HAL
open science

Robustness maximization of parallel multichannel systems

Jean-Yves Baudais, Fahad Syed Muhammad, Jean-François Hélard

To cite this version:

Jean-Yves Baudais, Fahad Syed Muhammad, Jean-François Hélard. Robustness maximization of parallel multichannel systems. Journal of Electrical and Computer Engineering, 2012, 2012, pp.ID 840513. 10.1155/2012/840513 . hal-00757635

HAL Id: hal-00757635

https://hal.science/hal-00757635v1

Submitted on 18 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hindawi Publishing Corporation
Journal of Electrical and Computer Engineering
Volume 2012, Article ID 840513, 16 pages
doi:10.1155/2012/840513

Research Article

Robustness Maximization of Parallel Multichannel Systems

Jean-Yves Baudais,[1] Fahad Syed Muhammad,[2] and Jean-François Hélard[2]

[1] *National Center for Scientific Research (CNRS), The Institute of Electronics and Telecommunications of Rennes (IETR), UMR 6164, 35708 Rennes, France*
[2] *Université Européenne de Bretagne, INSA, IETR, UMR 6164, 35708 Rennes, France*

Correspondence should be addressed to Jean-Yves Baudais, jean-yves.baudais@insa-rennes.fr

Received 27 February 2012; Accepted 10 May 2012

Academic Editor: Shuo Guo

Copyright © 2012 Jean-Yves Baudais et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Bit error rate (BER) minimization and SNR-gap maximization, two robustness optimization problems, are solved, under average power and bitrate constraints, according to the waterfilling policy. Under peak power constraint the solutions differ and this paper gives bit-loading solutions of both robustness optimization problems over independent parallel channels. The study is based on analytical approach, using generalized Lagrangian relaxation tool, and on greedy-type algorithm approach. Tight BER expressions are used for square and rectangular quadrature amplitude modulations. Integer bit solution of analytical continuous bitrates is performed with a new generalized secant method. The asymptotic convergence of both robustness optimizations is proved for both analytical and algorithmic approaches. We also prove that, in the conventional margin maximization problem, the equivalence between SNR-gap maximization and power minimization does not hold with peak-power limitation. Based on a defined dissimilarity measure, bit-loading solutions are compared over Rayleigh fading channel for multicarrier systems. Simulation results confirm the asymptotic convergence of both resource allocation policies. In nonasymptotic regime the resource allocation policies can be interchanged depending on the robustness measure and on the operating point of the communication system. The low computational effort leads to a good trade-off between performance and complexity.

1. Introduction

In transmitter design, a problem often encountered is resource allocation among multiple independent parallel channels. The resource can be the power, the bits or the data, and the number of channels. The resource allocation policies are performed under constraints and assumptions, and the independent parallel channels can be encountered in multitone transmission.

Independent parallel channels result from orthogonal design applied in time, frequency, or spatial domains [1]. They can either be obtained naturally or in a situation where the transmit and receive strategies are to orthogonalize multiple waveforms. The orthogonal design can also be applied in many communication scenarios when there are multiple transmit and receive dimensions. Orthogonal frequency-division multiplexing (OFDM) and digital multitone (DMT) are two successful commercial applications for wireless and wireline communications with orthogonality in the frequency domain.

To perform resource allocation, relations between various resources are needed, and one is the channel capacity. This capacity of n-independent parallel Gaussian channels is the well-known sum of the capacity of each channel

$$\mathcal{C} = \sum_{i=1}^{n} \mathcal{C}_i = \sum_{i=1}^{n} \log_2(1 + \mathrm{snr}_i). \tag{1}$$

This relation, which holds for memoryless channels, links the supremum bitrate $\mathcal{C}_i$, here expressed in bit per two dimensions, to the signal to noise ratio, snr_i, experienced by each channel or subchannel i. Any reliable and implementable system must transmit at a bitrate r_i below capacity $\mathcal{C}_i$ over each subchannel, and then the margin, or SNR-gap, γ_i is introduced to analyze such systems [2, 3]

$$\gamma_i = \frac{2^{\mathcal{C}_i} - 1}{2^{r_i} - 1}. \tag{2}$$

This SNR-gap is a convenient mechanism for analyzing systems that transmit below capacity or without Gaussian input, and

$$r_i = \log_2\left(1 + \frac{\mathrm{snr}_i}{\gamma_i}\right), \tag{3}$$

with r_i the bitrate in bits per two-dimensional symbol (bits per second per subchannel) which is also the number of bits per constellation symbol.

Resource allocation is performed using loading algorithms, and diverse criteria can be invoked to decide which portion of the available resource is allocated to each of the subchannels. From an information theory point of view, the criterion is the mutual information, and the optimal resource allocation under average power constraint was first devised in [4] for Gaussian inputs and later for non-Gaussian inputs [5]. Since the performance measure is the capacity, the SNR-gap in (3) is $\gamma_i = 1$ for all i. In other cases, γ_i is higher than 1, and (3) has been exploited into many optimal and suboptimal resource allocation policies. In fact, resource allocation is a constraint optimization problem, and generally two cases are of practical interest: rate maximization (RM) and margin maximization (MM), where the objective is the maximization of the data or bitrate, and the maximization of the system margin (or power minimization in practice), respectively [6]. The MM problem gathers all non RM problems including power minimization, margin maximization (in its strict sense), and other measures such as error probabilities or goodput. (In this paper MM abbreviation is related to the general family of non-RM problems and not only to the margin maximization problem in its strict sense. The expanded form is reserved for the margin maximization in its strict sense.) It is not necessary to study all the resource allocation strategies, and equivalence or duality can be found. Families of approaches are defined, and unified processes have been used [7–10]. The loading algorithms are also split in to two families. The first is based on greedy-type approach to iteratively distribute the discrete resources [11], and the second uses Lagrangian relaxation to solve continuous resource adaptation [12]. Both approaches have been compared in terms of performance and complexity [7, 12–14]. All these adaptive resource allocations are possible when channel state information (CSI) is known at both transmitter and receiver sides. This CSI can be perfect or imperfect, and full or partial. The effects of channel estimation error and feedback delay on the performance of adaptive modulated systems can also be considered in the resource allocation process [15–17].

In this paper we shall focus henceforth on MM problems, and the main contributions are as follows:

(i) the new resource allocation algorithm and

(ii) the comparison of the different resource allocation strategies.

It is assumed that the channel estimation is perfect, and feedback CSI delay and overhead are negligible. The considered peak-power constraint, instead of the conventional average power constraint or sum power constraint, results from power mask limitation and has been taken into account in resource allocation problem [13, 18–20]. With this peak power constraint, each channel must satisfy a power constraint. Note that the sum power constraint is historically the first considered constraint [4]. Bitrate constraint comes from communication applications or service requirements, where different flows can exist, but one of them is chosen at the beginning of the communication. In this configuration, the remaining parameter to optimize is then the SNR-gap γ_i which is also related to the error probability of the communication system.

Two similar problems of MM have the same objective, that is, to maximize the system robustness. What we call robustness in this paper is the capability of a system to maintain acceptable performance with unforeseen disturbances.

The first measure of robustness is the SNR gap, or system margin, and its maximization ensures protection against unforeseen channel impairments or noise. The system margin maximization is the maximization of the minimal SNR-gap γ_i in (3) over the n subchannels. In that case the conventional equivalence between margin maximization and power minimization in MM problems is not generally true. In this paper we show that this equivalence can nevertheless be obtained in particular configurations.

The second robustness measure is the bit error rate (BER) and its minimization can reduce the packet error rate and the data retransmissions. In transmitter design, the BER minimization can be realized using uniform bit-loading and adaptive precoding [21, 22]. Analytical studies have been performed with peak-BER or average BER (computed as arithmetic mean) approaches [15, 17]. With nonuniform bit loading, the average BER must be computed as weighted arithmetic mean, and the resource allocation has been performed using a greedy-type algorithm [23]. The first main contribution of this paper is the analytical solution of the resource allocation problem in the case of weighted arithmetic mean BER minimization.

To perform the analytical study, based on a generalized Lagrangian relaxation tool, we develop a new method for finding roots of functions. This method generalizes the secant method to better fit the function-depending weight and to speed up the search of the roots. Both robustness polices are compared using a new measure. This measure evaluates the difference of the bit distributions instead of the bitrates. We also prove that both robustness policies provide the same bit distribution in asymptotic regime, which is defined for high SNR and high bitrate regimes, and this is the second main contribution in this paper. The proof is given in the case of unconstrained modulations (i.e., continuous bitrates and analytical solution) and also for QAM constellations and greedy-type algorithms. The convergence is exemplified by simulation in multicarrier communications systems.

