Convexity Analysis and Splitting Algorithm for the Sum-Rate Maximization Problem

Guidance

Assistant Professor Shunsuke HAYASHI
Professor Masao FUKUSHIMA

Daisuke YAMAMOTO

Department of Applied Mathematics and Physics

Graduate School of Informatics

Kyoto University

February 2010

Abstract

We consider the dynamic spectrum management (DSM) system in which many users share the same frequency band and it is divided into multiple tones. In such a system, each user or central office allocates the transmission power to each tone adaptively in response to channel conditions. However, if multiple users allocate their transmission powers to the same tone, then their data rate may decrease due to the electromagnetic interference called crosstalk. Therefore, it is required to allocate the transmission power so that every user can achieve a sufficient data rate.

Although the sum-rate (the sum of all users' data rates) is the most typical measure of overall system performance, it is nonconcave in general, and hence, the sum-rate maximization problem may have multiple local maxima. Thus, the algorithms that have been proposed so far could find nothing more than a "near-optimal" solution under general conditions.

In this paper, we provide conditions under which the sum-rate function is guaranteed to be concave over the feasible region. Moreover, by utilizing the concavity, we propose some splitting algorithms that can be implemented in a distributed manner. From the numerical experiments, we observe that the splitting algorithms solve the sum-rate maximization problems efficiently.

Contents

1 Introduction

Recently, the dynamic spectrum management (DSM) [22] is attracting much attention in the multiuser communication systems such as the digital subscriber line (DSL) and wireless system. Although the current system called static spectrum management (SSM) cannot achieve sufficient data rate, DSM has a potential to improve the system performance drastically due to its adaptive allocation scheme for each user's transmission power.

In this paper, we focus on the system in which many users share the same frequency band and it is divided into multiple parallel subchannels called *tones*. In such a system, the tone-wise allocation of the transmission power can be regarded as each user's power spectrum. Therefore, in the DSM system, each user or central office allocates the transmission power to each tone adaptively in response to channel conditions.

When multiple users allocate their transmission powers to the same tone, the electro-magnetic interference called *crosstalk* arises and may decrease their data rates quite a bit. Therefore, it is important to allocate the transmission power so that the effect of the crosstalk can be mitigated and each user can achieve a sufficient data rate. In this paper, we adopt the *sum-rate* (the sum of all users' data rates) as the measure of overall system performance. However, the problem of maximizing the sum-rate may have multiple local maxima because of the nonconcavity. In addition, it is known to be NP-hard to find a global optimum of the sum-rate maximization problem [13]. Therefore, the major research objective in this field is to provide an approximation method for finding a "near-optimal" solution.

So far, a number of algorithms have been proposed for solving power allocation problems such as the sum-rate maximization. Particularly, existing algorithms can be classified into the following three types: (i) algorithms based on game-theoretical concepts [8, 16, 19, 21, 20, 22, 25, 26], (ii) algorithms based on the duality theory [5, 6, 15, 27], and (iii) algorithms based on the idea of frequency division multiple access (FDMA)[1] [13, 24]. In the game-theoretical methods, the power allocation problem is regarded as a non-cooperative game, in which each user's power spectrum and achievable data rate correspond to the player's strategy and payoff, respectively. Iterative water-filling algorithm (IWFA) [26] is one of popular game-theoretical methods, in which each user maximizes his/her own data rate successively so that the generated sequence of the power allocations converges to a Nash equilibrium. In fact, when the crosstalk coefficients are small, IWFA is guaranteed to be convergent and provides a sufficiently large sum-rate [8, 16, 26]. On the other hand, when the crosstalk coefficients are large, IWFA may not converge and the obtained solution can be far from optimal. The duality theory based algorithms are to solve the Lagrangian dual problem instead of the original sum-rate maximization. Since the dual problem itself is convex and can be decomposed into lower dimensional problems, we may not suffer from the local (non-global) optimality or high computational cost. However, in order to evaluate the dual function, we have to maximize a certain non-concave function. Moreover, the duality gap is positive[2] in general. Thus, the global optimality of the obtained solution cannot be guaranteed at all. Very recently, Hayashi and Luo [13] proposed algorithms based on the idea of FDMA. They showed that, if the crosstalk coefficients for all tones are larger than a certain threshold, then the global optimum of the sum-rate maximization problem is FDMA. In addition, based on the FDMA idea, they proposed several efficient algorithms which can find a solution with high sum-rate especially when the crosstalk coefficients are large. On the other hand, their FDMA

[1] In an FDMA system, each tone can be used at most one user.

[2] In [27, 17], it is claimed that the duality gap converges to zero as the number of tones goes to infinity.

based algorithms show relatively poor performance when the crosstalk coefficients are small. Improving Hayashi and Luo's algorithms, Yamamoto [24] proposed some algorithms based on a partial FDMA[3] idea. This type of algorithms is efficient particularly for a system where some tones have strong crosstalk and others have weak crosstalk.

The above-mentioned methods are developed on the premise that the sum-rate function is nonconcave. Hence, they cannot find a global optimum in general. However, is the sum-rate function always nonconcave over the feasible region? Or, does it become concave under certain conditions? Actually, we can easily see that the sum-rate function is strictly concave when all crosstalk coefficients are zero. Moreover, the feasible set of the sum-rate maximization problem is compact (See Section 2). Therefore, it can be expected that, even if the crosstalk coefficients are not zero, the sum-rate function may still be concave over the feasible region under weak crosstalk conditions. The purpose of this paper is to provide such conditions in an explicit manner, by analyzing the Hessian matrix of the sum-rate function. In addition, we propose splitting algorithms for solving the sum-rate maximization problem in a distributed fashion.

This paper is organized as follows. In Section 2, we review the system model and the mathematical formulation of the sum-rate maximization problem. In Section 3, we derive sufficient conditions under which the sum-rate function becomes concave over the feasible region. In Section 4, we propose algorithms that utilize the special structure of the sum-rate maximization problem. Some numerical results are reported in Section 5, and the concluding remarks are given in Section 6.

Throughout the paper, we use the following notations. $E(z)$ denotes the expected value of random variable z. For a vector $\boldsymbol{x} = (x_1, \ldots, x_m)^\top$, $\mathrm{diag}(\boldsymbol{x})$ denotes the diagonal matrix with diagonal entries $x_1, \ldots, x_m$. For square matrices $X_1, \ldots, X_m$, $\mathrm{diag}(X_1, \ldots, X_m)$ denotes the block-diagonal matrix with diagonal entries $X_1, \ldots, X_m$. For three numbers a, b and $c \in \mathbb{R} \cup \{\pm\infty\}$, $\mathrm{med}(a, b, c)$ denotes the median.

2 Channel model and the sum-rate maximization problem

In this section, we describe the frequency selective Gaussian interference channel model and mathematical formulation of the sum-rate maximization problem.

Suppose that there exist K users ($K \geq 2$) sharing a common frequency band divided into N tones. For notational simplicity, we denote the sets of tones and users $\mathcal{N} := \{1, \ldots, N\}$ and $\mathcal{K} := \{1, \ldots, K\}$, respectively. Then, signal y_k^n that user $k \in \mathcal{K}$ receives at tone $n \in \mathcal{N}$ is given by

$$y_k^n := \sum_{l=1}^{K} h_{lk}^n x_l^n + z_k^n,$$

where $x_l^n \in \mathbb{C}$ denotes the signal from user $l \in \mathcal{K}$ at tone $n \in \mathcal{N}$, $h_{lk}^n \in \mathbb{C}$ represents the channel gain coefficient depending on the distance between user l and user k, and $z_k^n \in \mathbb{C}$ denotes the complex Gaussian channel noise following the complex normal distribution with zero mean and variance N_0. Then, the signal to noise ratio (SNR) for user k at tone n is given by

$$\mathrm{SNR} = \frac{E(\,|h_{kk}^n x_k^n|^2)}{E(\,|\sum_{l \neq k} h_{lk}^n x_l^n + z_k^n|^2\,)}.$$

[3]A power allocation is called partial FDMA if some (not all) tones are used at most one user.

By Shannon's information theory [10], user k's achievable data rate R_k^n at tone n can be written as

$$\begin{aligned} R_k^n(\boldsymbol{S}^n) &:= \log(1+\mathrm{SNR}) \\ &= \log\left(1+\frac{|h_{kk}^n|^2 S_k^n}{N_0+\sum_{l\neq k}|h_{lk}^n|^2 S_l^n}\right), \end{aligned} \tag{2.1}$$

where $S_k^n := E(\,|x_k^n|^2\,)$ represents user k's signal power at tone n, and $\boldsymbol{S}^n := (S_1^n,\ldots,S_K^n)^\top \in \mathbb{R}^K$ denotes the vector of all users' signal powers at tone n. Dividing both the numerator and denominator in (2.1) by $|h_{kk}^n|^2$, we further obtain

$$R_k^n(\boldsymbol{S}^n) = \log\left(1+\frac{S_k^n}{\sigma_k^n+\sum_{l\neq k}\alpha_{lk}^n S_l^n}\right),$$

where $\sigma_k^n := N_0/|h_{kk}^n|^2$ denotes user k's normalized background noise power at tone n and $\alpha_{lk}^n := |h_{lk}^n|^2/|h_{kk}^n|^2$ is the normalized crosstalk coefficient[4] from user l to user k at tone n. Therefore, user k's total achievable data rate R_k is given by

$$\begin{aligned} R_k(\boldsymbol{S}) &:= \sum_{n=1}^N R_k^n(\boldsymbol{S}^n) \\ &= \sum_{n=1}^N \log\left(1+\frac{S_k^n}{\sigma_k^n+\sum_{l\neq k}\alpha_{lk}^n S_l^n}\right), \end{aligned} \tag{2.2}$$

where $\boldsymbol{S} := ((\boldsymbol{S}^1)^\top,\ldots,(\boldsymbol{S}^N)^\top)^\top \in \mathbb{R}^{NK}$ denotes the signal power vector in the whole system.

Throughout this paper, we assume that, for each $n \in \mathcal{N}$ and $k \in \mathcal{K}$, the signal power S_k^n is bounded by the spectral mask $S_{\max,k}^n \geq 0$ and that user k's total signal power is bounded by the power budget $P_k > 0$, i.e.,

$$0 \leq S_k^n \leq S_{\max,k}^n, \tag{2.3}$$

$$\sum_{n=1}^N S_k^n \leq P_k. \tag{2.4}$$

Without loss of generality, we can assume that

$$P_k \leq \sum_{n=1}^N S_{\max,k}^n, \quad S_{\max,k}^n \leq P_k, \tag{2.5}$$

since we have only to redefine $P_k := \sum_{n=1}^N S_{\max,k}^n$ or $S_{\max,k}^n := P_k$ without changing the feasible region if $P_k > \sum_{n=1}^N S_{\max,k}^n$ or $S_{\max,k}^n > P_k$, respectively.

