A Method to Compute All Fixed Points of Iterative MUD of IDMA Systems

Mark S. K. Lau, Wuyi Yue, Ping Li

Memoirs of Konan University. Science and engineering series

Volume 53

Number 2

Page range 123-142

Year 2006-12-25

URL http://doi.org/10.14990/00000132

<table>
<thead>
<tr>
<th>著者（英）</th>
<th>Mark S. K. Lau, Wuyi Yue, Ping Li</th>
</tr>
</thead>
<tbody>
<tr>
<td>創刊者</td>
<td>藤間章男</td>
</tr>
<tr>
<td>創刊年月</td>
<td>1968年1月</td>
</tr>
<tr>
<td>報紙種類</td>
<td>毎月発行</td>
</tr>
<tr>
<td>本社</td>
<td>藤間章男</td>
</tr>
<tr>
<td>電子版</td>
<td>あり</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://doi.org/10.14990/00000132">http://doi.org/10.14990/00000132</a></td>
</tr>
</tbody>
</table>
A Method to Compute All Fixed Points of Iterative MUD of IDMA Systems

Mark S. K. Lau*, Wuyi Yue† and Li Ping‡

* Institute of Intelligent Information and Communications Technology
  Konan University, Kobe 658-8501 Japan
† Department of Information Science and Systems Engineering
  Konan University, Kobe 658-8501 Japan
‡ Department of Electronic Engineering, City University of Hong Kong, Hong Kong
  sklau@center.konan-u.ac.jp, yue@konan-u.ac.jp, eeliping@cityu.edu.hk

(Received October 13, 2006)

Abstract
In this paper, we propose a simple method to compute all fixed points of the iterative MUD of IDMA systems. The task of finding all fixed points has previously been difficult due to the necessity to solve a set of nonlinear equations with multiple variables. We show that this set of equations can be reduced to an equivalent equation with only one variable. As a result, all fixed points can be obtained by simply locating the roots of this univariate equation. Two applications of the proposed method in optimal power allocation are presented as well.

1 Introduction

An interleave-division multiple-access (IDMA) scheme for wireless communications has been proposed recently [1], [2], [3], [4]. In a system with IDMA, all users share the same channel and signals transmitted from different users are superimposed. Each user is assigned a unique interleaving sequence for signal separation at the receiver. The choice of spreading sequences is unimportant for signal separation, and all users can be assigned the same spreading sequence. Therefore, an IDMA is considered as a special form of the conventional code-division multiple-access (CDMA) scheme. A major advantage of IDMA over the conventional CDMA is that signal separation can be performed at a very low computation cost, which is independent of the number of users.
In systems with IDMA, signal separation is accomplished using an iterative multiuser detection (MUD) technique [2]. The MUD is performed in the receiver which consists of an elementary signal estimator (ESE) and a bank of single-user decoders. The ESE first provides coarse estimates of the transmitted signals, which are then used in the decoders to generate extrinsic information. The information is used in the ESE in the next iteration to refine the estimates. This process is repeated for a sufficiently large number of iterations.

The performance of a system can be measured by the signal-to-noise-ratio (SNR) observed at the ESE. After each iteration, the SNR for each user is increased. However, it converges as the number of iterations tends to infinity [5]. As a result, the SNR will reach a steady value after a large number of iterations. A vector of such steady values of all users is often referred to as a fixed point. A fixed point specifies the maximum achievable SNR of each user. It plays an important role in performance analysis and power allocation. In such applications, fixed points are often required to satisfy certain performance requirements so as to ensure that services of acceptable quality can be delivered to the users.

In practice, the SNR of all users are initialized to zero in the iterative detection process. Different initial values may result in different fixed points. For analysis purposes, it would be desirable if we can examine all of the fixed points. However, obtaining all fixed points is very difficult because a set of nonlinear equations with multiple variables has to be solved. Moreover, the equations involve a highly nonlinear function without explicit expression. A delicately devised numerical solver would be required to obtain all fixed points.

In this paper, we propose a simple method which guarantees to obtain all fixed points with any number of users. The method is devised based on a crucial observation that the set of equations with multiple variables can be reduced to an equivalent equation with just one variable. We call this univariate equation the characteristic equation. Moreover, we prove that fixed points can be computed from the roots of the characteristic equation. Because the roots can be located easily, the method is simple and computationally cheap.

To demonstrate the usefulness of the characteristic equation, we present two applications in optimal power allocation. In power allocation, each user is assigned a received power level such that the resulted system performance satisfies a given requirement. A vector of such power levels is called a feasible power vector. Optimal power allocation is achieved if the assigned power vector has minimum total power among all feasible power vectors. In the first application, we use the characteristic equation as a tool to study the feasibility and optimality of a simple, computationally cheap allocation method. In the second application, we propose a procedure to obtain near-optimal power vectors. The procedure is devised based on the characteristic equation.
This paper is organized as follows. In the next section, we derive some properties of fixed points. We use these properties to derive the characteristic equation. We also provide a simple example for illustration. In Section 3, we discuss two applications of the characteristic equation in optimal power allocation and give some numerical examples. We conclude this paper in Section 4.

2 Fixed Point and Characteristic Equation

Consider the SNR observed at the ESE of an IDMA system with \( n \) users. It was shown in [4] that the SNR for user \( k \) after \( g \geq 1 \) iterations is not less than \( \gamma_k^g \) where

\[
\gamma_k^g = \frac{p_k}{\sum_{j \neq k} p_j f(\gamma_j^{g-1}) + \sigma^2},
\]

\[
\gamma_k^0 = 0, \quad k = 1, 2, \ldots, n.
\]

The vector \( \mathbf{p} = (p_1, p_2, \ldots, p_n) \) is given, where \( 0 < p_k < \infty \) denotes the power of the signal of user \( k \) at the ESE. The vector \( \mathbf{p} \) is called a power vector. The channel noise is additive white Gaussian with zero mean and variance \( \sigma^2 \).