The organization of the paper is as follows. In Section 2, the quantities to be used throughout the paper are introduced, and the robustness optimization problem is formulated in a general way for both system margin maximization and BER minimization. The equivalences between margin maximization and power minimization are worked out.

Section 3 presents the considered expressions of accurate BER, the new measure of the bit distribution differences, and the new search method of roots of functions. The solutions of formulated problems are given in Section 4 in the form of an optimum resource allocation policy based on greedy-type algorithms. The conditions of equivalence of both margin maximization and BER minimization are given in this section. Section 5 presents the analytical solution and both greedy-type and analytical methods are compared in Section 6. This Section 6 exemplifies the application of robustness optimization to multicarrier communication systems. Finally, the paper concludes in Section 7 with the proofs of several results relegated to the appendices.

Notation. The bitrates $\{r_i\}_{i=1}^n$ are defined as a number of bits per two dimensions and they are simply given by a number of bits (undertone per constellation).

2. Problem Formulation

Consider n parallel subchannels. On the i-th subchannel, the input-output relationship is

$$Y_i = h_i S_i + W_i, \tag{4}$$

where S_i is the transmitted symbol, Y_i is the received one, and h_i the complex scalar channel gain. The complex Gaussian noise W_i is a proper complex random variable with zero-mean and variance equal to $\sigma_{W_i}^2$.

The conventional average power constraint is

$$\frac{1}{n}\sum_{i=1}^{n} E\left[|S_i|^2\right] \le P, \tag{5}$$

whereas the peak-power constraint, or power spectrum density constraint, considered in this paper is

$$E\left[|S_i|^2\right] \le P, \quad i = 1,\ldots,n. \tag{6}$$

It is convenient to use normalized unit-power symbol $\{X_i\}_{i=1}^n$ such that

$$S_i = \sqrt{p_i P} X_i, \tag{7}$$

which leads to the peak-power constraint

$$p_i \le 1, \quad i = 1,\ldots,n. \tag{8}$$

It is also convenient to introduce two other variables. The first one is the conventional SNR

$$\mathrm{snr}_i = |h_i|^2 p_i \frac{P}{\sigma_{W_i}^2} \tag{9}$$

and the second is called power spectrum density noise ratio (PSDNR)

$$\mathrm{psdnr} = \frac{1}{n}\sum_{i=1}^{n} |h_i|^2 \frac{P}{\sigma_{W_i}^2}, \tag{10}$$

which is the mean signal to noise ratio over the n subchannels if and only if $p_i = 1$ for all i. This PSDNR is the ratio between the power mask at the receiver side (the transmitted power mask through the channel) and the power spectrum density of the noise. The system performance will be given according to this parameter to point out the ability of a system to exploit the available power under peak-power constraint.

Using the previous notations, (3) becomes

$$r_i = \log_2\left(1 + \frac{|h_i|^2 p_i P}{\gamma_i \sigma_{W_i}^2}\right). \tag{11}$$

With $p_i/\gamma_i = 1$ for all i, r_i is the subchannel capacity under power constraint P. With unconstrained modulations, r_i is defined in $\mathbb{R}$, but constrained modulations are used in practice and r_i takes a finite number of nonnegative values. Noninteger number of bits per symbol can also be used with fractional bit constellations [24, 25]. In this paper, modulations defined by discrete points are used with integer number of bits per symbol. Typically, $r_i \in \{0, \beta, 2\beta, \ldots, r_{\max}\}$, where β is the granularity in bits and $r_{\max}$ is the number of bits in the richest available constellation. The peak-power and bitrate constraints are then

$$p_i \le 1 \quad \forall i, \quad \sum_{i=1}^{n} r_i = R, \quad r_i \in \{0, \beta, 2\beta, \ldots, r_{\max}\} \ \forall i. \tag{12}$$

Obviously, the exploitation of the available power leads to $p_i = 1$ for all i and the constraint is simplified as

$$\sum_{i=1}^{n} r_i = R, \quad r_i \in \{0, \beta, 2\beta, \ldots, r_{\max}\} \ \forall i. \tag{13}$$

With peak-power and bitrate constraints, the resource allocation strategy is then to use all available power and to optimize the robustness.

The problem we pose is to determine the optimal bitrate allocation $\{r_i^*\}_{i=1}^n$ that maximizes a robustness measure, or inversely minimizes a frailness measure, under constraints given in (13). In its general form, this problem can be written as

$$[r_1^*, \ldots, r_n^*] = \arg \min_{\substack{\sum_{i=1}^n r_i = R \\ r_i \in \{0,\beta,2\beta,\ldots,r_{\max}\}}} \phi\left(\{r_i\}_{i=1}^n\right), \tag{14}$$

where $\phi(\cdot)$ is the frailness measure. In this paper, this measure is given by the SNR gap or the BER. In addition to the bitrate allocation, the receiver is presumed to have knowledge of the magnitude and phase of the channel gain $\{h_i\}_{i=1}^n$, whereas the transmitter needs only to know the magnitude $\{|h_i|\}_{i=1}^n$. The objective is to find the data vector $[r_1^*, \ldots, r_n^*]$ which is the final relevant information for the transmitter. The resource allocation can then be computed on the receiver side to reduce the feedback data rate from n real numbers to n finite integer numbers. Furthermore, the integer nature of the data rates allows a full CSI at the transmitter, which is not possible with real numbers.

2.1. System Margin Maximization. The SNR-gap γ_i of the subchannel i is (3)

$$\gamma_i = \frac{\mathrm{snr}_i}{2^{r_i} - 1}. \tag{15}$$

With reliable communications, γ_i is higher than 1 for all subchannels. Let the system margin, or system SNR-gap, be the minimal value of the SNR gap in each subchannel

$$\gamma = \min_i \gamma_i. \tag{16}$$

Let γ_{init} be the initial system margin of one communication system ensuring a given QoS. Let γ be the optimized system margin of this system. Then, the system margin improvement ensures system protection in unforeseen channel impairment or noise, for example, impulse noise; bitrate and system performance targets are always reached for an unforeseen SNR reduction of $\gamma/\gamma_{\text{init}}$ over all subchannels. This robustness optimization does not depend on constellation and channel-coding types. The system margin γ is defined and optimized without knowledge of used constellations and coding, and the proposed robustness optimization works for any coding and modulation scheme.

The objective is the maximization of the system margin which is equivalent to the minimization of γ^{-1}. We note $\gamma_i(r_i)$ the function that associates r_i to γ_i. The function $\phi(\cdot)$ in (14) is then given by

$$\phi(\{r_i\}_{i=1}^n) = \max_i \frac{1}{\gamma_i(r_i)}, \tag{17}$$

$$[r_1^*, \ldots, r_n^*] = \arg \min_{\substack{\sum_{i=1}^n r_i = R \\ r_i \in \{0,\beta,2\beta,\cdots,r_{\max}\}}} \max_i \gamma_i^{-1}. \tag{18}$$

This problem is the inverse problem of bitrate maximization under peak-power and SNR-gap constraints. The solution of the bitrate maximization problem is obvious under the said constraints and given by

$$r_i^* = \beta \left\lfloor \frac{1}{\beta} \log_2\left(1 + \frac{\text{snr}_i}{\gamma_i}\right) \right\rfloor \quad \forall i. \tag{19}$$

Following the conventional SNR-gap approximation [2], the symbol error rate (SER) of QAM depending on the SNR-gap is constellation size independent with

$$\text{ser}_i(r_i) = 2\,\text{erfc}\left(\sqrt{\frac{3}{2}\gamma_i}\right) \quad \forall r_i, \tag{20}$$

where the complementary error function is usually defined as

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt. \tag{21}$$

The system margin maximization is then equivalent to the peak-SER minimization in high-SNR regime. Note that, with (16), the system margin maximization can also be called a trough-SNR-gap maximization, and it is strongly related to the peak-power minimization. Whereas the bit-loading solution is the same for power minimization and margin maximization with sum-margin or sum-power constraints, instead of peak constraints, the following lemma gives sufficient conditions for equivalence in the case of peak constraints.

Lemma 1. *The bit allocation that maximizes the system margin under peak-power constraint* $\{p_i^{margin}\}_{i=1}^n$ *minimizes the peak-power under SNR-gap constraint* $\{\gamma_i^{power}\}_{i=1}^n$ *if* p_i^{margin} $\gamma_i^{power} = \alpha$ *for all i.*

Proof. It is straightforward using (11) and (18). Both problems have the same expression and therefore the same solution. □

This lemma provides a sufficient but not necessary condition for the equivalence of solutions, and it says that if the power and the SNR-gap constraints have proportional distributions for margin maximization and peak-power minimization problems, respectively, then both problems have the same optimal bitrate allocation. In the general case, we cannot conclude that both problems have the same solution.