As is shown in formula (2.2), if the system tries to increase a certain user's data rate only, then other users' data rates may decrease because of the crosstalk. Therefore, it is required to introduce a reasonable utility function representing the overall system performance. Although many functions have been proposed so far (e.g., arithmetic mean, geometric mean, harmonic

[4]Due to the normalization, we always have $\alpha_{kk}^n = 1$ for all k.

mean, minimum of all users' data rates, etc. [7, 17]), we adopt the most typical one, sum-rate function R defined by

$$\begin{aligned} R(\boldsymbol{S}) &:= \sum_{k=1}^{K} R_k(\boldsymbol{S}) \\ &= \sum_{n=1}^{N}\sum_{k=1}^{K} \log\left(1 + \frac{S_k^n}{\sigma_k^n + \sum_{l\neq k} \alpha_{lk}^n S_l^n}\right). \end{aligned} \tag{2.6}$$

Note that function R can also be expressed in a tone-wise manner, that is, $R(\boldsymbol{S}) = \sum_{n=1}^{N} R^n(\boldsymbol{S}^n)$ with "tone-rate" function R^n defined by

$$R^n(\boldsymbol{S}^n) := \sum_{k=1}^{K} \log\left(1 + \frac{S_k^n}{\sigma_k^n + \sum_{l\neq k} \alpha_{lk}^n S_l^n}\right). \tag{2.7}$$

Hence, the sum-rate maximization problem can be written as

$$\begin{aligned} \underset{\boldsymbol{S}^1,\dots,\boldsymbol{S}^N}{\text{maximize}} \quad & \sum_{n=1}^{N} R^n(\boldsymbol{S}^n) \\ \text{subject to} \quad & \sum_{n=1}^{N} \boldsymbol{S}^n \leq \boldsymbol{P}, \\ & \boldsymbol{S}^n \in \Omega^n \ (n \in \mathcal{N}), \end{aligned} \tag{2.8}$$

where

$$\boldsymbol{P} := (P_1, \dots, P_K)^\top \in \mathbb{R}^K, \quad \Omega^n := \prod_{k=1}^{K} [0,\ S_{\max,k}^n].$$

In the remainder of the paper, we denote by $\mathcal{F}$ the feasible region of the sum-rate maximization problem (2.8), i.e.,

$$\mathcal{F} := \left\{ \boldsymbol{S} \in \mathbb{R}^{NK} \;\middle|\; \sum_{n=1}^{N} \boldsymbol{S}^n \leq \boldsymbol{P},\ \boldsymbol{S}^n \in \Omega^n \ (n \in \mathcal{N}) \right\}.$$

3 Concavity analyses for sum-rate function

There is no guarantee that the sum-rate maximization problem (2.8) becomes convex since the sum-rate function R is nonconcave in general. However, the sum-rate function has the following properties.

- If $\alpha_{lk}^n = 0$ for all $(n, k, l) \in \mathcal{N} \times \mathcal{K} \times \mathcal{K}$ with $k \neq l$, then the sum-rate function is strictly concave, i.e., $\nabla_{\boldsymbol{S}}^2 R(\boldsymbol{S}) \prec 0$.
- The sum-rate function and its Hessian matrix are continuous with respect to the crosstalk coefficients $(\alpha_{lk}^n)_{(n,k,l)\in\mathcal{N}\times\mathcal{K}\times\mathcal{K}}$.

Hence, noticing the compactness of the feasible region $\mathcal{F}$ of problem (2.8), we can expect that the sum-rate function R is still concave over $\mathcal{F}$ when all the crosstalk coefficients α_{lk}^n $(k \neq l)$ are positive but sufficiently small. Moreover, in general, the maximum eigenvalue of the Hessian matrix $\nabla^2_{\boldsymbol{S}} R(\boldsymbol{S})$ tends to be small as $\boldsymbol{S} \geq 0$ approaches the origin. This property implies that problem (2.8) is more likely to be convex as each spectral mask $(S^n_{\max,k})_{(n,k)\in\mathcal{N}\times\mathcal{K}}$ becomes small. In this section, we analyze the mathematical properties of the sum-rate maximization problem (2.8) and derive sufficient conditions for problem (2.8) to be convex.

We first provide some mathematical preliminaries. For notational simplicity, we define

$$
\begin{aligned}
\partial_k R^n(\boldsymbol{S}^n) &:= \frac{\partial}{\partial S^n_k} R^n(\boldsymbol{S}^n),\\
\partial_{kl} R^n(\boldsymbol{S}^n) &:= \frac{\partial^2}{\partial S^n_k \partial S^n_l} R^n(\boldsymbol{S}^n),\\
X^n_k(\boldsymbol{S}^n) &:= \sigma^n_k + \sum_{l\neq k} \alpha^n_{lk} S^n_l,\\
A^n_k(\boldsymbol{S}^n) &:= \frac{1}{X^n_k(\boldsymbol{S}^n)},\\
B^n_k(\boldsymbol{S}^n) &:= \frac{1}{X^n_k(\boldsymbol{S}^n) + S^n_k},\\
P^n_k(\boldsymbol{S}^n) &:= A^n_k(\boldsymbol{S}^n) - B^n_k(\boldsymbol{S}^n),\\
Q^n_k(\boldsymbol{S}^n) &:= A^n_k(\boldsymbol{S}^n)^2 - B^n_k(\boldsymbol{S}^n)^2.
\end{aligned}
$$

With these notations, the explicit formulas of the gradient and Hessian matrix of the tone-rate function R^n are given in the following proposition.

Proposition 1 ([13], Proposition 3.1)**.** *For every $(n,k,l) \in \mathcal{N}\times\mathcal{K}\times\mathcal{K}$ such that $k \neq l$, we have*

$$
\begin{aligned}
\partial_k R^n(\boldsymbol{S}^n) &= B^n_k - \sum_{r\neq k} \alpha^n_{kr} P^n_r,\\
\partial_{kk} R^n(\boldsymbol{S}^n) &= -(B^n_k)^2 + \sum_{r\neq k} (\alpha^n_{kr})^2 Q^n_r,\\
\partial_{kl} R^n(\boldsymbol{S}^n) &= -\alpha^n_{lk}(B^n_k)^2 - \alpha^n_{kl}(B^n_l)^2 + \sum_{r\neq k,l} \alpha^n_{kr}\alpha^n_{lr} Q^n_r.
\end{aligned}
$$

The following proposition claims that we have only to check the Hessian matrices of each tone-rate function to analyze the concavity of the sum-rate function.

Proposition 2. (i) *The sum-rate function R is concave over the feasible region $\mathcal{F}$ of problem* (2.8) *if and only if* (ii) $-\nabla^2_{\boldsymbol{S}^n} R^n(\boldsymbol{S}^n) \succeq 0$ *for any $\boldsymbol{S}^n \in \Omega^n$ and $n \in \mathcal{N}$.*

Proof. First we show (ii) ⇒ (i). If (ii) holds, then we have $-\nabla^2_{\boldsymbol{S}} R(\boldsymbol{S}) \succeq 0$ over $\Omega := \prod_{n=1}^N \Omega^n$ since

$$\nabla^2_{\boldsymbol{S}} R(\boldsymbol{S}) = \operatorname{diag}\left(\nabla^2_{\boldsymbol{S}^1} R^1(\boldsymbol{S}^1), \ldots, \nabla^2_{\boldsymbol{S}^N} R^N(\boldsymbol{S}^N)\right).$$

Noticing $\mathcal{F} \subseteq \Omega$, we immediately have (i).

Next we show (i) ⇒ (ii) by contraposition. If (ii) does not hold, then there exist $n' \in \mathcal{N}$ and $\bar{\boldsymbol{S}}^{n'} \in \Omega^{n'}$ such that

$$-\nabla^2_{\boldsymbol{S}^{n'}} R^n(\bar{\boldsymbol{S}}^{n'}) \not\succeq 0. \tag{3.1}$$

Now, let $\boldsymbol{S} := ((\boldsymbol{S}^1)^\top, \ldots, (\boldsymbol{S}^N)^\top)^\top$ be defined as $\boldsymbol{S}^n := \boldsymbol{0}$ for all $n \neq n'$ and $\boldsymbol{S}^n := \bar{\boldsymbol{S}}^{n'}$ for $n = n'$. Then, we have $-\nabla^2_{\boldsymbol{S}} R(\boldsymbol{S}) \not\succeq 0$ from (3.1). This together with $\boldsymbol{S} \in \mathcal{F}$ from (2.5) implies that statement (i) does not hold. □

Proposition 2 indicates that, in order to derive the convexity conditions for the sum-rate function R, we have to check the positive semidefiniteness of $-\nabla^2_{\boldsymbol{S}^n} R^n(\boldsymbol{S}^n)$ for each $n \in \mathcal{N}$. However, in general, it is difficult to give the conditions directly since the matrix $-\nabla^2_{\boldsymbol{S}^n} R^n(\boldsymbol{S}^n)$ has a complicated structure. For this reason, in this paper, we make use of the diagonal dominance property of the matrices. In fact, we know that a symmetric diagonally dominant matrix with nonnegative diagonal entries is positive semidefinite.

Definition 1. *A symmetric matrix $A = [a_{ij}] \in \mathbb{R}^{m\times m}$ is said to be diagonally dominant if*

$$|a_{ii}| \geq \sum_{j\neq i} |a_{ij}|, \quad i = 1, \ldots, m.$$

Proposition 3 ([9], Proposition 2.2.20)**.** *If a symmetric matrix $A = [a_{ij}] \in \mathbb{R}^{m\times m}$ is diagonally dominant and has nonnegative diagonal entries, i.e.,*

$$a_{ii} \geq \sum_{j\neq i} |a_{ij}|, \quad i = 1, \ldots, m,$$

then A is positive semidefinite.

Now, using Propositions 1–3, we provide sufficient conditions for the sum-rate function R to be convex over the feasible region $\mathcal{F}$ of problem (2.8).

Theorem 1.

(a) *The tone-rate function R^n is concave over Ω^n if*

$$\frac{1}{(\sigma_k^n + \sum_{i\neq k} \alpha_{ik}^n S_{\max,i}^n + S_{\max,k}^n)^2} - \sum_{l\neq k} \left[\frac{\alpha_{lk}^n}{(\sigma_k^n)^2} + \frac{\alpha_{kl}^n}{(\sigma_l^n)^2}\right] - \sum_{l=1}^{K} \sum_{r\neq k,l} \left[\frac{\alpha_{kr}^n \alpha_{lr}^n}{(\sigma_r^n)^2} - \frac{\alpha_{kr}^n \alpha_{lr}^n}{(\sigma_r^n + S_{\max,r}^n)^2}\right] \geq 0 \tag{3.2}$$

for all $k \in \mathcal{K}$.

(b) *The sum-rate function R is concave over $\mathcal{F}$ if* (3.2) *holds for all $(n,k) \in \mathcal{N} \times \mathcal{K}$.*

Proof. From Proposition 2, we can easily see that (a) implies (b). Therefore, we only prove (a). For all $k \in \mathcal{K}$, we have the following inequalities for any $\boldsymbol{S}^n \in \Omega^n$:

$$\sigma_k^n \leq X_k^n(\boldsymbol{S}^n) \leq \sigma_k^n + \sum_{l\neq k} \alpha_{lk}^n S_{\max,l}^n, \tag{3.3}$$

$$\frac{1}{\sigma_k^n + \sum_{l\neq k} \alpha_{lk}^n S_{\max,l}^n} \leq A_k^n(\boldsymbol{S}^n) \leq \frac{1}{\sigma_k^n}, \tag{3.4}$$

$$\frac{1}{\sigma_k^n + \sum_{l\neq k} \alpha_{lk}^n S_{\max,l}^n + S_{\max,k}^n} \leq B_k^n(\boldsymbol{S}^n) \leq \frac{1}{\sigma_k^n}. \tag{3.5}$$

In addition, the following inequalities hold for any $\boldsymbol{S}^n \in \Omega^n$ and $k \in \mathcal{K}$:

$$\begin{aligned} 0 \leq Q_k^n(\boldsymbol{S}^n) &= \frac{1}{(X_k^n(\boldsymbol{S}^n))^2} - \frac{1}{(X_k^n(\boldsymbol{S}^n) + S_k^n)^2} \\ &\leq \frac{1}{(X_k^n(\boldsymbol{S}^n))^2} - \frac{1}{(X_k^n(\boldsymbol{S}^n) + S_{\max,k}^n)^2} \\ &\leq \frac{1}{(\sigma_k^n)^2} - \frac{1}{(\sigma_k^n + S_{\max,k}^n)^2}, \end{aligned} \tag{3.6}$$

where the last inequality follows from the fact that, for given $\beta > 0$, function $f(x) := 1/x^2 - 1/(x+\beta)^2$ is monotonically decreasing on $\{x \,|\, x > 0\}$.

