i.e., $\tilde{P} < P^*$. This means ($\tilde{\mathbf{F}}$, $\{\tilde{b}_k\}$) can achieve bit rate B^* with a smaller power $\tilde{P}$. It contradicts the assumption that ($\mathbf{F}^*$, $\{b^*_k\}$) is the optimal solution for the problem $\mathscr{A}_{rate,int}$ that has the smallest transmit power. Hence we have $\tilde{P} = P^*$ and the solution ($\mathbf{F}^*$, $\{b_k^*\}$) is optimal for $\mathscr{A}_{pow,int}$. $\triangle\triangle\triangle$

Theorem 1 shows that the optimal solution obtained for the power-minimizing problem is also an optimal solution for the rate-maximizing problem. Theorem 2 shows that the solution with the smallest transmit power for the rate-maximizing problem is also optimal for the power-minimizing problem. In this paper, the duality is derived when the subchannel symbol error rate is constrained. Discussion on the case when the averaged bit error rate is constrained can be found in [17].

4. OPTIMAL SOLUTION FOR $\mathscr{A}_{rate,int}$

For the MIMO transceiver designs without integer constraint on bit allocation, the optimal solutions that minimize the transmit power are proposed in [8, 9, 12] and bit rate maximization systems are developed in [10, 11]. For the power minimization problem with integer bit allocation, the solution has been found in [12]. There is no solution yet for the rate maximization problem with integer bit allocation. For $\mathscr{A}_{rate,int}$ with power constraint P_0, if the maximal rate $B^*_{int}(P_0)$ is known, we can solve it using the solution of $\mathscr{A}_{pow,int}$ based on Theorem 1. We can find $B^*_{int}(P_0)$ using an iterative search. For example, starting from $B_0 = 1$ we compute $P^*_{int}(B_0)$. If $P^*_{int}(B_0) \le P_0$, we increase B_0 by one and compute $P^*_{int}(B_0)$ again until $P^*_{int}(B_0) > P_0$. Then $B^*_{int}(P_0) = B_0 - 1$. To reduce the number of iterations we note that $B^*_{int}(P_0) \le B^*(P_0)$, where $B^*(P_0)$ is the maximal bit rate of the rate maximization problem without integer constraint on bit constraint. As a result, $B^*_{int}(P_0) \le \lfloor B^*(P_0) \rfloor$, where the notation $\lfloor x \rfloor$ denotes the largest integer that is less than or equal to x. Using this property and Theorem 1 we have the following algorithm.

Algorithm for finding the solution of $\mathscr{A}_{rate,int}$:

1. Initially, given the power constraint P_0, compute the maximal bit rate $B^*(P_0)$. Then set $B_0 = \lfloor B^*(P_0) \rfloor$.
2. Given the target bit rate B_0, find the optimal bit allocation and transceiver for minimizing transmit power in $\mathscr{A}_{pow,int}$. Compute the minimal power $P^*_{int}(B_0)$.
3. If $P^*_{int}(B_0) > P_0$, set $B_0 = B_0 - 1$ and go to step 2. If $P^*_{int}(B_0) \le P_0$, then the maximal bit rate $B^*_{int}(P_0) = B_0$.

In this algorithm, the number of iterations is equal to $\lfloor B^*(P_0) \rfloor - B^*_{int}(P_0)$. This number is in fact less than M as we explain below. Suppose $\lfloor B^*(P_0) \rfloor - B^*_{int}(P_0) \ge M$. Let $\{b_k^*\}$ be the optimal real-valued bit allocation for the rate maximizing problem without integer constraint, i.e., $B^*(P_0) = \sum_{k=0}^{M-1} b_k^*$. Then $\{\lfloor b_k^* \rfloor\}$ is also a valid integer bit allocation that satisfies the error rate constraint. Since $\{b_k^*\}$ is real, we have $\lfloor B^*(P_0) \rfloor - \sum_{k=0}^{M-1} \lfloor b_k^* \rfloor \le B^*(P_0) - \sum_{k=0}^{M-1} \lfloor b_k^* \rfloor < M$. This implies $\sum_{k=0}^{M-1} \lfloor b_k^* \rfloor > B^*_{int}(P_0)$, which contradicts the definition of $B^*_{int}(P_0)$. Therefore we have $\lfloor B^*(P_0) \rfloor - B^*_{int}(P_0) < M$.