2.2. BER Minimization. In communication systems, the error rate of the transmitted bits is a conventional robustness measure. By definition, the BER is the ratio between the number of wrong bits and the number of transmitted bits. With a multidimensional system, there exists several BER expressions [15, 23]. Let the BER be evaluated over the transmission of m multidimensional symbols. (We suppose that m is high enough to respect the ergodic condition and to make possible use of error probability.) In our case, the multidimensional symbols are the symbols sent over n subchannels. Let e_i be the number of erroneous bits received over subchannel i during the transmission. The BER is then given as

$$\text{ber} = \frac{\sum_{i=1}^n e_i}{m \sum_{i=1}^n r_i} = \frac{\sum_{i=1}^n r_i (e_i/m r_i)}{\sum_{i=1}^n r_i}. \tag{22}$$

The BER over subchannel i is e_i/mr_i, and the BER of n subchannels is then

$$\text{ber}(\{r_i\}_{i=1}^n) = \frac{\sum_{i=1}^n r_i \text{ber}_i(r_i)}{R} \tag{23}$$

with $\text{ber}_i(r_i)$ the function that associates the BER of channel i with the bitrate r_i. The BER of multiple variable bitrate r_i is then not the arithmetic mean of BER but is the weighted mean BER. Weighted mean BER and arithmetic mean BER are equal if $r_i = r_j$ for all i, j or if $\text{ber}_i = 0$ for all i. As there exists $\text{ber}_i \neq 0$, then weighted mean BER and arithmetic mean BER are equal if and only if $r_i = r_j$ for all i, j. Note that if the number m of transmitted multidimensional symbols depends on the subchannel i, (23) does not hold anymore. These obvious results on mean measures are not taken into account, and mean BER is erroneously used instead of mean weighted BER [15, 17].

The function $\phi(\cdot)$ in (14) is then given by

$$\phi(\{r_i\}_{i=1}^n) = \frac{1}{R} \sum_{i=1}^n r_i \text{ber}_i(r_i), \tag{24}$$

$$[r_1^*, \ldots, r_n^*] = \arg \min_{\substack{\sum_{i=1}^n r_i = R \\ r_i \in \{0,\beta,2\beta,\ldots,r_{\max}\}}} \text{ber}(\{r_i\}_{i=1}^n). \tag{25}$$

To simplify the notations, let $\mathrm{ber}(R)$ be the BER of the system. In high SNR regime with Gray mapping, $r_i \mathrm{ber}_i(r_i) = \mathrm{ser}_i(r_i)$, and then weighted mean BER can be approximated by arithmetic mean SER divided by the number of transmitted bits.

Contrary to system margin maximization, the BER minimization needs the knowledge of constellation and coding schemes, and it is based on accurate expressions of BER functions. In this paper, the used constellations are QAM, and the optimization is performed without a channel coding scheme. When dealing with practical coded systems, the ultimate measure is the coded BER and not the uncoded BER. However, the coded BER is strongly related to the uncoded BER. It is then generally sufficient to focus on the uncoded BER when optimizing the uncoded part of a communication system [26].

3. Interludes

Before solving the optimization problem, the BER approximation of QAM is presented. This approximation plays a chief role in BER minimization, and a good approximation is therefore needed. Since this paper deals with bitrate allocation, a measure of difference in the bitrate distribution is proposed and presented in this section. This section also presents a new research method of roots of functions. This method generalizes the secant method and converges faster than the secant one.

3.1. BER Approximation. Conventionally, the BER approximation of square QAM has been performed by either calculating the symbol error probability or by simply estimating it using lower and upper bounds [27]. This conventional approximation tends to deviate from the exact values when the SNR is low and cannot be applied for rectangular QAM. Exact and general closed-form expressions are developed in [28] for arbitrary one and two-dimensional amplitude modulation schemes.

An approximate BER expression for QAM can be obtained by neglecting the higher-order terms in the exact closed-form expression [28].

$$\mathrm{ber}_i \simeq \frac{1}{r_i}\left(2 - \frac{1}{I_i} - \frac{1}{J_i}\right) \mathrm{erfc}\left(\sqrt{\frac{3}{I_i^2 + J_i^2 - 2}\ \mathrm{snr}_i}\right) \tag{26}$$

with $I_i = 2^{\lfloor r_i/2 \rfloor}$, $J_i = 2^{\lceil r_i/2 \rceil}$, and $r_i = \log_2(I_i \cdot J_i)$. By symmetry, I_i and J_i can be inverted. The BER can also be expressed using the SNR-gap γ_i. Using (3) and (26), the BER is written as

$$\mathrm{ber}_i \simeq \frac{1}{r_i}\left(2 - \frac{1}{I_i} - \frac{1}{J_i}\right) \mathrm{erfc}\left(\sqrt{\frac{3(I_i J_i - 1)}{I_i^2 + J_i^2 - 2}\ \gamma_i}\right). \tag{27}$$

These two approximations allow the extension of the $\mathrm{ber}_i(r_i)$ function from $\mathbb{N}$ to $\mathbb{R}_+$ which is useful for analytical studies. Figure 1 gives the theoretical BER curves and the approximated ones from the binary phase shift keying (BPSK) to the 32768-QAM. For BER lower than $5 \cdot 10^{-2}$, the relative error is lower than 1% for all modulations.

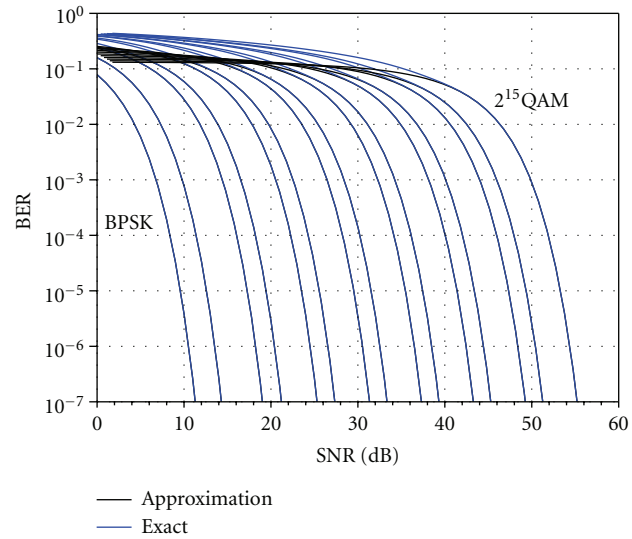

Figure 1: Exact BER curves and approximations (26).

3.2. Dissimilar Resource Allocation Measure. Two resource allocations can have the same bitrate, but this does not mean that the bitrates per subchannel are the same. To measure the difference in the bit distribution between different resource allocation strategies, we need to evaluate the *dissimilarity*. This dissimilarity measure must verify the following properties: (1) if two resource allocations lead to the same bit distribution, then the measure of dissimilarity must be null, whereas (2) if two resource allocations lead to two completely different bit distributions in loaded subchannels, then the measure of dissimilarity must be equal to one, and (3) the measure is symmetric; that is, the dissimilarity between the resource allocations X and Y must be the same as the dissimilarity between the resource allocations Y and X. We choose that the empty subchannels do not impact the measure.

Definition 2. The dissimilarity measure between the resource allocations X and Y is

$$\mu(X,Y) = \frac{\sum_{i=1}^{n} \delta(r_i(X) - r_i(Y))}{\max_{j\in\{X,Y\}} \sum_{i=1}^{n} \delta(r_i(j))}, \tag{28}$$

where $\delta(x) = 1$ if $x \neq 0$ else $\delta(x) = 0$.

This dissimilarity has the following properties.

Property 1. $\mu(X,Y) = 0$ iff $r_i(X) = r_i(Y)$ for all i.

Property 2. $\mu(X,Y) = 1$ iff $r_i(X) \neq r_i(Y)$ or $r_i(X) = r_i(Y) = 0$ for all i.

Property 3. $\mu(X,Y) = \mu(Y,X)$.

Property 4. If $\mu(X,Y) = 0$, then, for all resource allocation Z, $\mu(X,Z) = \mu(Y,Z)$.

All these properties are direct consequences of Definition 2. For a null dissimilarity, $\mu(X,Y) = 0$, all

the subchannels transmit the same number of bits, that is, $r_i(X) = r_i(Y)$ for all i. For a full dissimilarity, $\mu(X,Y) = 1$, all the nonempty subchannels of both resource allocations X and Y transmit a different number of bits, that is, for all i such as $r_i(X) \neq 0$ and $r_i(Y) \neq 0$, then $r_i(X) \neq r_i(Y)$. It is obvious that the measure is symmetric $\mu(X,Y) = \mu(Y,X)$. If two resource allocations have a null dissimilarity $\mu(X,Y) = 0$, then they are identical and for any resource allocation Z $\mu(X,Z) = \mu(Y,Z)$. The converse of this last property is not true. Note that the dissimilarity is not defined for two empty resource allocations.