Then, for all $k \in \mathcal{K}$, we have

$$
\begin{aligned}
&-\partial_{kk} R^n(\boldsymbol{S}^n) - \sum_{l \neq k} |-\partial_{kl} R^n(\boldsymbol{S}^n)| \\
&= \left[(B_k^n)^2 - \sum_{r \neq k} (\alpha_{kr}^n)^2 Q_r^n \right] - \sum_{l \neq k} \left[\left| -\alpha_{lk}^n (B_k^n)^2 - \alpha_{kl}^n (B_l^n)^2 + \sum_{r \neq k,l} \alpha_{kr}^n \alpha_{lr}^n Q_r^n \right| \right] \\
&\geq \left[(B_k^n)^2 - \sum_{r \neq k} (\alpha_{kr}^n)^2 Q_r^n \right] - \sum_{l \neq k} \left[\alpha_{lk}^n (B_k^n)^2 + \alpha_{kl}^n (B_l^n)^2 + \sum_{r \neq k,l} \alpha_{kr}^n \alpha_{lr}^n Q_r^n \right] \\
&= (B_k^n)^2 - \sum_{l \neq k} \left(\alpha_{lk}^n (B_k^n)^2 + \alpha_{kl}^n (B_l^n)^2 \right) - \sum_{l=1}^{K} \sum_{r \neq k,l} \alpha_{kr}^n \alpha_{lr}^n Q_r^n \\
&\geq \frac{1}{(\sigma_k^n + \sum_{i \neq k} \alpha_{ik}^n S_{\max,i}^n + S_{\max,k}^n)^2} - \sum_{l \neq k} \left[\frac{\alpha_{lk}^n}{(\sigma_k^n)^2} + \frac{\alpha_{kl}^n}{(\sigma_l^n)^2} \right] \\
&\qquad - \sum_{l=1}^{K} \sum_{r \neq k,l} \left[\frac{\alpha_{kr}^n \alpha_{lr}^n}{(\sigma_r^n)^2} - \frac{\alpha_{kr}^n \alpha_{lr}^n}{(\sigma_r^n + S_{\max,r}^n)^2} \right] \\
&\geq 0,
\end{aligned} \tag{3.7}
$$

where the first inequality follows from the triangle inequality, the second inequality follows from (3.3)–(3.6), and the last inequality follows from (3.2). Thus, (3.7) together with Proposition 3 yields $-\nabla^2 R^n(\boldsymbol{S}^n) \succeq 0$, i.e., the tone-rate function R^n is concave. □

Remark 1. *If $\alpha_{lk}^n = 0$ for all $(n,k,l) \in \mathcal{N} \times \mathcal{K} \times \mathcal{K}$ with $k \neq l$, then we can easily see that* (3.2) *automatically holds. In addition,* (3.2) *is more likely to hold as $S_{\max,k}^n$ becomes smaller since $S_{\max,k}^n$ appears only in the denominators of the positive terms. However, even if $S_{\max,k}^n = 0$ for all $k \in \mathcal{K}$, i.e.,$\mathcal{F} = \{0\}$, inequality* (3.2) *may not hold. If $S_{\max,k}^n = 0$ for all $k \in \mathcal{K}$, then* (3.2) *can be rewritten as*

$$
\frac{1}{(\sigma_k^n)^2} - \sum_{l \neq k} \left[\frac{\alpha_{lk}^n}{(\sigma_k^n)^2} + \frac{\alpha_{kl}^n}{(\sigma_l^n)^2} \right] \geq 0, \tag{3.8}
$$

which is invalid when $1 - \sum_{l \neq k} \alpha_{lk}^n < 0$ for some $(n,k) \in \mathcal{N} \times \mathcal{K}$. Anyhow, we emphasize that the sum-rate function R can be concave over $\mathcal{F}$ even if (3.2) *does not hold, since Theorem* 1 *provides nothing more than sufficient conditions.*

Although inequality (3.2) only involves the parameters of the sum-rate maximization problem (2.8), it is still difficult to see the effects of the noise coefficients σ_k^n, the crosstalk coefficients α_{lk}^n, and the spectral masks $S_{\max,k}^n$. In order to understand (3.2) more intuitively, we give a sufficient condition for (3.2) by using the maximum and minimum values of σ_k^n, α_{lk}^n, and $S_{\max,k}^n$.

Let us define

$$\begin{aligned}
\alpha^n_{\max} &:= \max_{l,k,\ l\neq k} \alpha^n_{lk}, \\
\alpha^n_{\min} &:= \min_{l,k,\ l\neq k} \alpha^n_{lk}, \\
\sigma^n_{\max} &:= \max_k \sigma^n_k, \\
\sigma^n_{\min} &:= \min_k \sigma^n_k, \\
S^n_{\max} &:= \max_k S^n_{\max,k}.
\end{aligned}$$

Then, we have the following corollary to Theorem 1.

Corollary 1. *The tone-rate function R^n is concave over Ω^n if*

$$\frac{1}{\left(\sigma^n_{\max} + ((K-1)\alpha^n_{\max} + 1)\, S^n_{\max}\right)^2} + \frac{(K-1)^2(\alpha^n_{\min})^2}{(\sigma^n_{\max} + S^n_{\max})^2} - 2(K-1)\frac{\alpha^n_{\max}}{(\sigma^n_{\min})^2} - (K-1)^2\frac{(\alpha^n_{\max})^2}{(\sigma^n_{\min})^2} \geq 0 \tag{3.9}$$

for all $k \in \mathcal{K}$.

Proof. From Theorem 1(a), we have only to show (3.2). By (3.9), we have for any $k \in \mathcal{K}$

$$\begin{aligned}
&\frac{1}{(\sigma^n_k + \sum_{i\neq k} \alpha^n_{ik} S^n_{\max,i} + S^n_{\max,k})^2} - \sum_{l\neq k}\left[\frac{\alpha^n_{lk}}{(\sigma^n_k)^2} + \frac{\alpha^n_{kl}}{(\sigma^n_l)^2}\right] \\
&\qquad - \sum_{l=1}^{K}\sum_{r\neq k,l}\left[\frac{\alpha^n_{kr}\alpha^n_{lr}}{(\sigma^n_r)^2} - \frac{\alpha^n_{kr}\alpha^n_{lr}}{(\sigma^n_r + S^n_{\max,r})^2}\right] \\
&\geq \frac{1}{(\sigma^n_{\max} + \sum_{i\neq k} \alpha^n_{\max} S^n_{\max} + S^n_{\max})^2} - \sum_{l\neq k}\left[\frac{\alpha^n_{\max}}{(\sigma^n_{\min})^2} + \frac{\alpha^n_{\max}}{(\sigma^n_{\min})^2}\right] \\
&\qquad - \sum_{l=1}^{K}\sum_{r\neq k,l}\left[\frac{(\alpha^n_{\max})^2}{(\sigma^n_{\min})^2} - \frac{(\alpha^n_{\min})^2}{(\sigma^n_{\max} + S^n_{\max})^2}\right] \\
&= \frac{1}{\left(\sigma^n_{\max} + ((K-1)\alpha^n_{\max} + 1)\, S^n_{\max}\right)^2} + \frac{(K-1)^2(\alpha^n_{\min})^2}{(\sigma^n_{\max} + S^n_{\max})^2} \\
&\qquad - 2(K-1)\frac{\alpha^n_{\max}}{(\sigma^n_{\min})^2} - (K-1)^2\frac{(\alpha^n_{\max})^2}{(\sigma^n_{\min})^2} \\
&\geq 0.
\end{aligned}$$

Thus, we have (3.2). □

From Corollary 1, we may expect that the tone-rate function R^n is more likely to be concave over Ω^n as $\alpha^n_{\max}$, $\sigma^n_{\max}$, and $S^n_{\max}$ become smaller. Moreover, by considering a sufficient condition for (3.9), we can prove that the tone-rate function becomes concave over Ω^n when all crosstalk coefficients α^n_{lk} $(l \neq k)$ are smaller than a certain threshold.

Corollary 2. *The tone-rate function R^n is concave over Ω^n if*

$$\alpha^n_{\max} \leq \frac{(\sigma^n_{\min})^2}{(K^2-1)(\sigma^n_{\max} + KS^n_{\max})^2} \tag{3.10}$$

for all $k \in \mathcal{K}$.

Proof. By Corollary 1, it suffices to show (3.9). When (3.10) holds, we can easily see

$$\alpha^n_{\max} \le 1, \tag{3.11}$$

$$\frac{1}{(\sigma^n_{\max} + KS^n_{\max})^2} - \frac{(K^2-1)\alpha^n_{\max}}{(\sigma^n_{\min})^2} \ge 0. \tag{3.12}$$

Hence, for all $k \in \mathcal{K}$, we have

$$\begin{aligned}
&\frac{1}{(\sigma^n_{\max} + ((K-1)\alpha^n_{\max}+1)\, S^n_{\max})^2} + \frac{(K-1)^2(\alpha^n_{\min})^2}{(\sigma^n_{\max} + S^n_{\max})^2} \\
&\qquad\qquad - 2(K-1)\frac{\alpha^n_{\max}}{(\sigma^n_{\min})^2} - (K-1)^2\frac{(\alpha^n_{\max})^2}{(\sigma^n_{\min})^2} \\
&\ge \frac{1}{(\sigma^n_{\max} + ((K-1)+1)\, S^n_{\max})^2} - 2(K-1)\frac{\alpha^n_{\max}}{(\sigma^n_{\min})^2} - (K-1)^2\frac{\alpha^n_{\max}}{(\sigma^n_{\min})^2} \\
&= \frac{1}{(\sigma^n_{\max} + KS^n_{\max})^2} - \frac{(K^2-1)\alpha^n_{\max}}{(\sigma^n_{\min})^2} \\
&\ge 0,
\end{aligned}$$

where the first inequality follows from (3.11) and the last inequality follows from (3.12). Thus, we have (3.9). □

From Corollary 2, we may expect that the tone-rate function R^n is more likely to be concave over Ω^n as the number of users K becomes smaller.

4 Splitting methods for the sum-rate maximization problem

In the last section, we showed that the sum-rate maximization problem (2.8) becomes concave when all the crosstalk coefficients are sufficiently small. In this section, we propose splitting algorithms which utilize the special structure of problem (2.8) under the concavity assumption.