5. SIMULATION

In the simulations, we will compute the optimal solution for $\mathscr{A}_{rate,int}$ using the proposed algorithm Section 4. The number of subchannels M is 4. The noise vector $\mathbf{q}$ is assumed to be complex white Gaussian with $\mathrm{E}[\mathbf{q}\mathbf{q}^\dagger] = \mathbf{I}_4$. The symbol error rate constraint ϵ is assumed to be 10^{-4}. The channel is of size 4×4 and the elements are complex Gaussian random variables. The results are averaged over 10^6 channel realizations. In Table 1, we compute the maximal bit rate of $\mathscr{A}_{rate,int}$. For comparison, we also show the maximal bit rate $B^*(P_0)$ of the case without integer constraint. The gap is less than 0.15 bits per symbol.

P_0 (dB)	$B^*(P_0)$ (bits)	$B^*_{int}(P_0)$ (bits)
2	2.0305	1.4572
4	2.7096	2.1557
6	3.5429	2.9865
8	4.5391	3.9700
10	5.7103	5.1333
12	7.0629	6.4794

Table 1. Bit rate of $\mathscr{A}_{rate,int}$ and the case without integer constraint on bit allocation when the power constraint is $P_0 = 2, 4, 6, 8, 10, 12, 14, 16$ dB.

6. CONCLUSION

In this paper, two commonly used transceiver design criteria were considered: power minimization criterion and bit rate maximization criterion. Both problems are considered with an integer constraint on the bit allocation. The duality between these two problems was derived. Using the duality, the optimal solution of the rate maximization problem can be found. In the simulation, we have computed the optimal solution for the rate maximization problem with integer bit constraint using the proposed algorithm.

Appendix

Lemma 1. *For the MIMO transceiver in Fig. 1, suppose the channel matrix $\mathbf{H}$ is given and the transmitter of the system is replaced by $\mathbf{FD}$, where $\mathbf{F}$ is an $N \times M$ matrix and $\mathbf{D}$ is a diagonal matrix. The diagonal elements of $\mathbf{D}$ are given by*

$$[\mathbf{D}]_{ll} = \begin{cases} 1, & l \neq k_0; \\ \mu, & l = k_0, \end{cases} \tag{13}$$

for some k_0, where μ is a positive real number. Then for $l \neq k_0$, the error variances $\sigma^2_{e_l}$ are increasing and continuous functions of μ. For $l = k_0$ the error variances $\sigma^2_{e_{k_0}}$ is a decreasing and continuous function of μ.

Proof: Given the channel matrix $\mathbf{H}$ and the transmitter $\mathbf{FD}$, the MSE matrix becomes $\tilde{\mathbf{E}} = (N_0^{-1}\mathbf{D}\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{F}\mathbf{D} + \mathbf{I}_M)^{-1}$.

The noise variance $\tilde{\sigma}^2_{e_l}$ is given by $\tilde{\sigma}^2_{e_l} = [\mathbf{D}^{-1}(N_0^{-1}\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{F}+\mathbf{D}^{-2})^{-1}\mathbf{D}^{-1}]_{ll}$. The derivative of $\tilde{\sigma}^2_{e_l}$ with respect to μ is $\partial\tilde{\sigma}^2_{e_l}/\partial\mu = \partial[\tilde{\mathbf{E}}]_{ll}/\partial\mu$. Define $\mathbf{B} = (N_0^{-1}\mathbf{F}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{F} + \mathbf{D}^{-2})^{-1}$. For $l \neq k_0$, we have $\partial[\tilde{\mathbf{E}}]_{ll}/\partial\mu = [\partial\mathbf{B}/\partial\mu]_{ll}$. The derivative of $\mathbf{B}$ with respect to μ is

$$\frac{\partial\mathbf{B}}{\partial\mu} = -\mathbf{B}\frac{\partial\mathbf{B}^{-1}}{\partial\mu}\mathbf{B}^\dagger \tag{14}$$

$$= 2\mu^{-3}\mathbf{b}_{k_0}\mathbf{b}^\dagger_{k_0}, \tag{15}$$

where $\mathbf{b}_{k_0}$ is the k_0-th column of $\mathbf{B}$. Then for $l \neq k_0$ we have $\partial\tilde{\sigma}^2_{e_l}/\partial\mu = \left[2\mu^{-3}\mathbf{b}_{k_0}\mathbf{b}^\dagger_{k_0}\right]_{ll} = 2\mu^{-3}|[\mathbf{B}]_{lk_0}|^2 \geq 0$. Thus we conclude that $\tilde{\sigma}^2_{e_l}$ is an increasing function of μ for $l \neq k_0$. For $l = k_0$, we have

$$\frac{\partial\tilde{\sigma}^2_{e_{k_0}}}{\partial\mu} = \left[\frac{\partial(\mathbf{D}^{-1}\mathbf{B}\mathbf{D}^{-1})}{\partial\mu}\right]_{k_0k_0} = \left[\frac{\partial(\mu^{-2}\mathbf{B})}{\partial\mu}\right]_{k_0k_0}. \tag{16}$$