For example, let $n = 4$ and $[r_1(X),\ldots,r_4(X)] = [4\ 3\ 3\ 0]$. If $[r_1(Y),\ldots,r_4(Y)] = [3\ 2\ 2\ 2]$ or $[r_1(Y),\ldots,r_4(Y)] = [5\ 5\ 0\ 0]$, then $\mu(X,Y) = 1$. If $[r_1(Y),\ldots,r_4(Y)] = [4\ 3\ 2\ 1]$, then $\mu(X,Y) = 1/2$. The measure $\mu(X,Y)$ is null if and only if $[r_1(Y),\ldots,r_4(Y)] = [4\ 3\ 3\ 0]$. The dissimilarity does not evaluate the total bitrate differences but only the bit distribution differences; the contribution of two bitrates $r_i(X)$ and $r_j(Y)$ in the dissimilarity measure is independent of the bitrate difference $|r_i(X) - r_j(Y)|$.

3.3. Generalized Secant Method. There are many numerical methods for finding roots of functions. We propose a new method, called the generalized secant method, that is, based on the secant method. This new method better fits the function-depending weight than secant method do and then improves the speed of the convergence. Before explaining this new method, a brief overview of the secant method is given.

In our case, the objective function $f(x)$ is monotonous, nondifferentiable and computable over $x \in [x_1, x_2]$ with $f(x_1)/|f(x_1)| = -f(x_2)/|f(x_2)|$. The secant method is as follows for an increasing function $f(x)$:

(1) $i = 0$, $y_0 = f(x_1)$;

(2) $x_0 = (x_2 f(x_1) - x_1 f(x_2))/(f(x_1) - f(x_2))$, $y_{i+1} = f(x_0)$;

(3) if $|y_{i+1} - y_i| \leq \epsilon$, then x_0 is the root of $f(x)$, else $\left\{\begin{matrix} y_{i+1}<0 \text{ then } x_1=x_0 \\ y_{i+1}>0 \text{ then } x_2=x_0 \end{matrix}\right\}$, $i \to i+1$ and go to step 2.

The objective of the secant method is to approximate $f(x)$ by a linear function $g_i(x) = a_i x + b_i$ at each iteration i, with $g_i(x_1) = f(x_1)$ and $g_i(x_2) = f(x_2)$, and to set x_0 as the root of $g_i(x)$. The search for the root of $f(x)$ is completed when the desired precision ϵ is reached. The precision is given for y_i, but it can also be given for x_i.

As the function $f(x)$ is computable, it can be plotted and an *a posteriori* simple algebraic or elementary transcendental invertible function over $[x_1, x_2]$ can be used to better fit the function $f(x)$. The *a posteriori* information is then used to improve the search for the root. The function $f(x)$ is iteratively approximated by $a_i h(x) + b_i$ instead of $a_i x + b_i$, where $h(x)$ is the invertible function. This method is then given as follows for an increasing function $f(x)$:

(1) $i = 0$, $y_0 = f(x_1)$;

(2) $x_0 = h^{-1}((x_2 f(x_1) - x_1 f(x_2))/(f(x_1) - f(x_2)))$, $y_{i+1} = f(x_0)$;

(3) if $|y_{i+1} - y_i| \leq \epsilon$, then x_0 is the root of $f(x)$, else $\left\{\begin{matrix} y_{i+1}<0 \text{ then } x_1=x_0 \\ y_{i+1}>0 \text{ then } x_2=x_0 \end{matrix}\right\}$, $i \to i+1$ and go to step 2.

Compared to the secant method, only step 2 differs and the computation of x_0 is performed taking into account the approximated shape $h(x)$ of the function $f(x)$.

This generalized secant method is used in Section 5 to find the root of the Lagrangian and is compared to the conventional secant method. In our case, $f(x)$ is the sum of logarithmic functions, and the function $h(x)$ is then the logarithmic one.

4. Optimal Greedy-Type Resource Allocations

The general problem is to find the optimal resource allocation $[r_1^*,\ldots,r_n^*]$ that minimizes $\phi(\cdot)$, the inverse robustness measure, or frailness. This is a combinatorial optimization problem or integer programming problem. The core idea in this iterative resource allocation is that a sequential approach can lead to a globally optimum discrete loading. Greedy-type methods then converge to the optimal solution. Convexity is not required for the convergence of the algorithm and monotonicity is sufficient [29]. This monotonicity ensures that the removal or addition of β bits at each iteration converges to the optimal solution. In this paper the used functions $\phi(\cdot)$ are monotonic increasing functions.

In its general form and when the objective function $\phi(\cdot)$ is not only a weighted sum function, the iterative algorithm is as follows:

(1) start with allocation $[r_1^{(0)},\ldots,r_n^{(0)}] = 0$,

(2) $k = 0$,

(3) allocate one more bit to the subchannel j for which

$$\phi\left(\left\{r_i^{(k+1)}\right\}_{i=1}^{n}\right) \tag{29}$$

is minimal, with $r_j^{(k+1)} = r_j^{(k)} + \beta$ and $r_i^{(k+1)} = r_i^{(k)}$ for all $i \neq j$,

(4) if $\sum_i r_i^{(k+1)} = R$, terminate; otherwise $k \to k+1$ and go to step 3.

The obtained resource allocation is then optimal [29] and solves (14). This algorithm needs R/β iterations. The target bitrate R is supposed to be feasible; that is, R is a multiple of β. Note that an equivalent formulation can be given starting with $r_i^{(0)} = r_{\max}$ for all i and using bit removal instead of bit addition with maximization instead of minimization. For bitrates higher than $(n/2)r_{\max}$, the number of iterations with bit removal is lower than with bit addition. The opposite is true with bitrate lower than $(n/2)r_{\max}$.

Iterative resource allocations have been firstly applied to bitrate maximization under power constraint [11]. Many works have been devoted to complexity reduction of greedy-type algorithms; see, for example, [6, 12, 30, 31] and references therein. In this section, only greedy-type algorithms are presented in order to compare the analytical resource allocation to the optimal iterative one. Note that the analytical solution can also be used as an input of the greedy-type algorithm to initialize the algorithm and to reduce the number of iterations.

4.1. System Margin Maximization. The system margin, or system SNR gap, maximization under bitrate and peak-power constraints is the inverse problem of the bitrate maximization under SNR-gap and peak-power constraints. This inverse problem has been solved, for example, in [18]. To comply with the general problem formulation, the inverse system margin minimization is presented instead of the system margin maximization.

Lemma 3. *Under bitrate and peak-power constraints, the greedy-type resource allocation that minimizes the inverse system margin* γ^{-1} (16) *allocates sequentially* β *bits to the subchannel i bearing* r_i *bits and for which*

$$\frac{2^{r_i+\beta}-1}{\mathrm{snr}_i} \tag{30}$$

is minimum.

Proof. It is straightforward using (17) and (29). See Appendix A for an original proof. □

The main advantage of system margin maximization is that the optimal resource allocation can be reached independently of the SNR regime. Resource allocation is always possible even for very low SNR, but it can lead to unreliable communication with SNR gap lower than 1. Lemma 3 is given with unbounded modulation orders, that is, $r_{\max} = \infty$ and $r_i \in \beta\mathbb{N}$ for all i. With full constraints (13), the subchannels that reach $r_{\max}$ are simply removed from the iterative process.

4.2. BER Minimization. The system BER minimization under bitrate and peak-power constraints is the inverse problem of bitrate maximization under peak-power and BER constraints. This inverse problem has been solved, for example, in [23]. Using (29) and (24), the solution of BER minimization is straightforward, and the corresponding greedy-type algorithm is also known as Levin-Campello algorithm [5, 32, 33]. The main drawback of this solution is that it requires good approximated BER expressions even in low-SNR regime. This constraint can be relaxed, and the following lemma gives the optimal greedy-type resource allocation for the BER minimization.

Lemma 4. *In high SNR regime and under bitrate and peak-power constraints, the greedy-type resource allocation that minimizes the BER minimizes* $(r_i+\beta)\mathrm{ber}_i(r_i+\beta)$ *at each step.*

Proof. See Appendix B. □

Lemma 4 states how to allocate bits without mean BER computation at each step. It is given without modulation order limitation. Like system margin maximization solution, the bounded modulation order is simply taken into account using $r_{\max}$ and subchannel removal.