For \( 0 \leq y < \infty \), the value \( f(y) \) represents the variance of the output of a decoder driven by an input sequence with the SNR \( y \). The decoders have no outputs before the first iteration, therefore \( f(\gamma_k^0) \) is assumed to be maximal for all \( k \). This implies the initial condition in Eq. (2) since \( f(y) \) attains its maximum value at \( y = 0 \) as we will see shortly.

The function \( f \) does not have an explicit expression. In this paper, we use the experimental method of [4] to evaluate \( f \) at a large number of points. The values between two adjacent points are then estimated by linear interpolation. To facilitate analysis, we impose the following assumptions:

1. \( f \geq 0 \) on \([0, \infty)\).
2. \( f \) is continuous and strictly decreasing on \([0, \infty)\).
3. \( f(0) = 1 \) and \( f(y) \to 0 \) as \( y \to \infty \).

The function \( f \) depends on the error control code being used. Our numerical examples will use codes with rates \( 1/2, 1/4, 1/8, 1/16 \) and \( 1/32 \). For \( \mathcal{L} = 1, 2, \ldots, 5 \), the code with rate \( 1/2^\mathcal{L} \) will be a serial concatenation of a rate-1/2 convolutional code with generators \((23,35)\) and a repetition code of length \( 2^{\mathcal{L}-1} \). For ease of reference, the code with rate \( 1/2^\mathcal{L} \) will be called Code \( \mathcal{L} \). The functions \( f \) for these codes are shown in Fig. 1.
The value $\gamma_k^q = p_k / \sigma^2$ is finite, and $\gamma_k^q$ increases with $q$ because $f$ is decreasing and $\gamma_k^0 = 0$ for all $k$. Therefore, the vector $(\gamma_1^q, \gamma_2^q, \ldots, \gamma_n^q)$ converges as $q \to \infty$. The limit

$$\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n) = \lim_{q \to \infty} (\gamma_1^q, \gamma_2^q, \ldots, \gamma_n^q)$$

is a solution of the following set of $n$ equations

$$\gamma_k = \frac{p_k}{\sum_{j=1}^n p_j f(\gamma_j)} + \sigma^2, \quad k = 1, 2, \ldots, n. \quad (3)$$

A solution of this set of equations is called a fixed point for $\mathbf{p}$. Because Eq. (3) may have multiple solutions for a given $\mathbf{p}$, fixed points are not unique. Different fixed points can be obtained using the iteration of Eq. (1) with different initial conditions. Fixed points that can be realized in a real system are those associated with the initial condition in Eq. (2).

For analysis purposes, it would be desirable if we can examine all fixed points. A possible approach is to solve Eq. (3) using a numerical solver. However, the effectiveness of the solver depends heavily on the structure of the equations and the initial guesses of the solutions. The solver has to be delicately devised. Another approach is to employ Eq. (1). Because different fixed points can be obtained with different initial conditions, it is possible to identify all fixed points by exhausting all combinations of $\gamma_1^0, \gamma_2^0, \ldots, \gamma_n^0$. But this approach is extremely inefficient.
2.1 Characteristic Equation
We propose a simple method to compute fixed points. It is devised based on a crucial observation that the set of multivariate equations of Eq. (3) can be reduced to an equivalent equation with one variable.

The derivation starts with a necessary and sufficient condition for a fixed point. For the sake of convenience, we define a function as follows:

\[ g(y) = \frac{y}{1 + yf(y)}, \quad 0 \leq y < \infty. \]  \hspace{1cm} (4)

For a power vector \( p = (p_1, p_2, \ldots, p_n) \), we define

\[ I(\gamma) = \sum_{k=1}^{n} p_k f(\gamma_k) + \sigma^2 \quad \text{and} \quad A(\gamma) = \sum_{k=1}^{n} f(\gamma_k)g(\gamma_k) \]  \hspace{1cm} (5)

where \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n) \). If \( \gamma \) is a fixed point for \( p \), we can use Eq. (3) to obtain \( p_k = g(\gamma_k)I(\gamma) \). This implies

\[ f(\gamma_k)g(\gamma_k) = \frac{p_kf(\gamma_k)}{I(\gamma)} \quad \text{and hence} \quad I(\gamma) = \frac{\sigma^2}{1 - A(\gamma)}. \]

Putting the last equality into \( p_k = g(\gamma_k)I(\gamma) \), we derive

\[ p_k = \frac{\sigma^2 g(\gamma_k)}{1 - A(\gamma)}, \quad k = 1, 2, \ldots, n. \]  \hspace{1cm} (6)