When the tone-rate function R^n becomes concave over Ω^n, problem (2.8) is a so-called separable convex programming problem. In fact, if the power budget constraint $\sum_{n=1}^N (S^n_k)^{(\nu+1)} \le P_k$ in problem (2.8) is removed, we can decompose problem (2.8) into N independent subproblems. To make the most of such a special structure, we introduce splitting algorithms. There have been proposed many splitting algorithms such as the dual decomposition method (DDM) [1], the alternating minimization algorithm (AMA) [23], the alternating direction method of multipliers (ADMM) [3, 11], and the primal Douglas-Rachford splitting algorithm (PDRSA) [12, 14]. The DDM solves the dual of the original problem by using appropriate algorithms such as the subgradient method. However, the algorithm may be slow since the dual function is nondifferentiable in general. On the other hand, the AMA and the ADMM solve the original problem by minimizing auxiliary functions such as the Lagrangian function and the augmented Lagrangian function. The AMA and the ADMM are guaranteed to be globally convergent under the assumption that the objective function is strongly convex and convex, respectively. The PDRSA is originally designed to solve multi-valued equations involving a class of monotone mappings, but it is also applicable to separable convex programming problems.

In this paper, we adopt the ADMM and the PDRSA. First, in Subsection 4.1, we apply the ADMM to the sum-rate maximization problem (2.8). Next, in Subsection 4.2, we apply the ADMM to the dual of problem (2.8). Finally, in Subsection 4.3, we apply the PDRSA to problem (2.8). Throughout this section, we only consider the case where the sum-rate function R is concave over the feasible region $\mathcal{F}$ of problem (2.8).

4.1 Alternating direction method of multipliers (primal version)

In this subsection, we apply the ADMM to the sum-rate maximization problem (2.8). The ADMM is designed to solve a problem of the form

$$\begin{array}{ll} \text{minimize} & G_1(\boldsymbol{y}) + G_2(\boldsymbol{z}) \\ \text{subject to} & \boldsymbol{y} \in C_1,\ \boldsymbol{z} \in C_2,\ A\boldsymbol{y} = \boldsymbol{z}, \end{array} \tag{4.1}$$

where $G_1 : \mathbb{R}^m \to \mathbb{R}$ and $G_2 : \mathbb{R}^l \to \mathbb{R}$ are convex functions, $A \in \mathbb{R}^{l\times m}$ is a given matrix, and $C_1 \subseteq \mathbb{R}^m$ and $C_2 \subseteq \mathbb{R}^l$ are convex sets. For solving (4.1), the ADMM generates the sequence $\{\boldsymbol{y}^{(\nu)}, \boldsymbol{z}^{(\nu)}, \boldsymbol{\mu}^{(\nu)}\}$ by

$$\boldsymbol{y}^{(\nu+1)} := \operatorname*{argmin}_{\boldsymbol{y}\in C_1} L_c(\boldsymbol{y}, \boldsymbol{z}^{(\nu)}, \boldsymbol{\mu}^{(\nu)}), \tag{4.2}$$

$$\boldsymbol{z}^{(\nu+1)} := \operatorname*{argmin}_{\boldsymbol{z}\in C_2} L_c(\boldsymbol{y}^{(\nu+1)}, \boldsymbol{z}, \boldsymbol{\mu}^{(\nu)}), \tag{4.3}$$

$$\boldsymbol{\mu}^{(\nu+1)} := \boldsymbol{\mu}^{(\nu)} + c\left(A\boldsymbol{y}^{(\nu+1)} - \boldsymbol{z}^{(\nu+1)}\right), \tag{4.4}$$

where $L_c(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\mu}) := G_1(\boldsymbol{y}) + G_2(\boldsymbol{z}) + \boldsymbol{\mu}^\top(A\boldsymbol{y} - \boldsymbol{z}) + \frac{c}{2}\|A\boldsymbol{y} - \boldsymbol{z}\|^2$ is the augmented Lagrangian function for $G_1(\boldsymbol{y}) + G_2(\boldsymbol{z})$.

By introducing an auxiliary variable $\boldsymbol{S}^{N+1} \in \mathbb{R}^K$, we can rewrite problem (2.8) as

$$\begin{array}{ll} \underset{\boldsymbol{S}^1,\dots,\boldsymbol{S}^{N+1}}{\text{maximize}} & \displaystyle\sum_{n=1}^{N+1} R^n(\boldsymbol{S}^n) \\ \text{subject to} & \displaystyle\sum_{n=1}^{N+1} \boldsymbol{S}^n = \boldsymbol{P}, \\ & \boldsymbol{S}^n \in \Omega^n,\ n \in \mathcal{N} \cup \{N+1\}, \end{array} \tag{4.5}$$

where

$$R^{N+1}(\boldsymbol{S}^{N+1}) \equiv 0, \quad \Omega^{N+1} := \prod_{k=1}^{K} [0,\ P_k].$$

We can apply the ADMM to problem (4.5) by setting

$$\boldsymbol{y} := ((\boldsymbol{S}^1)^\top, \dots, (\boldsymbol{S}^{N+1})^\top)^\top \in \mathbb{R}^{(N+1)K},$$

$$\boldsymbol{z} := (z_1^1, \dots, z_K^1, z_1^2, \dots, z_K^{N+1})^\top \in \mathbb{R}^{(N+1)K},$$

$$G_1(\boldsymbol{y}) := -\sum_{n=1}^{N+1} R^n(\boldsymbol{S}^n),$$

$$G_2(\boldsymbol{z}) := 0,$$

$$C_1 := \prod_{n=1}^{N+1} \Omega^n, \tag{4.6}$$

$$C_2 := \left\{ \boldsymbol{z} \;\middle|\; \sum_{n=1}^{N+1} z_k^n = P_k\ ,\ k = 1, \dots, K \right\}, \tag{4.7}$$

$$A := I_{(N+1)K}. \tag{4.8}$$

Then, the ADMM for the sum-rate maximization problem (2.8) is given as follows. (For more details, see Appendix A.)

Algorithm 1.

Step 0 Choose a scalar $c > 0$, and vectors $\boldsymbol{S}^{(0)} \in \mathbb{R}^{NK}$, $(\boldsymbol{S}^{N+1})^{(0)} \in \mathbb{R}^{K}$ and $\boldsymbol{\lambda}^{(0)} \in \mathbb{R}^{K}$. Let $\nu := 0$.

Step 1 For each $n \in \mathcal{N} \cup \{N+1\}$, find the solution $(\boldsymbol{S}^n)^{(\nu+1)} \in \mathbb{R}^K$ of the problem

$$\begin{aligned} \text{minimize} \quad & -R^n(\boldsymbol{S}^n) + \sum_{k=1}^{K} \Bigg\{ \lambda_k^{(\nu)} S_k^n \\ & + \frac{c}{2}\left(S_k^n - (S_k^n)^{(\nu)} + \frac{1}{N+1}\left(\sum_{n=1}^{N+1} (S_k^n)^{(\nu)} - P_k \right) \right)^2 \Bigg\} \\ \text{subject to} \quad & \boldsymbol{S}^n \in \Omega^n. \end{aligned} \tag{4.9}$$

Step 2 For each $k \in \mathcal{K}$, put

$$\lambda_k^{(\nu+1)} := \lambda_k^{(\nu)} + \frac{c}{N+1}\left(\sum_{n=1}^{N+1} (S_k^n)^{(\nu+1)} - P_k \right). \tag{4.10}$$

Step 3 Terminate if a certain criterion is satisfied. Otherwise, set $\nu := \nu + 1$ and return to Step 1.

Note that $\lambda_k^{(\nu+1)}$ corresponds to the Lagrange multiplier for the power budget constraint $\sum_{n=1}^{N+1} (S_k^n)^{(\nu+1)} = P_k$. Note also that problem (4.9) has a unique solution and can be solved independently for each $n \in \mathcal{N} \cup \{N+1\}$. In general, it is less expensive to solve multiple lower-dimensional problems than a single higher-dimensional problem. Thus, it can be expected that the ADMM is suitable for the sum-rate maximization problem.

The ADMM for problem (4.1) is guaranteed to be globally convergent when the following assumptions hold [3, Proposition 4.2]:

(i) The optimal solution set of problem (4.1) is nonempty.

(ii) C_1 and C_2 are nonempty polyhedral sets.

(iii) Either C_1 is bounded or the matrix $A^\top A$ is invertible.

Note that, in case of problem (4.5), assumptions (i)–(iii) hold since C_1, C_2 and A are given by (4.6)–(4.8). Therefore, the global convergence of Algorithm 1 is guaranteed when the sum-rate function R is concave over the feasible region $\mathcal{F}$.

Theorem 2. *Suppose that the sum-rate function R is concave over the feasible region F of problem* (2.8)*. Then, Algorithm* 1 *is globally convergent.*

4.2 Alternating direction method of multipliers (dual version)

In this subsection, we apply the ADMM to the dual of the sum-rate maximization problem (2.8). Fukushima [11] proposed the ADMM for solving the dual of the following separable convex programming problem:

$$\begin{aligned} &\text{minimize} && \sum_{n=1}^{N} f^n(\boldsymbol{x}^n) \\ &\text{subject to} && \sum_{n=1}^{N} c_k^n(\boldsymbol{x}^n) \le 0, \quad k \in \mathcal{K}, \\ &&& \boldsymbol{x}^n \in X^n \subset \mathbb{R}^{d^n},\ n \in \mathcal{N}, \end{aligned} \tag{4.11}$$

where $f^n : \mathbb{R}^{d^n} \to \mathbb{R}$ and $c_k^n : \mathbb{R}^{d^n} \to \mathbb{R}$ are convex functions and $X^n \subseteq \mathbb{R}^{d^n}$ are nonempty closed convex sets for all $n \in \mathcal{N}$. By setting

$$\begin{aligned} \boldsymbol{x}^n &:= \boldsymbol{S}^n, \\ f^n(\boldsymbol{x}^n) &:= -R^n(\boldsymbol{S}^n), \\ c_k^n(\boldsymbol{x}^n) &:= S_k^n - \frac{P_k}{N}, \\ X^n &:= \Omega^n, \end{aligned}$$

the sum-rate maximization problem (2.8) reduces to problem (4.11).