4.3. Comparison of Resource Allocations. To compare the two optimization policies, we call $\mathcal{B}$ the resource allocation that maximizes the system margin and $\mathcal{C}$ the resource allocation that minimizes the BER. Table 1 gives an example of bitrate allocation over 20 subchannels where the SNR follows a Rayleigh distribution and with $\beta = 1$. In this example, the PSDNR defined in (10) is equal to 25 dB, and the maximum allowed bitrate per subchannel is never reached. As expected, the system margin minimization leads to a minimal SNR gap, $\min_i \gamma_i$, higher than that provided by the BER minimization policy with a gain of 0.3 dB. On the other hand, the BER minimization policy leads to BER lower than that provided by system margin minimization ($2.6 \cdot 10^{-5}$ versus $3.1 \cdot 10^{-5}$). In this example, the dissimilarity is $\mu(\mathcal{B}, \mathcal{C}) = 0.1$, and two subchannels convey different bitrates. All these results are obtained with $r_{\max} = 10$.

Table 1: Example of system margin and BER with $n = 20$, $R = 100$, psdnr = 25 dB, and $\beta = 1$.

	System margin maximization ($\mathcal{B}$)	BER minimization ($\mathcal{C}$)
$\min_i \gamma_i$	6.9 dB	6.6 dB
ber	$3.1 \cdot 10^{-5}$	$2.6 \cdot 10^{-5}$

This example shows that the difference between the resource allocation policies can be small. The question is whether both resource allocations converge and if they converge then in what cases. The following theorem answers the question.

Theorem 5. *In high-SNR regime with square QAM and under bitrate and peak-power constraints, the greedy-type resource allocation that maximizes the system margin converges to the greedy-type resource allocation that minimizes the BER.*

Proof. See Appendix D. □

The consequence of Theorem 5 is that the dissimilarity between the resource allocation that maximizes the system margin and the resource allocation that minimizes the BER is null in high-SNR regime and with square QAM. With square QAM, β should be a multiple of 2. Note that with square modulations, β can also be equal to 1 if the modulations are, for example, those defined in ADSL [34]. Figure 6 exemplifies the convergence with $\beta = 2$ as we will see later in Section 6.

5. Optimal Analytical Resource Allocations

The analytical method is based on convex optimization theory [35]. Unconstrained modulations lead to bitrates r_i defined in $\mathbb{R}$. With $r_i \in \mathbb{R}_+$ the solution is the waterfilling one. With bounded modulation order, that is, $0 \le r_i \le r_{\max}$, the solution is quite different from the waterfilling one. The solution is obtained in the framework of generalized Lagrangian relaxation using Karush-Kuhn-Tucker (KKT) conditions [36].

As the bitrates are continuous and not only integers in this analytical analysis, the constraints (13) do not hold anymore and become

$$\sum_{i=1}^{n} r_i = R\ , \quad 0 \le r_i \le r_{\max}\ \forall i. \tag{31}$$

The KKT conditions associated to the general problem (14) with (31) instead of (13) write [36]

$$-r_i \le 0\ , \quad i = 1, \ldots, n, \tag{32}$$

$$r_i - r_{\max} \le 0\ , \quad i = 1, \ldots, n, \tag{33}$$

$$R - \sum_{i=1}^{n} r_i = 0, \tag{34}$$

$$\mu_i \ge 0\ , \quad i = 1, \ldots, n, \tag{35}$$

$$\nu_i \ge 0\ , \quad i = 1, \ldots, n, \tag{36}$$

$$\mu_i r_i = 0\ , \quad i = 1, \ldots, n, \tag{37}$$

$$\nu_i (r_i - r_{\max}) = 0\ , \quad i = 1, \ldots, n, \tag{38}$$

$$\frac{\partial}{\partial r_i} \phi\left(\left\{r_j\right\}_{j=1}^{n}\right) - \lambda - \mu_i + \nu_i = 0\ , \quad i = 1, \ldots, n, \tag{39}$$

where λ, μ_i and ν_i are the Lagrange multipliers. The first three equations (32)–(34) represent the primal constraints, (35) and (36) represent the dual constraints, (37) and (38) represent the complementary slackness, and (39) is the cancellation of the gradient of Lagrangian with respect to r_i. When the primal problem is convex and the constraints are linear, the KKT conditions are sufficient for the solution to be primal and dual optimal. For the system margin maximization problem, the function $\phi(\cdot)$ is convex over all input bitrates and SNR whereas this function is no longer convex for the BER minimization problem. Appendix C gives the convex domain of the function $\phi(\cdot)$ in the case of BER minimization problem.

The properties of the studied function $\phi(\cdot)$ are such that

$$\frac{\partial}{\partial r_i} \phi\left(\left\{r_j\right\}_{j=1}^{n}\right) = \psi_i(r_i) \tag{40}$$

is independent of r_j for all $j \ne i$. The optimal solution that solves (32)–(39) is then [36]

$$r_i^*(\lambda) = \begin{cases} 0, & \text{if } \lambda \le \psi_i(0), \\ \psi_i^{-1}(\lambda), & \text{if } \psi_i(0) < \lambda < \psi_i(r_{\max}), \\ r_{\max}, & \text{if } \lambda \ge \psi_i(r_{\max}) \end{cases} \tag{41}$$

for all $i = 1, \ldots, n$ and with λ verifying the constraint

$$\sum_{i=1}^{n} r_i^*(\lambda) = R. \tag{42}$$

It is worthwhile noting that the above general solution is the waterfilling one if $r_{\max} \ge R$. The waterfilling is also the solution in the following case. Let $\mathcal{I}'$ be the subset index such that

$$\mathcal{I}' = \{i \mid r_i^* \ne \{0, r_{\max}\}\}, \tag{43}$$

and let R' the target bitrate over $\mathcal{I}'$. In this subset, $\{r_i^*\}_{i \in \mathcal{I}'}$ are solutions of

$$\begin{cases} \dfrac{\partial}{\partial r_i} \phi\left(\left\{r_j\right\}_{j=1}^{n}\right) - \lambda = 0\ , \quad \forall i \in \mathcal{I}' \\ R' - \sum\limits_{i \in \mathcal{I}'} r_i(\lambda) = 0. \end{cases} \tag{44}$$

This is the solution of (14) with unbounded modulations over the subchannel index subset $\mathcal{I}'$. If $\mathcal{I}' = \{1, \cdots, n\}$ and $R' = R$, and (44) is also the solution of (14) with unconstrained modulations.

5.1. System Margin Maximization

Theorem 6. *Under bitrate and peak-power constraints, the asymptotic bit allocation which minimizes the inverse system margin is given by*

$$r_i^* = \frac{R'}{|\mathcal{I}'|} + \frac{1}{|\mathcal{I}'|} \sum_{j \in \mathcal{I}'} \log_2 \frac{\mathrm{snr}_i}{\mathrm{snr}_j}, \quad \forall i \in \mathcal{I}'. \tag{45}$$

Proof. See Appendix E. □

The solution given by Theorem 6 holds for high modulation orders which defines the asymptotic regime, compare Appendix E. If the set $\mathcal{I}'$ is known, then Theorem 6 can be used directly to allocate the subchannel bitrates. Otherwise, $\mathcal{I}'$ should be found first.

The expression of r_i^* in Theorem 6 is a function of the target bitrate R', the number $|\mathcal{I}'|$ of subchannels, and the ratios of SNR. This expression is independent of the mean received SNR or PSDNR. It does not depend on the link budget but only on the relative distribution of subchannel coefficients $\{|h_i|^2\}_{i=1}^{n}$.

5.2. BER Minimization. The arithmetic mean BER minimization has been analytically solved, for example, in [22, 37]. This arithmetic mean measure needs to employ the same number of bits per constellation which limits the system efficiency. The following theorem gives the solution of the weighted mean BER minimization that allows variable constellation sizes in the multichannel system.

Theorem 7. *Under bitrate and peak-power constraints, the asymptotic bit allocation which minimizes the BER is given by*

$$r_i^* = \frac{R'}{|\mathcal{I}'|} + \frac{1}{|\mathcal{I}'|} \sum_{j \in \mathcal{I}'} \log_2 \frac{\mathrm{snr}_i}{\mathrm{snr}_j} \quad \forall i \in \mathcal{I}' \tag{46}$$

with equal in-phase and quadrature bitrates.

Proof. See Appendix F. □

The solution given by Theorem 7 holds for high modulation orders and for subchannel BER lower than 0.1,

and these parameters define the asymptotic regime in this case, compare Appendix F. The optimal asymptotic resource allocation leads to square QAM with $\sqrt{r_i^*}$ conveyed bitrate in each in-phase and quadrature components of the signal of subchannel i. It is important to note that, in asymptotic regime, BER minimization and system margin maximization lead to the same subchannel bitrate allocation. In that case, the asymptotic regime is defined by the more stringent context which is the BER minimization. As we will see in Section 6, this asymptotic behavior can be observed when $\beta = 2$.