Since \( g \) is strictly monotonic increasing, the inverse function \( g^{-1} \) is unique. Therefore, this equation is equivalent to

\[ p_k = \frac{\sigma^2 g(\gamma_k)}{1 - A(\gamma)} \quad \text{and} \quad \gamma_k = g^{-1}\left(\frac{p_k}{p_i} g(\gamma_i)\right) \]  \hspace{1cm} (7)

for each \( k = 2, 3, \ldots, n \). The converse is proved as follows. If Eq. (7) holds, then

\[ p_k = \frac{p_k}{p_i} p_i g(\gamma_k) = \frac{\sigma^2 g(\gamma_k)}{1 - A(\gamma)} \]

for each \( k = 1, 2, \ldots, n \). As a result,

\[ \sum_{j=k}^{n} p_j f(\gamma_j) = \sum_{j=k}^{n} \frac{\sigma^2 f(\gamma_j)}{[1 - A(\gamma)][1 + \gamma_j f(\gamma_j)]} \]

\[ = \frac{\sigma^2}{1 - A(\gamma)} \left[ A(\gamma) - \frac{\gamma_k f(\gamma_k)}{1 + \gamma_k f(\gamma_k)} \right] \]

\[ = \frac{\sigma^2}{[1 - A(\gamma)][1 + \gamma_k f(\gamma_k)]} - \sigma^2 \]

\[ = \frac{p_k}{\gamma_k} - \sigma^2. \]
We prove the following necessary and sufficient condition.

**Lemma 1.** Suppose that \( p = (p_1, p_2, \ldots, p_n) \) is a power vector. Then \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n) \) is a fixed point for \( p \) if and only if Eq. (7) holds.

This lemma suggests that, if \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n) \) is a fixed point, then \( \gamma_k \) is a function of \( \gamma_1 \) as expressed in Eq. (7) for each \( k = 2, 3, \ldots, n \). Thus Eq. (3) is equivalent to a set of equations with the variable \( \gamma_1 \) only. In the following, we show that this set of equations can be reduced to just one equation.

Again \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n) \) denotes a fixed point for the power vector \( p = (p_1, p_2, \ldots, p_n) \). Substituting Eq. (7) into the expression of \( A(\gamma) \) in Eq. (5), we obtain

\[
A(\gamma) = f(\gamma_1)g(\gamma_1) + \sum_{k=2}^{n} f\left(g^{-1}(q_k g(\gamma_1))\right)g(\gamma_1)q_k
\]

where \( q_k = \frac{p_k}{p_1} \) for \( k = 2, 3, \ldots, n \). Then

\[
p_1\left(1 - A(\gamma)\right) - \sigma^2 g(\gamma_1) = p_1 g(\gamma_1) \left[ \frac{1}{\gamma_1} - \sum_{k=2}^{n} f\left(g^{-1}(q_k g(\gamma_1))\right)q_k - \frac{\sigma^2}{p_1} \right] = 0.
\]

For a power vector \( p \) and \( 0 < y < \infty \), we define the characteristic equation as follows:

\[
C_p(y) = \frac{1}{y} - \sum_{k=2}^{n} f\left(g^{-1}(q_k g(y))\right)q_k - \frac{\sigma^2}{p_1} = 0.
\] (8)

From the above analysis, we see that \( y = \gamma_1 \) is a root of the characteristic equation. Since \( \gamma_k \) depends on \( \gamma_1 \) as recalled from Eq. (7), the behavior of fixed points is fully characterized by the roots of the equation.

**Theorem 2.** Suppose that \( p = (p_1, p_2, \ldots, p_n) \) is a power vector. Then \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n) \) is a fixed point for \( p \) if and only if \( C_p(\gamma_1) = 0 \) and \( \gamma_k = g^{-1}(p_k g(\gamma_1)/p_1) \) for each \( k = 2, 3, \ldots, n \).

This suggests that a fixed point can be obtained by simply finding a root of \( C_p = 0 \). Because \( C_p \) depends on just one variable, the root can be located easily.

**2.2 Illustrative Examples**

A system with two users is considered in these examples. In this simple system, we can obtain all fixed points using the graphical method shown in Fig. 2. The curves labeled with "A" and "B" contain, respectively, the points

\[
\left(\gamma_1, \frac{p_2}{p_1 f(\gamma_1) + \sigma^2}\right) \text{ and } \left(\gamma_2, \frac{p_1}{p_2 f(\gamma_2) + \sigma^2}\right)
\]
for $0 \leq \gamma_1, \gamma_2 \leq 3$, where $(p_1, p_2)$ is a power vector, and $\sigma^2 = 1$. The curves A and B intersect at fixed points, where

$$
\gamma_1 = \frac{p_1}{p_2 f(\gamma_2) + \sigma^2}, \quad \gamma_2 = \frac{p_2}{p_1 f(\gamma_1) + \sigma^2}.
$$

So, we can obtain fixed points from the intersection points.

For illustration purposes, we choose a power vector with many fixed points. The power vector $(2, 1.7)$ is associated with the fixed points $(0.83, 0.71)$, $(1.20, 1.09)$, $(1.79, 1.59)$ as seen from Fig. 2 (a). Theorem 2 suggests that these fixed points can also be obtained from the roots of the characteristic equation. As seen from Fig. 2 (b), the roots are $\gamma_1 = 0.83, 1.20$ and 1.79. Using Eq. (7) we obtain $\gamma_2 = 0.71, 1.09$ and 1.59, respectively. The result therefore agrees with the theorem.

![Graphs of fixed points and roots of characteristic equations](image)

Figure 2: Fixed points and roots of the characteristic equations for a two-user system.
Next we choose a power vector with only one fixed point. The power vector (2.3, 2.5) is associated with the unique fixed point (2.29, 2.47) as shown in Fig. 2 (c). Actually $\gamma_1 = 2.29$ is the unique root of the characteristic equation as shown in Fig. 2 (d). Then we use Eq. (7) to obtain $\gamma_2 = 2.47$. Again, this result agrees with the theorem.