Let $L(\boldsymbol{S}, \boldsymbol{y}) := -\sum_{n=1}^{N} R^n(\boldsymbol{S}^n) + \boldsymbol{y}^\top(\sum_{n=1}^{N} \boldsymbol{S}^n - \boldsymbol{P})$ be the Lagrangian function of problem (2.8). Then, the Lagrangian dual of problem (2.8) is given by

$$\begin{aligned} &\text{maximize} && d(\boldsymbol{y}) \\ &\text{subject to} && \boldsymbol{y} \ge 0, \end{aligned} \tag{4.12}$$

where $d : \mathbb{R}^K \to \mathbb{R}$ is the dual function defined by

$$\begin{aligned} d(\boldsymbol{y}) &:= \min\left\{ L(\boldsymbol{S}, \boldsymbol{y}) \;\middle|\; \boldsymbol{S}^n \in \Omega^n \ (n \in \mathcal{N}) \right\} \\ &= -\sum_{k=1}^{K} y_k P_k + \sum_{n=1}^{N} \min_{\boldsymbol{S}^n \in \Omega^n} \left\{ -R^n(\boldsymbol{S}^n) + \sum_{k=1}^{K} y_k S_k^n \right\}. \end{aligned}$$

Since $d(\boldsymbol{y})$ can be decomposed tone-wise, we can rewrite problem (4.12) as follows:

$$\begin{aligned} &\text{maximize} && \sum_{n=1}^{N} d^n(\boldsymbol{z}^n) \\ &\text{subject to} && \boldsymbol{y} - \boldsymbol{z}^n = 0,\ n \in \mathcal{N}, \\ &&& \boldsymbol{z}^n \ge 0,\ n \in \mathcal{N}, \end{aligned} \tag{4.13}$$

where $\boldsymbol{z}^n := (z_1^n, \dots, z_K^n)^\top \in \mathbb{R}^K$ $(n \in \mathcal{N})$ is an auxiliary variable and $d^n : \mathbb{R}^K \to \mathbb{R}$ is defined by

$$d^n(\boldsymbol{z}^n) := -\frac{1}{N} \sum_{k=1}^{K} z_k^n P_k + \min_{\boldsymbol{S}^n \in \Omega^n} \left\{ -R^n(\boldsymbol{S}^n) + \sum_{k=1}^{K} z_k^n S_k^n \right\}. \tag{4.14}$$

Thus, by letting

$$\begin{aligned}
\boldsymbol{y} &:= (y_1, \dots, y_K) \in \mathbb{R}^K, \\
\boldsymbol{z} &:= (\boldsymbol{z}^1, \dots, \boldsymbol{z}^N)^\top \in \mathbb{R}^{NK}, \\
G_1(\boldsymbol{y}) &:= 0, \\
G_2(\boldsymbol{z}) &:= -\sum_{n=1}^{N} d^n(\boldsymbol{z}^n), \\
C_1 &:= \mathbb{R}^K, \\
C_2 &:= \left\{ \boldsymbol{z} \in \mathbb{R}^{NK} \mid \boldsymbol{z}^n \geq 0,\ n \in \mathcal{N} \right\}, \\
A &:= [I, \dots, I]^\top \in \mathbb{R}^{NK \times K},
\end{aligned}$$

problem (4.13) reduces to problem (4.1).

Now, let $\boldsymbol{\mu}^n := (\mu_1^n, \dots, \mu_K^n)^\top \in \mathbb{R}^K$ be the Lagrange multiplier for the equality constraint $\boldsymbol{y} - \boldsymbol{z}^n = 0$. Then, the ADMM for the dual of the sum-rate maximization problem (2.8) is written as follows. (For more details, see Appendix B.)

Algorithm 2.

Step 0 Choose a scalar $c > 0$ and vectors $\boldsymbol{y}^{(0)} \in \mathbb{R}^K, \boldsymbol{z}^{(0)} \in \mathbb{R}^{NK}$ and $(\boldsymbol{\mu}^n)^{(0)} \in \mathbb{R}^N$ $(n \in \mathcal{N})$. Let $\nu := 0$.

Step 1 Let

$$\boldsymbol{y}^{(\nu+1)} := \frac{1}{N}\sum_{n=1}^{N} (\boldsymbol{z}^n)^{(\nu)} - \frac{1}{Nc}\sum_{n=1}^{N} (\boldsymbol{\mu}^n)^{(\nu)}. \tag{4.15}$$

Step 2 For each $n \in \mathcal{N}$, find a solution $(\boldsymbol{S}^n)^{(\nu+1)}$ of the minimization problem

$$\begin{aligned}
\underset{\boldsymbol{S}^n}{\text{minimize}} \quad & -R^n(\boldsymbol{S}^n) + \frac{c}{2}\sum_{k=1}^{K}\left[\max\left\{0, y_k^{(\nu+1)} + \frac{1}{c}\left((\mu_k^n)^{(\nu)} + S_k^n - \frac{1}{N}P_k\right)\right\}\right]^2 \\
\text{subject to} \quad & \boldsymbol{S}^n \in \Omega^n,
\end{aligned} \tag{4.16}$$

and determine $(\boldsymbol{z}^n)^{(\nu+1)}$ by

$$(\boldsymbol{z}^n)^{(\nu+1)} := \max\left\{0, \boldsymbol{y}^{(\nu+1)} + \frac{1}{c}\left((\boldsymbol{\mu}^n)^{(\nu)} + (\boldsymbol{S}^n)^{(\nu+1)} - \frac{1}{N}\boldsymbol{P}\right)\right\}.$$

Step 3 For each $n \in \mathcal{N}$, update $(\boldsymbol{\mu}^n)^{(\nu+1)}$ by

$$(\boldsymbol{\mu}^n)^{(\nu+1)} := (\boldsymbol{\mu}^n)^{(\nu)} + c\left(\boldsymbol{y}^{(\nu+1)} - (\boldsymbol{z}^n)^{(\nu+1)}\right).$$

Step 4 Terminate if a certain criterion is satisfied. Otherwise, set $\nu := \nu + 1$ and return to Step 1.

Since we consider the dual of the sum-rate maximization problem, (4.16) can be solved independently for each $n \in \mathcal{N}$. Note that, by applying the ADMM to the dual problem (4.13), we can update $\boldsymbol{z}^n$ without evaluating the dual function (4.14).

The ADMM for the dual of problem (4.11) is guaranteed to be globally convergent when the following assumptions hold [11, Theorem 1]:

(i) The solution set of problem (4.11) is nonempty and bounded.

(ii) The solution set of the dual of problem (4.11) is nonempty.

Note that assumption (i) clearly holds for problem (2.8). Also, when the sum-rate function R is concave over the feasible region $\mathcal{F}$, assumption (ii) holds for the following reasons:

- There is no duality gap and at least one optimal Lagrange multiplier for problem (2.8) since it has a feasible interior point and maximizes the concave function over the convex polyhedral set [2, Proposition 6.4.2].
- The set of optimal Lagrange multipliers for problem (2.8) coincides with the the set of dual optimal solutions since there is no duality gap [2, Proposition 6.2.3].

Therefore, we have the following theorem.

Theorem 3. *Suppose that the sum-rate function R is concave over the feasible region $\mathcal{F}$ of problem (2.8). Then, Algorithm 2 is globally convergent.*

4.3 Primal Douglas-Rachford splitting algorithm

In this subsection, we apply the PDRSA to the sum-rate maximization problem (2.8). Fukushima [12] proposed the PDRSA for the following multi-valued equation:

$$0 \in F(\boldsymbol{x}) + A^\top G(A\boldsymbol{x}), \tag{4.17}$$

where F and G are maximal monotone multi-valued mappings on R^l and R^m, respectively, and $A \in R^{m\times l}$ is a matrix such that $AA^\top$ is nonsingular. The PDRSA generates a sequence $\{\boldsymbol{z}^{(\nu)}\}$ by the iteration scheme

$$\boldsymbol{z}^{(\nu+1)} := J_F^\lambda\left(\left(2J^\lambda_{A^\top GA} - I\right)(\boldsymbol{z}^{(\nu)})\right) + \left(I - J^\lambda_{A^\top GA}\right)(\boldsymbol{z}^{(\nu)}), \tag{4.18}$$

where J_F^λ and $J^\lambda_{A^\top GA}$ denote the resolvents[5] of mappings F and $A^\top GA$, respectively.

When function R is concave over $\mathcal{F}$, the sum-rate maximization problem (2.8) reduces to problem (4.17) by setting

$$\begin{aligned} \boldsymbol{x} &:= \boldsymbol{S}, \\ F(\boldsymbol{x}) &:= -\nabla R(\boldsymbol{S}) + \partial\delta_\Omega(\boldsymbol{S}), \\ A &:= [I, \dots, I] \in \mathbb{R}^{K\times NK}, \\ G(A\boldsymbol{x}) &:= \partial\delta_W(A\boldsymbol{S}), \end{aligned}$$

where the rectangle W is given by $W := \prod_{k=1}^K [0,\, P_k]$, δ_Ω and δ_W are the indicator functions[6] of Ω and W, respectively, and $\partial\delta_\Omega$ denotes the subdifferential of function δ_Ω. Indeed, $AA^\top$ equals NI and hence is nonsingular. Moreover, multi-valued mappings $-\nabla R(\boldsymbol{S}) + \partial\delta_\Omega(\boldsymbol{S})$ and δ_W are maximal monotone since the subdifferential mapping of any closed proper convex function is maximal monotone [18]. Thus, the PDRSA applied to problem (2.8) can be written as follows. (For more details, see Appendix C.)

[5]For a maximal monotone mapping T and a positive number λ, the mapping $J_T^\lambda = (I + \lambda T)^{-1}$ is said to be the resolvent of T.

[6]For a given set $C \subseteq \mathbb{R}^n$, the indicator function $\delta_C : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is defined as follows: $\delta_C(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \in C$, and $\delta_c(\boldsymbol{x}) = +\infty$ for $\boldsymbol{x} \notin C$.

Algorithm 3.

Step 0 Choose a scalar $c > 0$ and a vector $\boldsymbol{z}^{(0)} := ((\boldsymbol{z}^1)^{(0)}, \ldots, (\boldsymbol{z}^N)^{(0)})^\top \in \mathbb{R}^{NK}$. Let $\nu := 0$.

Step 1 For each $n \in \mathcal{N}$, find the unique solution $(\boldsymbol{S}^n)^{(\nu)} \in \mathbb{R}^K$ of the problem

$$\begin{aligned} &\text{minimize} \quad -R^n(\boldsymbol{S}^n) + \frac{1}{2c}\|\boldsymbol{S}^n - (\boldsymbol{z}^n)^{(\nu)}\|^2 \\ &\text{subject to} \quad \boldsymbol{S}^n \in \Omega^n. \end{aligned} \tag{4.19}$$

Step 2 Let

$$\begin{aligned} \boldsymbol{u}^{(\nu)} &:= \sum_{n=1}^{N} \left(2(\boldsymbol{S}^n)^{(\nu)} - (\boldsymbol{z}^n)^{(\nu)} \right), \\ \boldsymbol{w}^{(\nu)} &:= \operatorname{med}\{0,\ \boldsymbol{u}^{(\nu)},\ \boldsymbol{P}\}, \\ \boldsymbol{v}^{(\nu)} &:= \frac{1}{Nc}(\boldsymbol{u}^{(\nu)} - \boldsymbol{w}^{(\nu)}). \end{aligned}$$

Step 3 For each $n \in \mathcal{N}$, let

$$(\boldsymbol{z}^n)^{(\nu+1)} := (\boldsymbol{S}^n)^{(\nu)} - c\boldsymbol{v}^{(\nu)}.$$

Terminate if a certain criterion is satisfied. Otherwise, set $\nu := \nu + 1$ and return to Step 1.

Note that problem (4.19) can be solved independently for each $n \in \mathcal{N} \cup \{N+1\}$.

The PDRSA for problem (4.17) is known to be convergent if the solution set is nonempty [14, Theorem 1]. Hence, we have the following theorem.

Theorem 4. *Suppose that the sum-rate function R is concave over the feasible region $\mathcal{F}$ of problem* (2.8)*. Then, Algorithm* 3 *is globally convergent.*

5 Numerical results

In this section, we report some numerical results. Particularly, we compare Algorithms 1–3 with the well-known iterative water-filling algorithm (IWFA) [26]. All programs were coded in Matlab 2009a and run on a machine with an Intel® Core i7 920 2.67GHz CPU and 3GB RAM.