The main drawback of the formulas in Theorems 7 and 6 is that the subset $\mathcal{I}'$ must be known. To find this subset, the negative subchannel bitrates and those higher than $r_{\max}$ should be clipped, and $\mathcal{I}'$ can be found iteratively [18]. But clipping negative bitrates first can decrease those higher than $r_{\max}$, and clipping bitrates higher than $r_{\max}$ first can increase the negative ones. It is then not possible to apply first the waterfilling solution and after that to clip the bitrates r_i greater than $r_{\max}$ to converge to the optimal solution. Finding the set $\mathcal{I}'$ requires many comparisons, and we propose a fast iterative solution based on the generalized secant method.

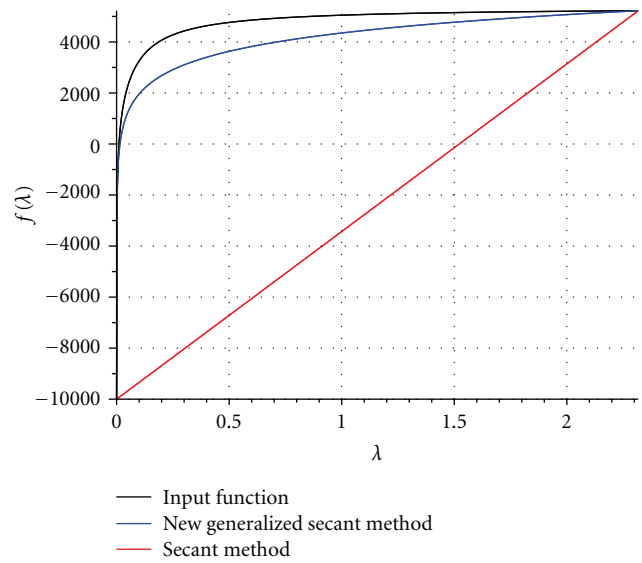

Figure 2: Approximation of the input function $f(\lambda)$ with the generalized secant method and the secant method, $n = 1024$ and $r_{\max} = 15$.

5.3. Lagrangian Resolution. To solve (41), numerical iterative methods are required. It is important to observe that the function defined in (41) is not differentiable, and, thus, methods like Newton's cannot be used [18]. We use the proposed generalized secant method to better fit the function-depending weight and increase the speed of the convergence. An important point for the iterative method is that the initialization value must lead to feasible solution and should be as close as possible to the final solution.

The root of the function defined by (42) is now calculated. Let

$$f(\lambda) = \sum_{i=1}^{n} r_i(\lambda) - R. \tag{47}$$

Theorems 6 and 7 show that $r(\lambda)$ is the sum of $\log_2(\cdot)$ functions. This is the reason why the function $\log_2(\cdot)$ is used in the generalized secant method. Figure 2 shows three functions versus the parameter λ. The first function is the input function $f(\lambda)$, the second one is the function used by the generalized secant method, and the last one if the linear function used by the secant method. In this example, the common points are $\lambda = 0$ and $\lambda = 2.3$. As it is shown, the generalized secant method better fits the input function than the secant method and therefore can improve the speed of the convergence to find the root which is around $\lambda = 1/80$ in this example.

To ensure the convergence of the secant methods, the algorithm should be initialized with λ_1 and λ_2 such as $f(\lambda_1) < 0$ and $f(\lambda_2) > 0$. For both optimization problems, system margin maximization and BER minimization, the parameter λ is given by the function $\psi_i(r_i)$, and it can be reduced to $\lambda = 2^{r_i}/\mathrm{snr}_i$, as shown in Appendices E and F. Parameters $\{\lambda_1, \lambda_2\}$ are then chosen as

$$\lambda_1 = \frac{1}{\max_i \mathrm{snr}_i}, \qquad \lambda_2 = \frac{2^{r_{\max}}}{\min_i \mathrm{snr}_i}. \tag{48}$$

Using (41), $\lambda \le \lambda_1$ leads to $r_i(\lambda) = 0$ for all i, and $\lambda \ge \lambda_2$ leads to $r_i(\lambda) = r_{\max}$ for all i. Then, it follows that $f(\lambda_1) < 0$ and $f(\lambda_2) > 0$ if $R \in (0, nr_{\max})$.

Figure 3 shows the needed number of iterations for the convergence of the generalized and conventional secant methods versus the target bitrate R. Results are given over a Rayleigh distribution of the subchannel SNR with 1024 subchannels. The possible bitrates are then $R \in [0, n \times r_{\max}]$ and $\beta = 2$. Here, $r_{\max} = 15$ and then $R \le 15360$ bits per multidimensional symbol. For comparison, the number of iterations needed by the greedy-type algorithm is also plotted. Note that the greedy-type algorithm can start by empty bitrate or by full bitrate limited by $r_{\max}$ for each subchannel. The number of iterations is then given by $\min\{R, nr_{\max} - R\}$. The iterative secant and generalized secant methods are stopped when the bitrate error is lower than 1. A better precision is not necessary since exact bitrates $\{r_i\}_{i\in\mathcal{I}'}$ can be computed using Theorems 6 and 7 when $\mathcal{I}'$ is known. As it is shown in Figure 3, the generalized secant method converges faster than the secant method, except for the very low target bitrates R. For very high target bitrates, near from $n \times r_{\max}$, the number of iterations with the generalized secant method can be higher than that with the greedy-type algorithm. Except for these particular cases, the generalized secant method needs no more than 4-5 iterations to converge. In conclusion, we can say that with Rayleigh distribution of $\{\mathrm{snr}_i\}_{i=1}^{n}$ and for target bitrates R such that $3\% \le R/nr_{\max} \le 97\%$, the generalized secant method converges faster than the secant method or the greedy-type algorithm.

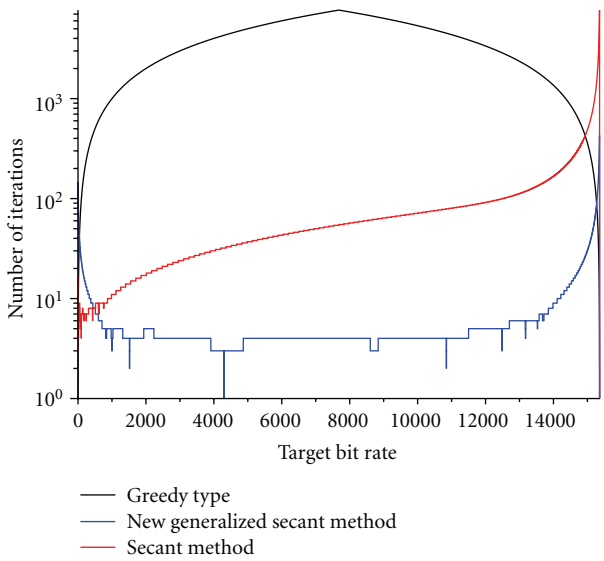

FIGURE 3: Number of iterations of the secant and generalized secant methods, and greedy-type algorithm versus the target bitrate, $n = 1024$, $r_{\max} = 15$.

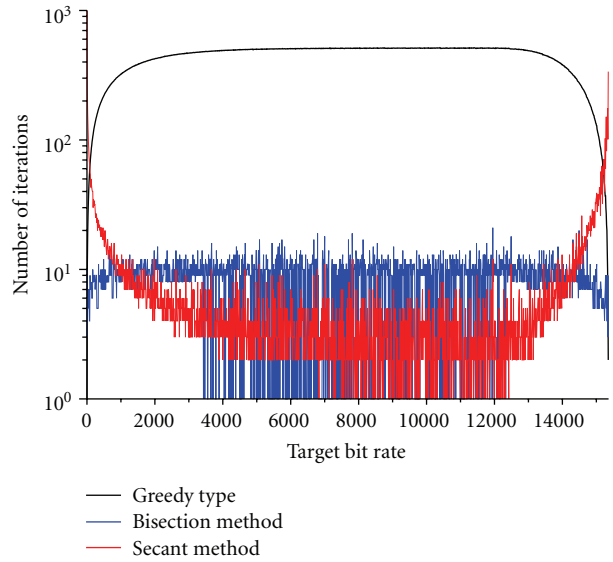

FIGURE 4: Number of iterations of the bisection and secant methods, and greedy-type algorithm for integer-bit solution versus target bitrate, $n = 1024$, $r_{\max} = 15$.

Using the generalized secant method, the bitrates are not integers and for all i, $r_i^* \in [0, r_{\max}]$. These solutions have to be completed to obtain integer bitrates.