3 Applications in Optimal Power Allocation

Two applications of the characteristic equation in optimal power allocation are discussed in this section. Optimal power allocation means to look for a power vector that has the minimum total power [4], [6]. The power vector moreover has to satisfy a given performance requirement in order to maintain the quality of services delivered to the users. It requires that the SNR at the ESE for each user is above a pre-defined level. An optimal power vector can be obtained by solving a global optimization problem, known as the power allocation problem. For a system with $n$ users, the problem can be formulated as follows:

$$\min_{p \in \mathbb{R}^n} \quad p_1 + p_2 + \cdots + p_n,$$

s.t. \quad $\gamma_k (p) \geq \Gamma, \quad k = 1, 2, \ldots, n,$

$\quad p_k \leq p_{k+1}, \quad k = 1, 2, \ldots, n - 1,$

$\quad 0 < p_k \leq \bar{p}_k, \quad k = 1, 2, \ldots, n.$

(9)

Here $p = (p_1, p_2, \ldots, p_n)$ is a vector in the Euclidean $n$-space $\mathbb{R}^n$, and $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ is a fixed point for $p$ which can be obtained using Eq. (1) with the initial condition of Eq. (2). Note that $p$ is the vector of decision variables and $\gamma_k$ varies with $p$. Moreover $0 < \bar{p}_k < \infty$ is the maximum allowable power for user $k$. Finally $0 < \Gamma$ is chosen according to the final bit-error rate (BER) of the error control code after decoding. In this paper, the values selected for Codes 1–5 are shown in Table 1, which correspond to the final BER less than $10^{-4}$. The value $10^{-4}$ is a popular choice in the literature (see [4] and [6] for instances).

| Table 1: The values of $\Gamma$ associated with Codes 1–5. |
|-----------------|---|---|---|---|---|
| Code | 1  | 2  | 3  | 4  | 5  |
| $\Gamma$ | 2.617 | 1.309 | 0.625 | 0.328 | 0.164 |

A vector lying in the feasible set of Eq. (9) is called feasible. A solution of Eq. (9) is called an optimal power vector, which has minimum total power among all
feasible power vectors. A few results on the feasibility of power vectors will be presented in this section. Nevertheless, these results are difficult to derive based on Eq. (9). A major reason is that the function \( \gamma_k(p) \) is defined through the iteration in Eq. (1) and is difficult to express explicitly. The lack of explicit expression would complicate our derivations. Fortunately, the following theorem suggests that feasible power vectors can also be characterized by the characteristic equation. The new characterization will make our derivations much easier and will be the basis of our study in this section.

**Theorem 3.** Let \( p = (p_1, p_2, \ldots, p_n) \) be a power vector satisfying the second and third constraints of Eq. (9). Then, \( p \) is feasible if and only if \( C_p(y) > 0 \) for each \( 0 < y < \Gamma \).

In proving this theorem, we will need to use the following result.

**Lemma 4.** Suppose that \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n) \) is a fixed point for a power vector. It can be obtained by the iteration in Eq. (1) with the initial condition in Eq. (2) if and only if \( \gamma_k \leq \alpha_k \) for each \( k = 1, 2, \ldots, n \) and each fixed point \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \) for the same power vector.

**Proof of Lemma 4.** Let \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n) \) be a fixed point which can be obtained with the initial condition in Eq. (2). We have seen in Section 2 that \( \gamma_k \) is the limit of the sequence \( \gamma^0_k, \gamma^1_k, \gamma^2_k, \ldots \) with \( \gamma^0_k = 0 \). Because \( \gamma^0_k \leq \alpha_k \) for all \( k \), we have

\[
\gamma^k = \frac{p_k}{\sum_{k \neq j} p_j f(\gamma^j) + \sigma^2} \leq \frac{p_k}{\sum_{k \neq j} p_j f(\alpha_j) + \sigma^2} = \alpha_k, \quad k = 1, 2, \ldots, n.
\]

Then, it can be proven by induction that \( \gamma^q_k \leq \alpha_k \) for \( q = 1, 2, \ldots \). Thus \( \gamma_k = \lim_{q \to \infty} \gamma^q_k \leq \alpha_k \). The converse is proven as follows. If \( \gamma \) is a fixed point that cannot be obtained with the initial condition in Eq. (2), then there is another fixed point, say, \( \beta = (\beta_1, \beta_2, \ldots, \beta_n) \), that can be obtained with the initial condition. From the result we have derived, we know that \( \beta_k \leq \gamma_k \) for \( k = 1, 2, \ldots, n \). Because \( \beta \neq \gamma \), we conclude that \( \beta_k < \gamma_k \) for some \( k = 1, 2, \ldots, n \).

**Proof of Theorem 3.** Assume that \( p \) is a feasible power vector. Then, the fixed point for \( p \) obtained from Eq. (1) and Eq. (2) has all entries not less than \( \Gamma \). It follows from Lemma 4 that all fixed points for \( p \) have all entries not less than \( \Gamma \). Therefore all roots of \( C_p \) are not less than \( \Gamma \). Moreover, \( C_p(y) \to \infty \) as \( y \to 0 \). Thus, we conclude that \( C_p(y) > 0 \) for each \( 0 < y < \Gamma \). The converse is proven as follows. Let \( p \) be a power vector satisfying the second and third constraint of Eq. (9) and \( C_p(y) > 0 \) for each \( 0 < y < \Gamma \). Suppose that \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n) \) is a fixed point for \( p \) obtained from Eq. (1) and Eq. (2). Recall from Theorem 2 that \( \gamma_1 \) is a root of
\[ C_p = 0 \text{ and} \]
\[ \gamma_k = g^{-1}\left(\frac{p_k}{p_i} g(\gamma_i)\right), \quad k = 2, 3, \ldots, n. \]