In implementing the algorithms, we set the initial points and the termination criterion as follows. Let γ_{rand} be a random variable in [0, 1]. For IWFA, we set the initial point as $(S_k^n)^{(0)} := \gamma_{\text{rand}} P_k S^n_{\max,k} / \sum_{n=1}^N S^n_{\max,k}$ for each $n \in \mathcal{N}$ and $k \in \mathcal{K}$, and terminate if $\|\boldsymbol{S}^{(\nu)} - \boldsymbol{S}^{(\nu+1)}\| \le 10^{-4}$ or $\nu \ge 300$. For Algorithm 1, we set the initial point as $(S_k^n)^{(0)} := \gamma_{\text{rand}} P_k S^n_{\max,k} / \sum_{n=1}^N S^n_{\max,k}$, $(S_k^{N+1})^{(0)} := P_k - \sum_{n=1}^N (S_k^n)^{(0)}$ for each $n \in \mathcal{N}$ and $k \in \mathcal{K}$, and $\boldsymbol{\lambda}^{(0)} := (0.01, \ldots, 0.01)^\top$. We terminate the algorithm if $\max(\|\boldsymbol{S}^{(\nu)} - \boldsymbol{S}^{(\nu+1)}\|, \|\boldsymbol{\lambda}^{(\nu)} - \boldsymbol{\lambda}^{(\nu+1)}\|) \le 10^{-4}$ or $\nu \ge 300$. For Algorithm 2, we set the initial point as $\boldsymbol{y}^{(0)} := (0.1, \ldots, 0.1)^\top$, $\boldsymbol{z}^{(0)} := (0.1, \ldots, 0.1)^\top$, $\boldsymbol{\mu}^{(0)} := (0.1, \ldots, 0.1)^\top$, and terminate if $\max(\|\boldsymbol{y}^{(\nu)} - \boldsymbol{y}^{(\nu+1)}\|, \|\boldsymbol{z}^{(\nu)} - \boldsymbol{z}^{(\nu+1)}\|, \|\boldsymbol{\mu}^{(\nu)} - \boldsymbol{\mu}^{(\nu+1)}\|) \le 10^{-4}$ or $\nu \ge 300$. For Algorithm 3, we set the initial point as $(z_k^n)^{(0)} := \gamma_{\text{rand}} P_k S^n_{\max,k} / \sum_{n=1}^N S^n_{\max,k}$ for each $n \in \mathcal{N}$ and $k \in \mathcal{K}$ and terminate if $\|\boldsymbol{z}^{(\nu)} - \boldsymbol{z}^{(\nu+1)}\| \le 10^{-4}$ or $\nu \ge 300$.

In applying Algorithms 1–3, we first solve a problem with several choices of c, and then adopt the best one. If a solution $\bar{\boldsymbol{S}}$ generated by any of these algorithms is infeasible, then we reset $\bar{\boldsymbol{S}}$ as

$$\bar{S}_k^n := \frac{P_k}{\sum_{n=1}^N \bar{S}_k^n} \bar{S}_k^n, \ n \in \mathcal{N}, \ k \in \mathcal{K},$$

so that $\bar{\boldsymbol{S}}$ becomes feasible.

Experiment 1 In the first experiment, we generate test problems so that they satisfy the assumption of Theorem 1, and observe the performance of IWFA and Algorithms 1–3. We vary the number of tones N from 16 to 256, and generate 100 test problems for each value of N as follows.

- We consider the two-user case ($K = 2$).
- The noise coefficients σ_k^n are chosen randomly from the interval [10, 15] for each $n \in \mathcal{N}$ and $k \in \mathcal{K}$.
- The crosstalk coefficients α_{lk}^n are chosen randomly from the interval [0.1, 0.2] for each $n \in \mathcal{N}$ and $(k, l) \in \mathcal{K} \times \mathcal{K}$ with $k \neq l$.
- The power budgets P_k are chosen randomly from the interval $[N/2, N]$ for each $k \in \mathcal{K}$.
- We set the spectral masks as $S_{\max,k}^n = 2$ for each $n \in \mathcal{N}$ and $k \in \mathcal{K}$.

Tables 1 and 2 show the average sum-rate and CPU time for 100 trials, respectively. Note that the assumption of Theorem 1(b) holds for all problems. In addition, for all trials, the algorithms terminate within the maximum number of iterations.

Table 1: Average of sum-rate in Experiment 1

	$N = 16$	$N = 32$	$N = 64$	$N = 128$	$N = 256$
IWFA	1.813	3.745	7.392	15.11	29.38
ADMM(primal)	1.960	4.021	7.970	16.39	31.67
ADMM(dual)	1.960	4.021	7.970	16.39	31.67
PDRSA	1.960	4.021	7.970	16.39	31.67

Table 2: Average CPU time in Experiment 1

	$N = 16$	$N = 32$	$N = 64$	$N = 128$	$N = 256$
IWFA	0.0007	0.0012	0.0022	0.0041	0.0077
ADMM(primal)	0.4677	0.7704	1.300	2.390	5.213
ADMM(dual)	0.1682	0.3381	0.6946	1.363	2.771
PDRSA	0.1986	0.3874	0.7729	1.456	2.925

Note that, when the assumption of Theorem 1 holds, the sum-rate function is concave and Algorithms 1–3 are guaranteed to be globally convergent. In fact, from Table 1, we can see that Algorithms 1–3 yield a higher sum-rate compared to IWFA for the generated problems.

Although Table 2 shows that our algorithms are more expensive than IWFA, they solve the problems sufficiently fast. We further emphasize that, if we can use many high-end parallel computers to apply our algorithms in a distributed manner, then the problems can be solved much faster. We also note that Algorithm 2 is the fastest among Algorithms 1–3. This is probably the dual of the sum-rate maximization problem is easier to solve under the concavity assumption.

Experiment 2

In this experiment, we solve test problems that may not satisfy the assumption of Theorem 1, and observe the performance of IWFA and Algorithms 1–3. We vary the noise level β from -3 to 2, and generate 100 test problems for each value of β as follows.

- We consider the cases where $N = 32$ tones are shared by $K = 2$ users.
- The noise coefficients σ_k^n are chosen randomly from the interval $[1 \times 10^{\beta},\, 2 \times 10^{\beta}]$ for each $n \in \mathcal{N}$ and $k \in \mathcal{K}$.
- The crosstalk coefficients α_{lk}^n are chosen randomly from the interval $[0.05,\, 0.1]$ for each $n \in \mathcal{N}$ and $(k, l) \in \mathcal{K} \times \mathcal{K}$ with $k \neq l$.
- The power budgets P_k are chosen randomly from the interval $[N/2,\, N]$ for each $k \in \mathcal{K}$.
- We set the spectral masks as $S_{\max,k}^n = 2$ for each $n \in \mathcal{N}$ and $k \in \mathcal{K}$.

Tables 3 and 4 show the average sum-rate and CPU time for 100 trials, respectively. In Table 4, the numbers in parentheses show the number of times the algorithm terminated by reaching the maximum number of iterations. We note that the assumption of Theorem 1 holds for all problems with $\beta = 1$. In the case of $\beta = 0$, the assumption does not hold for all problems, but they seem to be convex since the Hessian matrices are actually negative definite for many vectors in the feasible region. When $\beta = -3, -2$ and -1, the assumption violated for all test problems, and the sum-rate function often becomes nonconcave.

Table 3: Average of sum-rate in Experiment 2

	$\beta = -3$	$\beta = -2$	$\beta = -1$	$\beta = 0$	$\beta = 1$
IWFA	169.9	156.9	97.43	26.39	3.708
ADMM(primal)	208.4	158.6	97.55	26.41	3.710
ADMM(dual)	212.6	158.7	97.57	26.41	3.711
PDRSA	215.4	158.7	97.57	26.41	3.711

Table 4: Average CPU time in Experiment 2

	$\beta = -3$	$\beta = -2$	$\beta = -1$	$\beta = 0$	$\beta = 1$
IWFA	0.0032	0.0015	0.0015	0.0015	0.0011
ADMM(primal)	2.814 (89)	2.057	0.6401	0.4912	0.7716
ADMM(dual)	2.977 (1)	1.390	0.2631	0.5989	0.2768
PDRSA	0.4836	0.7187	0.1973	0.5014	0.3435

As Table 3 shows, our algorithms generate a higher sum-rate than IWFA, especially when $\beta = -3$. However, we also observe that most of the generated test problems with $\beta = -3$ have FDMA solutions, which means that the sum-rate function is far from concave. On the other hand, in the case of $\beta = 0$ and 1, not only our algorithms but also IWFA finds near-optimal solutions. The reason seems to be that the effect of the crosstalk coefficients becomes smaller as β increases, and hence, each user optimizes his/her data rate without interfering other user's data rates.

Experiment 3

In this experiment, we consider a multiuser wireless communication system in a frequency selective environment. Let there be $N = 32$ tones shared by $K = 4$ users. We put each user's receiver on the corner of the 2-dimensional square with side length 10. We randomly generate the values of d_{kk} (the distance from transmitter k to receiver k) from the interval [0, 3] and put the transmitter randomly in the square. (See Figure 1.)

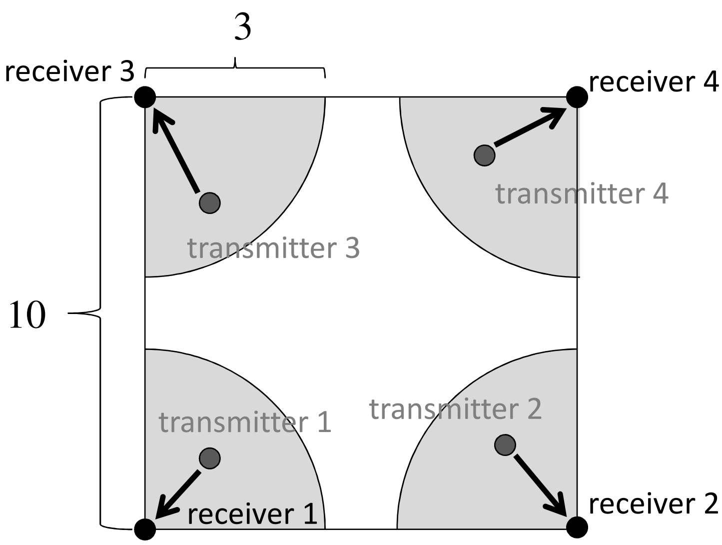

Figure 1: A wireless scenario

We let the noise level β vary from -8 to 1, and generate 100 test problems for each value of β as follows.

- We generate a channel gain coefficient by $h_{lk}^n := d_{lk}^{-1.8} g_{lk}^n$, where d_{lk} denotes the physical distance between transmitter l and receiver k, and g_{lk}^n is a complex normalized Gaussian random variable with zero mean and unit variance.
- The crosstalk coefficients are generated by $\alpha_{lk}^n := |h_{lk}^n|^2/|h_{kk}^n|^2$. (By this means, more than 95% of crosstalk coefficients become less than 0.1)
- The noise coefficients are generated by $\sigma_k^n := N_0/|h_{kk}^n|^2$, where the background noise level is set to be $N_0 = 10\beta$ dB.
- The power budget P_k are chosen randomly from the interval $[10, 13]$ dB for each $k \in \mathcal{K}$.
- We determine the spectral mask by $S_{\max,k}^n = 1$ for each $n \in \mathcal{N}$ and $k \in \mathcal{K}$.

Tables 5 and 6 show the average sum-rate and CPU time for 100 trials, respectively. Note that, for all trials, the algorithms terminate within the maximum number of iterations.