Using chain rule and the relation $\tilde{\sigma}^2_{e_{k_0}} = \mu^{-2}[\mathbf{B}]_{k_0k_0}$, we can obtain

$$\frac{\partial\tilde{\sigma}^2_{e_{k_0}}}{\partial\mu} = -2\mu^{-3}[\mathbf{B}]_{k_0k_0} + 2\mu^{-5}[\mathbf{B}]^2_{k_0k_0} \tag{17}$$

$$= -2\mu^{-1}\tilde{\sigma}^2_{e_{k_0}}(-1+\tilde{\sigma}^2_{e_{k_0}}). \tag{18}$$

Since $\tilde{\mathbf{E}}^{-1} - \mathbf{I}_M \geq 0$, we know $\mathbf{I}_M - \tilde{\mathbf{E}} \geq 0$. Then we have $\tilde{\sigma}^2_{e_{k_0}} \leq 1$ and thus $\partial\tilde{\sigma}^2_{e_{k_0}}/\partial\mu \leq 0$. As a result, we can conclude that $\tilde{\sigma}^2_{e_{k_0}}$ is a decreasing function of μ. △△△

A. REFERENCES

[1] E. Telatar, ”Capacity of multi-antenna Gaussian channels,” *European Transactions on telecommunications*, vol. 10, pp. 585-595, November/December 1999.

[2] W. Yu, T. Lan, “Transmitter Optimization for the Multi-Antenna Downlink With Per-Antenna Power Constraints,” *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2646-2660, June 2007.

[3] Y. Ding, T. N. Davidson, Z.-Q. Luo, and K.-M. Wong, “Minimum BER block precoders for zero-forcing equalization,” *IEEE Transactions on Signal Processing*, vol. 51, pp. 2410-2423, Sep. 2003.

[4] Y.-P. Lin and S.-M. Phoong, “BER minimized OFDM systems with channel independent precoders,” *IEEE Transactions on Signal Processing*, vol. 51, pp. 2369-2380, Sept. 2003.

[5] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, “Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints,” *IEEE Transactions on Signal Processing*, vol. 52, pp. 1179- 1197, May 2004.

[6] C.-H. F. Fung, W. Yu, T. J. Lim, “Precoding for the Multiantenna Downlink: Multiuser SNR Gap and Optimal User Ordering,” *IEEE Transactions on Communications*, vol. 55, no. 1, pp. 188-197, January 2007.

[7] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, “Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization,” *IEEE Transactions on Signal Processing*, vol. 51, pp. 2381-2401, Sept. 2003.

[8] S. Dasgupta and A. Pandharipande “Optimum multiflow biorthogonal DMT with unequal subchannel assignment,” *IEEE Transactions on Signal Processing*, vol. 53, pp. 3572-3582, September 2005.

[9] Y.-P. Lin and S.-M. Phoong, “Optimal ISI-free DMT transceivers for distorted channels with colored noise,” *IEEE Transactions on Signal Processing*, vol. 49, pp. 2702-2712, November 2001.

[10] P. P. Vaidyanathan, Yuan-Pei Lin, Sony Akkarakaran, See-May Phoong, “Discrete multitone modulation with principal component filter banks,” *IEEE Transactions on circuits and systems I: fundamental theory and applications* , vol. 49, pp. 1397-1412, october 2002.

[11] C.-C. Li, Y.-P. Lin, S.-H. Tsai, and P. P. Vaidyanathan, “Optimization of Transceivers with Bit Allocation to Maximize Bit Rate for MIMO Transmission,” *IEEE Transactions on Communications*, vol. 57, no. 11, pp. 3556-3560, Dec. 2009.

[12] D. P. Palomar, S. Barbarossa, “Designing MIMO communication systems: constellation choice and linear transceiver design,” *IEEE Transactions on Signal Processing*, vol. 53, pp. 3804-3818, October 2005.

[13] C.-C. Li and Y.-P. Lin, ”On the Duality of Transceiver Designs for MIMO Channels,” National Symposium on Telecommunications, Taiwain, December 2009.

[14] P. P. Vaidyanathan, See-May Phoong, and Yuan-Pei Lin, *Signal Processing and Optimization for Transceiver Systems.* Cambridge University Press, 2010.

[15] J. G. Proakis, *Digital Communications*, third edition New York: McGraw- Hill, 1995.

[16] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, G. D. Forney Jr., “MMSE decision-feedback equalizers and coding. I. Equalization results,” *IEEE Transactions on Communications*, vol. 43, no. 10, pp. 2582-2594, October 1995.

[17] C. C. Li and Y. P. Lin, “On the Duality of MIMO Transceiver Designs with Bit Allocation,” accepted and to be published in *IEEE Transactions on Signal Processing*.