Because \( C_p(y) > 0 \) holds for each \( 0 < y < \Gamma \), we have \( \gamma_k \geq \gamma_i \geq \Gamma \) for each \( k = 2, 3, \ldots, n. \) This implies that \( p \) is feasible. \( \square \)

### 3.1 Application I: Equal Power Allocation

Since IDMA is a relatively new scheme, just a few methods have been proposed to tackle the power allocation problem. A linear programming method and an interior-point method have been proposed in [4] and [6], respectively, to obtain sub-optimal solutions. The former solves a simplified problem which allows just a finite number of power levels. A branch-and-bound method has been proposed in [7] which is able to obtain exact solutions for systems with small number of users. The above-mentioned methods need a lot of calculations and the computation time required may be too long for real-time applications. Therefore, at the current stage, they are only suitable for non-real-time tasks such as performance evaluation and system analysis.

Using the characteristic equation as a tool, we study an allocation method that requires very low computation cost and hence has a greater potential for real-time applications. The method, which will be called equal power allocation, simply assigns all users the same received power levels. The method does not achieve optimal allocation in many cases (it does in some cases as we will see). But optimality is sacrificed in exchange for a significantly shorter computation time. Such a trade-off may be worthwhile in some applications. In fact, equal power allocation has been frequently employed in simulation examples for the sake of convenience and simplicity [8], [9].

So far, we are not aware of any systematic study on equal power allocation of IDMA in the literature. The issues of feasibility and optimality have not been addressed. In the following, we will use the characteristic equation to derive optimality and feasibility conditions.

#### 3.1.1 Feasibility Condition

When the power levels of all users are equal, the SNR values of some users can never meet the performance requirement, no matter how large the power levels are. Under such conditions, equal power allocation is unlikely. However, we show that equal power allocation is feasible when the number of users is less than a certain threshold and the power is higher than another threshold. Explicit expressions for these thresholds are derived.

Consider a system with \( n \) users and the following defined threshold values:
\[ \bar{n} = 1 + 1 / \max_{0 \leq y \leq 1} y f(y) \quad \text{and} \quad \pi(n) = \max_{0 \leq y \leq 1} \frac{y \sigma^2}{1 - (n - 1) y f(y)}. \quad (10) \]

The following theorem states that equal power allocation is feasible when the number of users is smaller than \( \bar{n} \) and the power level is greater than \( \pi(n) \).

**Theorem 5.** An \( n \)-vector \( (p, p, \ldots, p) \) is feasible if \( \pi(n) < p \leq \bar{n}_k \) for each \( k = 1, 2, \ldots, n \) and \( n < \bar{n} \).

A few remarks on the thresholds \( \bar{n} \) and \( \pi(n) \) are given below. The proof of Theorem 5 will appear after these remarks.

1. (Code-rate and \( \bar{n} \)) Recall that \( f(y) \) represents the variance of the output of a decoder driven by an input sequence with the SNR \( y \). The function \( f \) depends on the code being employed. A code with a lower rate results in an \( f \) which decreases more rapidly. Suppose that \( f \) and \( \bar{f} \) satisfy the three assumptions in Section 2 and \( f \leq \bar{f} \) on \( [0, \infty) \). In other words, \( f(y) \) decreases more rapidly than \( \bar{f}(y) \) as \( y \) increases. In this case, \( \bar{n} \) associated with \( f \) is larger, as seen from the definition of \( \bar{n} \). Therefore, the lower the code-rate is, the larger the threshold \( \bar{n} \) becomes. Table 2 shows the values of \( \bar{n} \) for various code-rates and variance of channel noise \( \sigma^2 = 1 \). It is interesting to note that \( \bar{n} \) tends to double when the code-rate is reduced by half.

<table>
<thead>
<tr>
<th>Code-rate</th>
<th>1/2</th>
<th>1/4</th>
<th>1/8</th>
<th>1/16</th>
<th>1/32</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>( \bar{n} )</td>
<td>2.678</td>
<td>5.540</td>
<td>11.361</td>
<td>23.054</td>
<td>46.390</td>
</tr>
</tbody>
</table>

2. (Growth of \( \pi(n) \)) Let \( y^* \in [0, 1] \) be a global maximum such that

\[ \frac{y^* \sigma^2}{1 - (n - 1) y^* f(y^*)} = \max_{0 \leq y \leq 1} \frac{y \sigma^2}{1 - (n - 1) y f(y)} = \pi(n). \]

For \( n < \bar{n} - 1 \), it follows that

\[ \pi(n + 1) - \pi(n) \geq \frac{\sigma^2}{1/y^* - n f(y^*)} - \frac{\sigma^2}{1/y^* - (n - 1) f(y^*)} > 0. \]

This suggests that \( \pi(n) \) grows with \( n \) faster than a reciprocal of \( 1/y^* - (n - 1) f(y^*) \). Therefore, if the number \( n \) of users in a system increases
(a new user is admitted for instance), the power $p$ of each user has to be increased to meet the condition $p > \pi(n)$. Moreover, the increase in $p$ can be considerable when $n$ is large. This trend can be observed in the numerical examples in Fig. 3, which shows $\pi(n)$ for Codes 2–5 and each $n < \bar{n}$. Code 1 is not shown here because $\bar{n} = 2.678$ is not large enough for showing this trend.