Table 5: Average of sum-rate in Experiment 3

	$\beta = -8$	$\beta = -5$	$\beta = -2$	$\beta = 1$
IWFA	800.0	815.8	423.4	55.91
ADMM(primal)	829.2	836.4	423.6	55.91
ADMM(dual)	829.5	836.4	423.6	55.91
PDRSA	829.3	836.4	423.6	55.91

Table 6: Average CPU time in Experiment 3

	$\beta = -8$	$\beta = -5$	$\beta = -2$	$\beta = 1$
IWFA	0.0064	0.0061	0.0066	0.00388
ADMM(primal)	4.719	5.402	5.344	3.950
ADMM(dual)	3.790	4.351	4.683	3.406
PDRSA	3.737	4.132	4.710	3.222

In the case of $\beta = 1$, some tone-rate functions R^n violate the assumption of Theorem 1(a). However, they seem to be nearly concave since the Hessian matrix of the sum-rate function becomes negative definite in a major part of the feasible region. Hence, we expect that our algorithms find global optima. Also, IWFA finds near-optimal solutions, since $\beta = 1$ implies that the noise level is larger than the crosstalk level. Table 6 shows that, for problems generated with $\beta = -8$, the computational costs of our algorithms are smaller than those for $\beta = -5$ and -2. Perhaps, this is because the power budget constraints are not active at the solutions.

6 Conclusion

In this paper, we have considered the sum-rate maximization problem which appears in the dynamic spectrum management for the multiuser communication systems. We provide some conditions under which the sum-rate function is guaranteed to be concave over the feasible region. In particular, we showed that the sum-rate function is concave over the feasible region when all the crosstalk coefficients are smaller than a certain value expressed by the problem parameters only. Moreover, by utilizing the concavity, we provided some splitting algorithms for the sum-rate maximization problem. From the numerical results, we observed that our algorithms could find solutions with higher sum-rates than IWFA

We still have some future issues to be addressed. First of all, the concavity conditions provided in this paper are nothing more than sufficient conditions. Hence, it is of interest to derive some weaker conditions. Second, it will be an important issue to develop new splitting algorithms more suitable for the sum-rate maximization problem.

Acknowledgments

First of all, I would like to express my sincere thanks and appreciation to Assistant Professor Shunsuke Hayashi. Although I sometimes troubled him, he always kindly looked after me, and read my poor draft manuscripts carefully. In addition, he often spared his precious time for me to discuss various issues in my study. I would also like to express my gratitude to Professor Masao Fukushima. He gave me a lot of constructive and precise advises. I would also like to tender my acknowledgement to Associate Professor Nobuo Yamashita. He gave me valuable comments from many various viewpoints. Finally, I would like to thank all members of Fukushima Laboratory, my friends and my family for their encouraging words.

References

[1] D. Bertsekas: *Nonlinear Programming: Second Edition*, Athena Scientific, 1999.

[2] D. P. Bertsekas, A. Nedić and A. E. Ozdaglar: *Convex Analysis and Optimization*, Athena Scientific, 2003.

[3] D. Bertsekas and J. Tsitsiklis: *Parallel and Distributed Computation: Numerical Methods*, Prentice Hall, 1989.

[4] S. Boyd and L. Vandenberghe: *Convex Optimization*, Cambridge University Press, 2004.

[5] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden and T. Bostoen, *Optimal multiuser spectrum balancing for digital subscriber lines*, IEEE Transactions on Communications, **54** (2006), pp. 922–933.

[6] V. M. K. Chan and W. Yu, *Joint multiuser detection and optimal spectrum balancing for digital subscriber lines*, EURASIP Journal on Applied Signal Processing, **2006** (2006), Article ID 80941, 13 pages.

[7] V. Chan and W. Yu, *Multiuser spectrum optimization for discrete multitone systems with asynchronous crosstalk*, IEEE Transactions on Signal Processing, **55** (2007), pp. 5425–5435.

[8] S. T. Chung, *Transmission schemes for frequency selective gaussian interference channels*, Doctral dissertation, Department of Electrical Engineering, Stanford University, Stanford, CA, USA, 2003.

[9] R. W. Cottle, J.-S. Pang and R. E. Stone: *The Linear Complementarity Problem*, Academic Press, 1992.

[10] T. M. Cover and J. A. Thomas: *Elements of Information Theory*, John Wiley & Sons, 1991.

[11] M. Fukushima, *Applications of the alternating direction method of multipliers to separable convex programming problems*, Computational Optimization and Applications, **1** (1992), pp. 93–111.

[12] M. Fukushima, *The primal Douglas-Rachford splitting algorithm for a class of monotone mappings with application to the traffic equilibrium problem*, Mathematical Programming, **72** (1996), pp. 1–15.

[13] S. Hayashi and Z.-Q. Luo, *Spectrum management for interference-limited multiuser communication systems*, IEEE Transactions on Information Theory, **55** (2009), pp. 1153–1175.

[14] P. L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM Journal on Numerical Analysis, **16** (1979), pp. 964–979.

[15] R. Lui and W. Yu, *Low-complexity near-optimal spectrum balancing for digital subscriber lines*, IEEE International Conference on Communications, **3** (2005), pp. 1947–1951.

[16] Z.-Q. Luo and J.-S. Pang, *Analysis of iterative waterfilling algorithm for multiuser power control in digital subscriber lines*, EURASIP Journal on Applied Signal Processing, **2006** (2006), Article ID 24012, 10 pages.

[17] Z.-Q. Luo and S. Zhang, *Dynamic spectrum management: Complexity and duality*, IEEE Journal of Selected Topics in Signal Processing, **2** (2008), pp. 57–73.

[18] R. T. Rockafellar: *Convex Analysis*, Princeton University Press, 1996.

[19] G. Scutari, D. Palomar and S. Barbarossa, *Asynchronous iterative water-filling for Gaussian frequency-selective interference channels*, IEEE Transactions on Information Theory, **54** (2008), pp. 2868–2878.

[20] G. Scutari, D. Palomar and S. Barbarossa, *Optimal linear precoding strategies for wideband noncooperative systems based on game theory-Part I: Nash equilibria*, IEEE Transactions on Signal Processing, **56** (2008), pp. 1230–1249.

[21] G. Scutari, D. Palomar and S. Barbarossa, *Optimal linear precoding strategies for wideband non-cooperative systems based on game theory-Part II: Algorithms*, IEEE Transactions on Signal Processing, **56** (2008), pp. 1250–1267.

[22] K. B. Song, S. T. Chung, G. Ginis and J. Cioffi, *Dynamic spectrum management for next-generation dsl systems*, IEEE Communications Magazine, **40** (2002), pp. 101–109.

[23] P. Tseng, *Applications of splitting algorithm to decomposition in convex programming and variational inequalities*, SIAM Journal on Control and Optimization, **29** (1991), pp. 119–138.

[24] D. Yamamoto, *Partial FDMA based approaches to sum-rate maximization problem for dynamic spectrum management*, (in Japanese), Bachelor Thesis, School of Informatics and Mathematical Science, Faculty of Engineering, Kyoto University, 2008.

[25] N. Yamashita and Z.-Q. Luo, *A nonlinear complementarity approach to multiuser power control for digital subscriber lines*, Optimization Methods and Software, **19** (2004), pp. 633–652.

[26] W. Yu, G. Ginis and J. Cioffi, *Distributed multiuser power control for digital subscriber lines*, IEEE Journal on Selected Areas in Communications, **20** (2002), pp. 1105–1115.

[27] W. Yu and R. Lui, *Dual methods for nonconvex spectrum optimization of multicarrier systems*, IEEE Transactions on Communications, **54** (2006), pp. 1310–1322.

A Derivation of Algorithm 1

Here, we describe how to derive Algorithm 1 when the ADMM is applied to the sum-rate maximization problem (2.8). By letting

$$\begin{aligned}
\boldsymbol{y} &:= ((\boldsymbol{S}^1)^\top, \dots, (\boldsymbol{S}^{N+1})^\top)^\top \in \mathbb{R}^{(N+1)K}, \\
\boldsymbol{z} &:= (z_1^1, \dots, z_K^1, z_1^2, \dots, z_K^{N+1})^\top \in \mathbb{R}^{(N+1)K}, \\
G_1(\boldsymbol{y}) &:= -\sum_{n=1}^{N+1} R^n(\boldsymbol{S}^n), \\
G_2(\boldsymbol{z}) &:= 0, \\
C_1 &:= \prod_{n=1}^{N+1} \Omega^n, \\
C_2 &:= \left\{ \boldsymbol{z} \;\middle|\; \sum_{n=1}^{N+1} z_k^n = P_k \ , \ k = 1, \dots, K \right\}, \\
A &:= I_{(N+1)K},
\end{aligned}$$

problem (4.5) reduces to problem (4.1). Moreover, the augmented Lagrangian function is given by

$$L_c(\boldsymbol{S}, \boldsymbol{z}, \boldsymbol{\mu}) = -\sum_{n=1}^{N+1} R^n(\boldsymbol{S}^n) + \boldsymbol{\mu}^\top(\boldsymbol{S} - \boldsymbol{z}) + \frac{c}{2}\|\boldsymbol{S} - \boldsymbol{z}\|^2,$$

where $\boldsymbol{\mu} = (\mu_1^1, \dots, \mu_K^1, \mu_1^2, \dots, \mu_K^{N+1})^\top \in \mathbb{R}^{(N+1)K}$ is the Lagrangian multiplier for the equality constraint $\boldsymbol{S} = \boldsymbol{z}$. Then, formulas (4.2)–(4.4) are rewritten as

$$(\boldsymbol{S}^n)^{(\nu+1)} = \operatorname*{argmin}_{\boldsymbol{S}^n \in \Omega^n} \left\{ -R^n(\boldsymbol{S}^n) + \sum_{k=1}^{K} \left((\mu_k^n)^{(\nu)} S_k^n + \frac{c}{2} \left(S_k^n - (z_k^n)^{(\nu)} \right)^2 \right) \right\}, \ n \in \mathcal{N} \cup \{N+1\}, \tag{A.1}$$

$$\boldsymbol{z}_k^{(\nu+1)} = \operatorname*{argmin}_{\boldsymbol{z}_k \in \mathbb{R}^{N+1}} \left\{ \sum_{n=1}^{N+1} \left(-(\mu_k^n)^{(\nu)} z_k^n + \frac{c}{2} \left((S_k^n)^{(\nu+1)} - z_k^n \right)^2 \right) \;\middle|\; \sum_{n=1}^{N+1} z_k^n = P_k \right\}, \ k \in \mathcal{K}, \tag{A.2}$$

$$(\mu_k^n)^{(\nu+1)} = (\mu_k^n)^{(\nu)} + c \left((S_k^n)^{(\nu+1)} - (z_k^n)^{(\nu+1)} \right), \ n \in \mathcal{N} \cup \{N+1\}, \ k \in \mathcal{K}, \tag{A.3}$$

where $\boldsymbol{z}_k := (z_k^1, \dots, z_k^{N+1})^\top \in \mathbb{R}^{N+1}$.