5.4. Integer-Bit Solution. Starting from the continuous bitrate allocations previously presented, a loading procedure is developed taking into account the integer nature of the bitrates to be allotted. A simple solution is to consider the integer part of $\{r_i^*\}_{i\in\mathcal{I}'}$ and to complete by a greedy-type algorithm to achieve the target bitrate R. The integer part of $\{r_i^*\}_{i\in\mathcal{I}'}$ is then used as a starting point for the greedy algorithm. This procedure can lead to a high number of iterations. Therefore, the secant or bisection methods are suitable to reduce the number of iterations. The problem to solve is then to find the root of the following function [18]:

$$g(\alpha) = \sum_{i\in\mathcal{I}'} \lfloor r_i^* + \beta\alpha \rfloor - R', \tag{49}$$

where r_i^*, $\mathcal{I}'$, and R' are given by the continuous Lagrangian solution. This is a suboptimal integer bitrate problem, and the optimal one needs to find $\{\alpha_i\}_{i=1}^n$ instead of a unique α. As the optimal solution leads to a huge number of iterations, it is not considered. The function (49) is a nondecreasing and nondifferentiable staircase function such that $g(0) < 0$, $g(1) > 0$ because $\sum_{i\in\mathcal{I}'} r_i^* = R'$. The iterative methods can then be initialized with $\alpha_1 = 0$ and $\alpha_2 = 1$.

Two iterative methods are compared: the bisection one and the secant one. Both methods are also compared to the greedy-type algorithm. Figure 4 presents the number of iterations of the three methods to solve the integer-bit problem of the Lagrangian solution with $\beta = 1$. Results are given over a Rayleigh distribution of the subchannel SNR, with 1024 subchannels and the target bitrates are between 0 and $n \times r_{\max} = 15360$. As it is shown, the convergence is faster with bisection method than with greedy-type algorithm. For target bitrates between 10% and 90% of the maximal loadable bitrate, the secant method outperforms the bisection one with a mean number of iterations around 4 whereas the number of iterations for bisection method is higher than 8. Figure 4 also shows that $|g(0)|$ is all the time lower than the half of number of subchannels and around this value for target bitrate between 10% and 90% of the maximal loadable bitrate. Then, if the number of iterations induces by the greedy-type algorithm to solve the integer-bit problem of the Lagrangian solution that is acceptable in a practical communication system, this greedy-type completion can be used and appears to lead to the optimal resource allocation. This result obtained without proof means that the greedy-type procedure has enough bits to converge to the optimal solution. If the number of iterations induced by the greedy-type algorithm is too high (this number is around $n/2$), the secant method can be used.

The overall analytical resolution of (14) needs few iterations compared to the optimal greedy-type algorithm. Whereas the continuous solution of (14) is optimal, the analytical integer bitrate solution is suboptimal.

6. Greedy-Type versus Analytical Resource Allocations

In the previous section, the numbers of iterations of the algorithms have been compared. In this section, robustness comparison is presented and the analytical solutions obtained in asymptotic regime are also applied in nonasymptotic regime which means that $\beta = 1$ and modulation orders can be low.

The evaluated OFDM communication system is composed of 1024 subcarriers without interferences between the symbols or the subcarriers. The channel is the Rayleigh

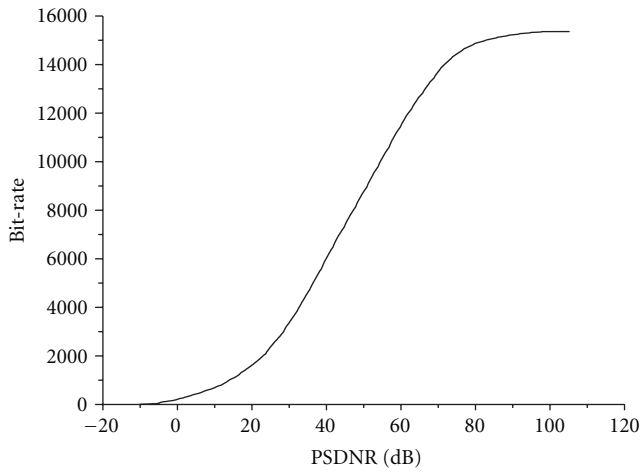

Figure 5: Target bitrate versus input PSDNR.

fading one with independent and identically distributed elements. The richest modulation order is $r_{\max} = 10$. The robustness measures are evaluated for different target bitrates which are given with the following arbitrary equation:

$$R = \left\lfloor \sum_{i=1}^{n} \min\left(\log_2\left(1 + \frac{\mathrm{snr}_i}{2}\right), r_{\max}\right) \right\rfloor. \tag{50}$$

This equation ensures reliable communications for all the input target bitrates or PSDNR. The empirical relationship between PSDNR and target bitrate is also given in Figure 5.

Figure 6 presents the output BER and the system margin of three resource allocation policies versus the target bitrate R. The first one, $\mathcal{A}$, is obtained using analytical optimization, the second, $\mathcal{B}$, is the solution of the greedy-type algorithm which maximizes the system margin, and the third, $\mathcal{C}$, is the solution of the greedy-type algorithm which minimizes the BER. Two cases are presented: one with $\beta = 1$ and the other with $\beta = 2$. All subchannel BER are lower than $2 \cdot 10^{-2}$ to use valid BER approximations. Note that, with $\beta = 1$, the system margin of allocation $\mathcal{B}$ is almost equal to 8.9 dB for all target-bit rates. This constant system margin γ is not a feature of the algorithm but is only a consequence of the relation between the target bitrate and the PSDNR.

To enhance the equivalences and the differences between the resource allocation policies, the dissimilarity is also given in Figure 6 with $\beta = 1$ and $\beta = 2$. As expected in both cases, $\beta = 1$ and $\beta = 2$, the minimal BER are obtained with allocation $\mathcal{C}$, and the maximal system margins with allocation $\mathcal{B}$.

With $\beta = 1$ and when the target bitrate increases, the Lagrangian solution converges faster to the optimal system margin maximization solution, $\mathcal{B}$, than to the optimal BER minimization solution, $\mathcal{C}$. Note that Theorem 7 is an asymptotic result valid for square QAM. With $\beta = 1$, the QAM can be rectangular, and the asymptotic result of Theorem 7 is not applicable, contrary to the result of Theorem 6 where there is not any condition on the modulation order.

The case $\beta = 2$ shows the equivalence between the optimal system margin maximization allocation and the

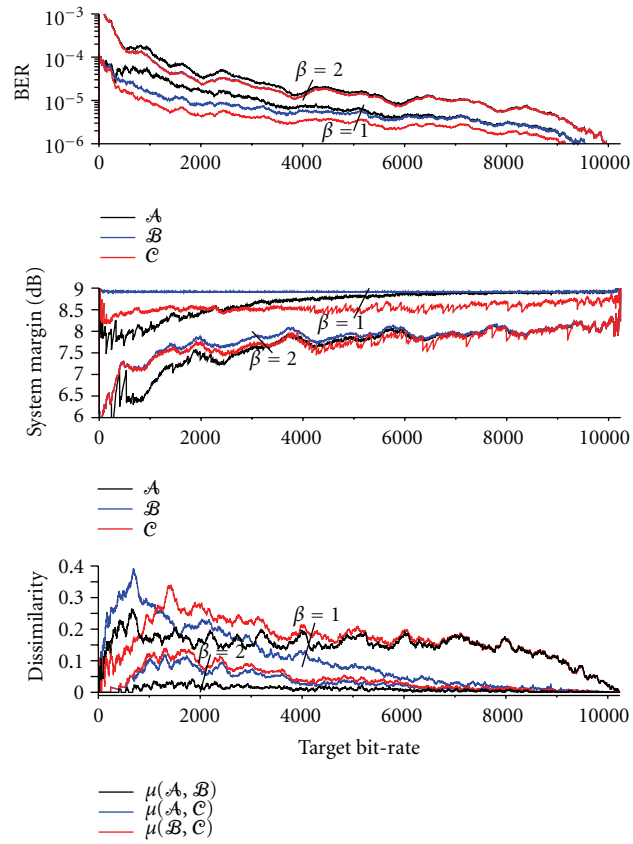

Figure 6: BER, system margin and dissimilarity versus target bitrate for Lagrangian ($\mathcal{A}$), greedy-type system margin maximization ($\mathcal{B}$) and greedy-type BER minimization ($\mathcal{C}$) algorithms, $n = 1024$, $r_{\max} = 10$, and $\beta \in \{1, 2\}$.

optimal BER minimization allocation. In this case, the asymptotic result given by Theorems 5 and 7 can be applied because the modulations are square QAM, and the convergence is ensured with high modulation orders, that is, high target bitrates. Beyond a mean bitrate per subchannel around 6, that corresponds to a target bitrate around 6000, all the allocations $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are equivalent, and the dissimilarity is almost equal to zero. In nonasymptotic regime, the differences in BER and system margin are low. The system margin differences are lower than 1 dB, and the ratios between two BER are around 3. In practical integrated systems, these low differences will not be significant and will lead to similar solutions for both optimization policies. Therefore, these resource allocations can be interchanged.