(3) (Existence and computation of thresholds) The thresholds $\bar{n}$ and $\pi(n)$ exist and can be obtained conveniently. Because $0 < y f(y)$ for $0 < y \leq \Gamma$, the denominator of the first equation in Eq. (10) is never zero. If $n < \bar{n}$ is

![Figure 3: The values of $\pi(n)$ for various codes and $n < \bar{n}$.](image)
satisfied, then the denominator of the second equation is not zero as well. Therefore, global maxima of \( yf(y) \) and \( \frac{y\sigma^2}{1-(n-1)yf(y)} \) exist in the closed interval \([0, \Gamma]\). Because the functions depend on one-variable only, the global maxima can be found conveniently. For example, we evaluate the function on a large number of points, and then perform a local search started from the point with the smallest function value.

**Proof of Theorem 5.** The inequality \( n \leq \bar{n} \) in the theorem implies that

\[
0 < \frac{1}{y} - (n - 1)f(y) \quad \text{for each} \quad 0 \leq y \leq \Gamma.
\]

It follows that \( \pi(n) \) is positive and hence \( \pi(n) \leq p \) implies that

\[
0 < \frac{1}{y} - (n - 1)f(y) - \frac{\sigma^2}{p} \quad \text{for each} \quad 0 \leq y \leq \Gamma.
\]

Because all entries of \( p = (p, p, \ldots, p) \) are equal, the right hand side of the above inequality is equal to \( C_p \), namely,

\[
C_p(y) = \frac{1}{y} - (n - 1)f(y) - \frac{\sigma^2}{p}.
\]

Thus \( C_p(y) > 0 \) for each \( 0 \leq y \leq \Gamma \). Moreover, the second and third constraints of Eq. (9) are satisfied. We use Theorem 3 to conclude that \( p \) is feasible. \( \square \)

### 3.1.2 Optimality Condition

If the number of users is smaller than a threshold \( \bar{n}^* \), then equal power allocation is not only feasible but also optimal. To show this, we define

\[
\bar{n}^* = 1 + \inf_{0 < y < \Gamma} \frac{y^{-1} - \Gamma^{-1}}{f(y) - f(\Gamma)} \quad \text{and} \quad \pi^*(n) = \frac{\sigma^2 \Gamma}{1 - (n - 1)f(\Gamma)}.
\]  

(11)

**Theorem 6.** Let \( p^* \) be an \( n \)-vector whose entries are all equal to \( \pi^*(n) \). Assume that \( f \) is differentiable at \( \Gamma \), and \( fg \) and \( g + \Gamma fg \) are both convex on \([\Gamma, \infty)\). Then \( p^* \) is an optimal solution of Eq. (9) if \( n < \bar{n}^* \) and \( \pi^*(n) \leq \bar{p}_k \) for each \( k = 1, 2, \ldots, n \).

A few remarks are given below, which is followed by the proof of Theorem 6.

(1) (Computation of \( \bar{n}^* \)) For convenience, \( \lambda \) denotes the following function:

\[
\lambda(y) = \frac{y^{-1} - \Gamma^{-1}}{f(y) - f(\Gamma)}, \quad 0 < y < \Gamma.
\]
If $\lambda$ has a global minimum point in $(0, \Gamma)$, then

$$\bar{n}^* = 1 + \inf_{0 < y < \Gamma} \lambda(y) = 1 + \min_{0 < y < \Gamma} \lambda(y).$$

(12)

In this case, $\bar{n}^*$ can be obtained by solving a one-variable minimization problem. A simple method is to evaluate $\lambda$ at a large number of points, and to perform a local search started from the point with the smallest function value. Now, we consider the case where $\lambda$ has no global minima in $(0, \Gamma)$. In this case, the infimum is simply the following limit:

$$\inf_{0 < y < \Gamma} \lambda(y) = \lim_{y \to \Gamma} \lambda(y).$$

Using L'Hospital's Rule, we obtain

$$\bar{n}^* = 1 + \inf_{0 < y < \Gamma} \lambda(y) = 1 + \lim_{y \to \Gamma} \frac{1}{y^{-1} y^2 f'(\Gamma)} = 1 - \frac{1}{\Gamma^2 f'(\Gamma)}$$

where $f'$ denotes the derivative of $f$. The values of $\bar{n}^*$ for various codes are given in Table 3. For these codes, $\lambda$ has a global minimum point in $(0, \Gamma)$ and thus $\bar{n}^*$ in the table has been computed using Eq. (12).

<table>
<thead>
<tr>
<th>Code</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{n}^*$</td>
<td>2.1667</td>
<td>4.2556</td>
<td>8.3695</td>
<td>17.0810</td>
<td>34.2375</td>
</tr>
</tbody>
</table>

(2) (Relation of $\bar{n}^*$ and $\bar{n}$) Recall that $\bar{n}$ denotes the threshold value defined in Eq. (10). It is guaranteed that equal power allocation is feasible when $n$ falls below $\bar{n}$. Because optimal power vector is feasible, it is expected that $\bar{n}^* \leq \bar{n}$. This can be shown as follows. Let $y^*$ be an interior global maximum of $yf(y)$ over $[0, \Gamma]$, namely, $y^* \in (0, \Gamma)$ and $yf(y) \leq y^*f(y^*)$ for each $0 \leq y \leq \Gamma$. Then

$$\bar{n}^* = \inf_{0 < y < \Gamma} \frac{y^{-1} - \Gamma^{-1}}{f(y) - f(\Gamma)} \leq \frac{y^*^{-1} - \Gamma^{-1}}{f(y^*) - f(\Gamma)} \leq \frac{1}{y^*f(y^*)} = 1 / \max_{0 < y < \Gamma} yf(y) = \bar{n}.$$