Moreover, (A.2) can be rewritten as

$$\begin{aligned}
\boldsymbol{z}_k^{(\nu+1)} &= \operatorname*{argmin}_{\boldsymbol{z}_k \in \mathbb{R}^{N+1}} \left\{ \frac{c}{2} \sum_{n=1}^{N+1} \left(z_k^n - (S_k^n)^{(\nu+1)} - \frac{1}{c}(\mu_k^n)^{(\nu)} \right)^2 \;\middle|\; \sum_{n=1}^{N+1} z_k^n = P_k \right\} \\
&= \operatorname*{argmin}_{\boldsymbol{z}_k \in \mathbb{R}^{N+1}} \left\{ \frac{c}{2} \left\| \boldsymbol{z}_k - \boldsymbol{S}_k^{(\nu+1)} - \frac{1}{c}\boldsymbol{\mu}_k^{(\nu)} \right\|^2 \;\middle|\; \sum_{n=1}^{N+1} z_k^n = P_k \right\},
\end{aligned} \tag{A.4}$$

where

$$\begin{aligned}
\boldsymbol{S}_k &:= (S_k^1, \dots, S_k^{N+1})^\top \in \mathbb{R}^{N+1}, \\
\boldsymbol{\mu}_k &:= (\mu_k^1, \dots, \mu_k^{N+1}) \in \mathbb{R}^{N+1}.
\end{aligned}$$

Since (A.4) implies that $\boldsymbol{z}_k^{(\nu+1)}$ is the projection of $\boldsymbol{S}_k^{(\nu+1)} + \frac{1}{c}\boldsymbol{\mu}_k^{(\nu)}$ onto the plane determined by $\sum_{n=1}^{N+1} z_k^n = P_k$ [4], it can be expressed explicitly as follows:

$$(z_k^n)^{(\nu+1)} = (S_k^n)^{(\nu+1)} + \frac{1}{c}(\mu_k^n)^{(\nu)} - \frac{1}{c}\lambda_k^{(\nu+1)},\ n \in \mathcal{N} \cup \{N+1\}, \tag{A.5}$$

$$\lambda_k^{(\nu+1)} = \frac{c}{N+1}\left\{\sum_{n=1}^{N+1}\left((S_k^n)^{(\nu+1)} + \frac{1}{c}(\mu_k^n)^{(\nu)}\right) - P_k\right\},\ k \in \mathcal{K}. \tag{A.6}$$

Substituting (A.5) into (A.3), we obtain

$$(\mu_k^n)^{(\nu+1)} = \lambda_k^{(\nu+1)},\ n \in \mathcal{N} \cup \{N+1\},\ k \in \mathcal{K}, \tag{A.7}$$

whereby (A.5) and (A.6) are rewritten as

$$(z_k^n)^{(\nu+1)} = (S_k^n)^{(\nu+1)} + \frac{1}{c}(\lambda_k^{(\nu)} - \lambda_k^{(\nu+1)}),\ n \in \mathcal{N} \cup \{N+1\},\ k \in \mathcal{K}, \tag{A.8}$$

$$\lambda_k^{(\nu+1)} = \lambda_k^{(\nu)} + \frac{c}{N+1}\left(\sum_{n=1}^{N+1}(S_k^n)^{(\nu+1)} - P_k\right),\ k \in \mathcal{K}. \tag{A.9}$$

Hence, we have (4.10). From (A.8) and (A.9), we have

$$(z_k^n)^{(\nu+1)} = (S_k^n)^{(\nu+1)} - \frac{1}{N+1}\left(\sum_{n=1}^{N+1}(S_k^n)^{(\nu+1)} - P_k\right),\ n \in \mathcal{N} \cup \{N+1\},\ k \in \mathcal{K}. \tag{A.10}$$

Hence, by substituting (A.7) and (A.10) with $\nu := \nu - 1$ into (A.1), we obtain (4.9).

B Derivation of Algorithm 2

Here, we describe how to derive Algorithm 2 when the ADMM is applied to the dual problem (4.13). By letting

$$\begin{aligned}
\boldsymbol{y} &:= (y_1, \ldots, y_K) \in \mathbb{R}^K, \\
\boldsymbol{z} &:= (\boldsymbol{z}^1, \ldots, \boldsymbol{z}^N)^\top \in \mathbb{R}^{NK}, \\
G_1(\boldsymbol{y}) &:= 0, \\
G_2(\boldsymbol{z}) &:= -\sum_{n=1}^{N} d^n(\boldsymbol{z}^n), \\
C_1 &:= \mathbb{R}^K, \\
C_2 &:= \left\{\boldsymbol{z} \in \mathbb{R}^{NK} \mid \boldsymbol{z}^n \geq 0,\ n \in \mathcal{N}\right\}, \\
A &:= [I, \ldots, I]^\top \in \mathbb{R}^{NK \times K},
\end{aligned}$$

problem (4.13) reduces to (4.1). Let $\boldsymbol{\mu}^n := (\mu_1^n, \ldots, \mu_K^n)^\top \in \mathbb{R}^K$ be the Lagrangian multiplier for the equality constraint $\boldsymbol{y} - \boldsymbol{z}^n = 0$. Then, since the augmented Lagrangian function is given by

$$L_c(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\mu}) = -\sum_{n=1}^{N} d^n(\boldsymbol{z}^n) + \sum_{n=1}^{N}\left(\boldsymbol{\mu}^{n\top}(\boldsymbol{y} - \boldsymbol{z}^n) + \frac{c}{2}\|\boldsymbol{y} - \boldsymbol{z}^n\|^2\right),$$

formulas (4.2)–(4.4) are rewritten as

$$
\begin{aligned}
\boldsymbol{y}^{(\nu+1)} &:= \operatorname*{argmin}_{\boldsymbol{y}\in\mathbb{R}^K} \left\{ \sum_{n=1}^{N} \left({(\boldsymbol{\mu}^n)^{(\nu)}}^\top \boldsymbol{y} + \frac{c}{2}\|\boldsymbol{y} - (\boldsymbol{z}^n)^{(\nu)}\|^2 \right) \right\}, && \text{(B.1)}\\
(\boldsymbol{z}^n)^{(\nu+1)} &:= \operatorname*{argmin}_{\boldsymbol{z}^n \geq 0} \left\{ -d^n(\boldsymbol{z}^n) - {(\boldsymbol{\mu}^n)^{(\nu)}}^\top \boldsymbol{z}^n + \frac{c}{2}\|\boldsymbol{y}^{(\nu+1)} - \boldsymbol{z}^n\|^2 \right\},\ n \in \mathcal{N},\\
(\boldsymbol{\mu}^n)^{(\nu+1)} &:= (\boldsymbol{\mu}^n)^{(\nu)} + c\left(\boldsymbol{y}^{(\nu+1)} - (\boldsymbol{z}^n)^{(\nu+1)} \right),\ n \in \mathcal{N}.
\end{aligned}
$$

Since $\boldsymbol{y}^{(\nu+1)}$ minimizes the convex quadratic function in (B.1), it can be expressed explicitly as (4.15). In addition, (4.16) is obtained by applying [11, Lemma 1] to (4.13) in a straightforward manner.

C Derivation of Algorithm 3

We describe how to derive Algorithm 3 when the PDRSA is applied to the sum-rate maximization problem (2.8). By letting

$$
\begin{aligned}
\boldsymbol{x} &:= \boldsymbol{S},\\
X &:= \mathbb{R}^{NK},\\
Y &:= \mathbb{R}^{K},\\
F(\boldsymbol{x}) &:= -\nabla R(\boldsymbol{S}) + \partial\delta_\Omega(\boldsymbol{S}),\\
A &:= [I, \dots, I] \in \mathbb{R}^{K\times NK}\\
W &:= \prod_{k=1}^{K} [0,\ P_k]\\
G(A\boldsymbol{x}) &:= \partial\delta_W(A\boldsymbol{S}),
\end{aligned}
$$

problem (2.8) reduces to (4.17). By applying Algorithm I in [12] directly, we have the following.

Algorithm 4.

Step 0 Choose a scalar $c > 0$ and a vector $\boldsymbol{z}^{(0)} := ((\boldsymbol{z}^1)^{(0)}, \dots, (\boldsymbol{z}^N)^{(0)})^\top \in \mathbb{R}^{NK}$. Let $\nu := 0$.

Step 1 Find the unique solution $(\boldsymbol{S})^{(\nu)} \in \mathbb{R}^{NK}$ of the equation

$$0 \in \frac{1}{c}\left(\boldsymbol{S} - \boldsymbol{z}^{(\nu)} \right) - \nabla R(\boldsymbol{S}) + \partial\delta_\Omega(\boldsymbol{S}). \tag{C.1}$$

Step 2 Put

$$\boldsymbol{u}^{(\nu)} := \sum_{n=1}^{N} \left(2(\boldsymbol{S}^n)^{(\nu)} - (\boldsymbol{z}^n)^{(\nu)} \right).$$

Find the unique solution $\boldsymbol{w}^{(\nu)} \in \mathbb{R}^K$ of the equation

$$0 \in \frac{1}{Nc}\left(\boldsymbol{w} - \boldsymbol{u}^{(\nu)} \right) + \partial\delta_W(\boldsymbol{w}) \tag{C.2}$$

and put

$$\boldsymbol{v}^{(\nu)} := \frac{1}{Nc}(\boldsymbol{u}^{(\nu)} - \boldsymbol{w}^{(\nu)}).$$

Step 3 For each $n \in \mathcal{N}$, put

$$(\boldsymbol{z}^n)^{(\nu+1)} = (\boldsymbol{S}^n)^{(\nu)} - c\boldsymbol{v}^{(\nu)}.$$

Terminate if a certain criterion is satisfied. Otherwise, set $\nu := \nu + 1$ and return to Step 1.

Since

$$\partial \left[-R(\boldsymbol{S}) + \frac{1}{2c}\|\boldsymbol{S} - (\boldsymbol{z})^{(\nu)}\|^2 + \delta_\Omega(\boldsymbol{S}) \right] = -\nabla R(\boldsymbol{S}) + \frac{1}{c}\left(\boldsymbol{S} - \boldsymbol{z}^{(\nu)}\right) + \partial\delta_\Omega(\boldsymbol{S}),$$

problem (C.1) is equivalent to the following problem:

$$\begin{array}{ll} \text{minimize} & -R(\boldsymbol{S}) + \dfrac{1}{2c}\|\boldsymbol{S} - \boldsymbol{z}^{(\nu)}\|^2 \\ \text{subject to} & \boldsymbol{S} \in \Omega, \end{array}$$

which can be decomposed into N independent problems

$$\begin{array}{ll} \text{minimize} & -R^n(\boldsymbol{S}^n) + \dfrac{1}{2c}\|\boldsymbol{S}^n - (\boldsymbol{z}^n)^{(\nu)}\|^2 \\ \text{subject to} & \boldsymbol{S}^n \in \Omega^n \end{array}$$

for $n = 1, \ldots, N$. Thus, we have (4.19). Moreover, problem (C.2) is equivalent to the following minimization problem:

$$\begin{array}{ll} \text{minimize} & \dfrac{1}{2Nc}\|\boldsymbol{w} - \boldsymbol{u}^{(\nu)}\|^2 \\ \text{subject to} & \boldsymbol{w} \in W, \end{array}$$

whose optimal solution is given by

$$\boldsymbol{w}^{(\nu)} = \text{med}\{0,\ \boldsymbol{u}^{(\nu)},\ \boldsymbol{P}\}.$$

As a result, the primal Douglas-Rachford splitting algorithm for the sum-rate maximization problem (2.8) is given by Algorithm 3.