(3) (Convexity of $fg$ and $g + \Gamma fg$) Because $f$ has no explicit expression, we cannot verify the convexity analytically. Nevertheless, as the functions depend on one variable only, we can establish the convexity by observing the graphs of the functions. The graphs of $fg$ for Codes 1–5 are given in
Fig. 4. We see that the function is indistinguishable from a convex function on \([\Gamma, \infty)\). Moreover, we obtain numerically that

\[
\max_{\gamma < \Lambda} |g'(y) + \Gamma f(\gamma) g(y) - y| \\
\approx 4.6 \times 10^{-4}, 5.39 \times 10^{-5}, 1.37 \times 10^{-5}, 1.90 \times 10^{-6}, 4.29 \times 10^{-7}
\]

for Codes 1, 2, 3, 4, 5, respectively, where \(\Lambda\) is a very large number. Thus, we can safely assume that \(g + \Gamma fg\) is linear on \([\Gamma, \infty)\) for these codes. As we observe, the requirement that \(fg \) and \(g + \Gamma fg\) are both convex is rather weak for practical codes, and codes with lower rate are more likely to satisfy the condition.

![Graph](image)

**Figure 4:** Convexity of \(fg\) for Codes 1–5.

(4) The value of \(\Gamma f(\Gamma)\) is usually much smaller than one. Actually \(\Gamma f(\Gamma) < 0.004\) for each of Codes 1–5. As a result, the change of \(\pi^*(n)\) with respect to \(n\) is rather mild for \(n < \bar{n}^*\). This can be observed in Fig. 5, in which \(\pi^*(n)\) is given for \(n < \bar{n}^*\) and \(\sigma^2 = 1\).

**Proof of Theorem 6.** Recall from Eq. (6) that, if \(\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)\) is a fixed point for a power vector \(p = (p_1, p_2, \ldots, p_n)\), then
Figure 5: The values of $\pi^*(n)$ for various codes and $n < \bar{n}$.

$$p_k = \frac{\sigma^2 g(\gamma_k)}{1 - A(\gamma)}, \quad k = 1, 2, \ldots, n.$$  

Therefore, the total power of $p$ can be expressed in terms of $\gamma$ as

$$\phi(\gamma) = \sum_{k=1}^{n} \frac{\sigma^2 g(\gamma_k)}{1 - A(\gamma)}.$$  

First of all, we show that $\phi$ is quasi-convex under the assumptions of the theorem. In the following, $\mathcal{R}$ denotes the rectangular region $[\Gamma, \infty)^n$. By a definition in [10],
\( \phi \) is quasi-convex on \( \mathcal{R} \) if the level set \( \{ \gamma \in \mathcal{R} : \phi(\gamma) \leq c \} \) is convex for each \( c \geq c_0 \), where \( c_0 \) is a lower bound of \( \phi \) over \( \mathcal{R} \). The bound can be chosen as \( c = n \Gamma \sigma^2 \) since \( p_k \geq \gamma_k \sigma^2 \geq \Gamma \sigma^2 \) for \( \gamma \in \mathcal{R} \). By our assumptions, \( fg \) and \( g + \Gamma fg \) are both convex. Therefore, \( \sigma^2 g + cf g \) is convex for all \( c \geq c_0 \). This implies that

\[
\sum_{k=1}^{n} \sigma^2 g(\gamma_k) + c \sum_{k=1}^{n} f(\gamma_k) g(\gamma_k) - c = \sum_{k=1}^{n} \sigma^2 g(\gamma_k) + c A(\gamma) - c
\]

is convex on \( \mathcal{R} \) for \( c \geq c_0 \) and hence the level set is convex for all \( c \geq c_0 \). Therefore \( \phi \) is quasi-convex on \( \mathcal{R} \).

Let \( \Gamma \) be an \( n \)-vector whose entries are all equal to \( \Gamma \). It is straight-forward to show that each entry of the gradient vector \( \nabla \phi(\Gamma) \) is equal to \( 1 + (n-1) \Gamma \sigma^2 f(\Gamma) \).

For \( n < \bar{n}^* \), we have

\[
n - 1 < \inf_{0 < y < \Gamma} \frac{y^{-1} - \Gamma^{-1}}{f(y) - f(\Gamma)} \leq \lim_{y \to \Gamma} \frac{y^{-1} - \Gamma^{-1}}{f(y) - f(\Gamma)} = -\frac{1}{\Gamma \sigma^2 f(\Gamma)}.\]

Therefore, \( (\gamma - \Gamma) \nabla \phi(\Gamma) \geq 0 \) for each \( \gamma \in \mathcal{R} \) when \( n < \bar{n}^* \). Because \( \phi \) is quasi-convex, an optimality condition in [10] implies that \( n \pi^*(n) = \phi(\Gamma) \leq \phi(\gamma) \) for each \( \gamma \in \mathcal{R} \). This suggests that any power vector, which has a fixed point lying in \( \mathcal{R} \), has total power larger than or equal to \( n \pi^*(n) \). The set of all such power vectors includes all feasible power vectors. Therefore, \( p^* \) is an optimal solution if \( p^* \) is feasible. From Eq. (8) we have

\[
C_{p^*}(y) = \frac{1}{y} - (n-1) f(y) - \sigma^2 \frac{1}{\pi^*(n)} = \frac{1}{y} - \frac{1}{\Gamma} + (n-1)[f(\Gamma) - f(y)].
\]

If \( n < \bar{n}^* \), then \( C_{p^*}(y) > 0 \) for each \( 0 < y < \Gamma \). If \( \pi^*(n) \leq \bar{p}_k \) for each \( k = 1, 2, \ldots, n \), we use Theorem 3 to conclude that \( p^* \) is feasible.

### 3.2 Application II: Correction Steps

The power allocation problem has been tackled in [7]. An approximating problem which can be solved effectively using a Branch-and-Bound method has been derived. The set of admissible power vectors of the approximating problem is a superset (or a relaxation) of the feasible set of Eq. (9). The authors of [7] showed that, when the approximation is sufficiently accurate, the two sets coincide and thus an optimal solution of Eq. (9) can be obtained.

To avoid high computation cost, approximation with modest accuracy should be employed. In this case, a power vector \( p' \) that solves the approximating problem may not lie in the feasible set of Eq. (9). This situation is depicted in Fig. 6 with \( \Omega \) denoting the feasible set of Eq. (9). Therefore, after obtaining \( p' \), a correction procedure (dotted-line) is needed in order to find a point \( p \in \Omega \) close to \( p' \). A possible procedure is to search over a neighborhood of \( p' \). Nevertheless, the
distance between $p'$ and $\Omega$ is usually unknown. To our best knowledge, there are no simple, assuring methods to select a suitable neighborhood that contains a point of $\Omega$.

Now we present a simple correction procedure which employs the characteristic equation. Let $p'$ be a power vector that solves the approximating problem. Moreover $p'_1$ and $\epsilon > 0$ denote the first coordinate of $p'$ and an arbitrarily small number, respectively. The procedure is as follows:

1. Obtain $C^* = \min \{ C_{p'}(y), 0 < y < \Gamma \}$.
2. Compute $a = \sigma^2/p'_1 + C^* - \epsilon$.
3. Compute $p = p'\sigma^2(p'_1a)^{-1}$.

Suppose that $p = (p_1, p_2, \ldots, p_n)$ was obtained in the last step of the procedure. Then

$$C_p(y) = \frac{1}{y} - \sum_{k=2}^{n} f\left(g^{-1}(q_k g(y))\right)q_k - \frac{\sigma^2}{p_1}$$

$$= \frac{1}{y} - \sum_{k=2}^{n} f\left(g^{-1}(q_k g(y))\right)q_k - a$$

$$= \frac{1}{y} - \sum_{k=2}^{n} f\left(g^{-1}(q_k g(y))\right)q_k - \frac{\sigma^2}{p_1} - C^* + \epsilon$$

$$= C_{p'} - \min \{ C_{p'}(y), 0 < y < \Gamma \} + \epsilon \geq \epsilon$$

for all $0 < y < \Gamma$. If moreover $p_k \leq \bar{p}_k$ for each $k = 1, 2, \ldots, n$, then we can apply Theorem 3 to conclude that $p \in \Omega$. However, the procedure cannot be applied when $a \leq 0$. In this case, the power levels of $p$ are either negative or infinite. Thus, we require that $C^* > \epsilon - \sigma^2/p'_1$ so that $a > 0$. Moreover, when $C^*$ is positive, $p'$ would be in $\Omega$ already and the procedure is unnecessary. Therefore, the procedure
is applicable whenever \(-\sigma^2/p'_r < C^r < 0\).

To illustrate this procedure, we now present an example. We consider a system which employs Code 1, \(\Gamma = 2.617\), \(n = 6\), and \(\sigma^2 = 1\). It was reported in [7] that \(p' = (2.6273, 2.6273, 8.9159, 15.5141, 41.0397, 73.0823)\) is a solution of the approximating problem for this system. The vector \(p'\) does not lie in the feasible set \(\Omega\) of Eq. (9) because \(C_{p'}(y) < 0\) for some \(0 < y < \Gamma\) as shown in Fig. 7 using dashed-line. Another method to verify \(p' \not\in \Omega\) is to examine the fixed point \(\gamma'\) obtained from Eq. (1) with the initial condition of Eq. (2). We obtain \(\gamma' = (0.1252, 0.1252, 0.5812, 0.9868, 1.8303, 3.0986)\), which has some coordinates smaller than \(\Gamma\). Thus we conclude that \(p' \not\in \Omega\).

We use the proposed procedure to obtain \(p \in \Omega\). The procedure gives \(p = (2.6917, 2.6917, 9.1344, 15.8943, 42.0455, 74.8734)\) with \(\varepsilon = 10^{-5}\). As seen from Fig. 7 (solid-line), \(C_p(y) > 0\) for all \(0 < y < \Gamma\) and hence \(p \in \Omega\). Again, we examine the fixed point \(\gamma\) obtained from Eq. (1) with the initial condition of Eq. (2). We obtain \(\gamma = (2.6831, 2.6831, 9.0765, 15.7935, 41.7789, 74.3986)\) and each coordinate of \(\gamma\) is larger than \(\Gamma\). Thus we conclude that \(p \in \Omega\).

![Figure 7: Functions \(C_p\) (solid-line) and \(C_{p'}\) (dashed-line).](image)

**Figure 7**: Functions \(C_p\) (solid-line) and \(C_{p'}\) (dashed-line).

4 Conclusions

We have proposed a simple method to compute all fixed points for the iterative multiuser detection (MUD) of interleave-division multiple-access (IDMA) systems. The characteristic equation has been derived and two applications in optimal power allocation have been discussed as well.
Acknowledgment

This work was supported in part by MEXT.ORC (2004-2008), Japan.

References