7. Conclusion

Two robustness optimization problems have been analyzed in this paper. Weighted mean BER minimization and minimal subchannel margin maximization have been solved under peak-power and bitrate constraints. The asymptotic convergence of both robustness optimizations has been proved for analytical and algorithmic approaches. In nonasymptotic regime, the resource allocation policies

can be interchanged depending on the robustness measure and the operating point of the communication system. We have also proved that the equivalence between SNR-gap maximization and power minimization in conventional MM problem does not hold with peak-power limitation without additional conditions. Integer bit solution of analytical continuous bitrates has been obtained with a new generalized secant method, and bit-loading solutions have been compared with a new defined dissimilarity measure. The low computational effort of the suboptimal resource allocation strategy, based on the analytical approach, leads to a good tradeoff between performance and complexity.

Appendices

A. Proof of Lemma 3

We prove that the optimal allocation is reached starting from empty loading with the same intermediate loading than starting from optimal loading to empty loading. To simplify the notation and without loss of generality, $\beta = 1$.

Let $[r_1^*, \ldots, r_n^*]$ be the optimal allocation that minimizes the inverse system margin $\gamma(R^*)^{-1}$ for the target bitrate R^*, and then

$$\gamma(R^*)^{-1} = \max_i \frac{2^{r_i} - 1}{\mathrm{snr}_i}. \tag{A.1}$$

Let $[r_1, \ldots, r_n]$ be the optimal allocation that minimizes the inverse system margin $\gamma(R+1)^{-1}$ for the target bitrate $R+1 \leq R^*$. The optimal allocation for target bitrate R is obtained iteratively by removing one bit at a time from the subchannel k with the highest inverse system margin [38]

$$k = \arg\ \max_i \frac{2^{r_i} - 1}{\mathrm{snr}_i} \tag{A.2}$$

or

$$\frac{2^{r_k} - 1}{\mathrm{snr}_k} \geq \frac{2^{r_i} - 1}{\mathrm{snr}_i}, \quad \forall i = 1, \ldots, n. \tag{A.3}$$

The last bit removed is from the subchannel with the lowest inverse-SNR, snr_i^{-1}, because the bits over the highest inverse-SNR are first removed.

Now, let $[r_1, \ldots, r_n]$ be the optimal allocation that minimizes the inverse system margin $\gamma(R)^{-1}$ for the target bitrate $R < R^*$. Following the algorithm strategy, the optimal allocation for target bitrate $R + 1$ is obtained adding one bit on subchannel j such that

$$j = \arg\ \min_i \frac{2^{r_i+1} - 1}{\mathrm{snr}_i}. \tag{A.4}$$

We first prove that

$$\gamma(R+1)^{-1} = \frac{2^{r_j+1} - 1}{\mathrm{snr}_j}. \tag{A.5}$$

Suppose that there exists j' such that

$$\frac{2^{r_{j'}} - 1}{\mathrm{snr}_{j'}} > \frac{2^{r_j+1} - 1}{\mathrm{snr}_j}, \tag{A.6}$$

then one bit must be added to subchannel j to obtain $r_j + 1$ bits before adding one bit to subchannel j' to obtain $r_{j'}$ bits which means that $[r_1, \ldots, r_n]$ is not optimal. As $[r_1, \ldots, r_n]$ is optimal by definition, it yields

$$\frac{2^{r_i} - 1}{\mathrm{snr}_i} \leq \frac{2^{r_j+1} - 1}{\mathrm{snr}_j} \quad \forall i = 1, \ldots, n \tag{A.7}$$

which proves (A.5). The first allocated bit is from the subchannel with the lowest inverse SNR given by (A.4) with $r_i = 0$ for all i.

Comparing (A.3) with (A.7) yields that $k = j$, and the index subchannel of the first added bit is the same as the last removed bit. All the intermediate allocations are then identical with bit-addition and bit-removal methods. There exists only one way to reach the optimal allocation R^* starting from the empty loading.

Proof of Lemma 3 can also be provided in the framework of matroid algebraic theory [19, 39].

B. Proof of Lemma 4

To simplify the notation and without loss of generality, the proof is given with $\beta = 1$. Let $[r_1, \cdots, r_n]$ be the optimal allocation for the target bitrate R such that $\sum r_i = R$. Let $R+1$ the new target bitrate. We first prove that

$$\Delta_i(r_i) = (r_i + 1)\mathrm{ber}_i(r_i + 1) - r_i \mathrm{ber}_i(r_i) \tag{B.1}$$

is a good measure at each step of the greedy-type algorithm for the BER minimization, and finally that $(r_i + 1)\mathrm{ber}_i(r_i + 1)$ can be used instead of $\Delta_i(r_i)$.

Starting from the optimal allocation of target bitrate R, the new target bitrate $R + 1$ is obtained by increasing r_j by one bit

$$\mathrm{ber}(R+1) = \frac{\left(r_j + 1\right)\mathrm{ber}_j\left(r_j + 1\right) + \sum_{i=1, i\neq j}^{n} r_i \mathrm{ber}_i(r_i)}{1 + \sum_{i=1}^{n} r_i} \tag{B.2}$$

and, using Δ_j,

$$\mathrm{ber}(R+1) = \frac{\Delta_j\left(r_j\right)}{R+1} + \frac{R}{R+1}\mathrm{ber}(R). \tag{B.3}$$

The $\mathrm{ber}(R + 1)$ which is equal to $\phi(\{r_i^{(k+1)}\}_{i=1}^n)$ in (29) is minimized only if $\Delta_j(r_j)$ is minimized. The minimum $\mathrm{ber}(R + 1)$ is then obtained with the increase of one bit in the subchannel j such that

$$j = \arg\ \min_i \Delta_i(r_i). \tag{B.4}$$

To complete the proof by induction, the relation must be true for $\mathrm{ber}(1)$. This is simply done by recalling that $\mathrm{ber}_i(0) = 0$, and then

$$\min\ \mathrm{ber}(1) = \min_i \mathrm{ber}_i(1) = \min_i\ \Delta_i(0). \tag{B.5}$$

The convergence of the algorithm to a unique solution needs the convexity of the function $r_i \mapsto r_i \mathrm{ber}(r_i)$. This convexity

is verified at high SNR. Appendix C provides a more precise domain of validity.

It remains to prove that $(r_i+1)\mathrm{ber}_i(r_i+1)$ can be used instead of $\Delta_i(r_i)$. In high SNR regime

$$\mathrm{ber}_i(r_i+1) \gg \mathrm{ber}_i(r_i) \tag{B.6}$$

and then

$$\lim_{\mathrm{snr}_i \to +\infty} \Delta_i(r_i) = (r_i+1)\mathrm{ber}_i(r_i+1) \tag{B.7}$$

which proves the lemma.

In low SNR regime, the approximation of Δ_i by $(r_i+1)\mathrm{ber}_i(r_i+1)$ remains valid; the dissimilarity between allocation using Δ_i (B.1) and allocation using $(r_i+1)\mathrm{ber}_i(r_i+1)$ is null in the domain of validity given by Appendix C.

C. Range of Convexity of $r_i\mathrm{ber}_i$

Let

$$f : \mathbb{N} \longrightarrow \mathbb{R}_+ r_i \longmapsto r_i\mathrm{ber}_i(r_i, \mathrm{snr}_i) \tag{C.1}$$

which equals the SER for high-SNR regime and Gray mapping. The function f is a strictly increasing function: $f(r_i) < f(r_i+1)$ for all snr_i, because $\mathrm{ber}(r_i, \mathrm{snr}_i) \le \mathrm{ber}(r_i+1, \mathrm{snr}_i)$ and $r_i < r_i+1$. Let $\Delta(r_i) = f(r_i+1) - f(r_i)$, and then

$$\begin{aligned}\Delta(r_i+1) - \Delta(r_i) &= f(r_i+2) - 2f(r_i+1) + f(r_i)\\ &\ge (r_i+1)(\mathrm{ber}_i(r_i+2) - 2\mathrm{ber}_i(r_i+1)).\end{aligned} \tag{C.2}$$

If $\mathrm{ber}_i(r_i+2) \ge 2\mathrm{ber}_i(r_i+1)$, then the function f is locally convex or defines a convex hull. This relation is verified for BER lower than 2×10^{-2} and for all $r_i \ge 0$.

D. Proof of Theorem 5

We prove that both metrics used in Lemmas 3 and 4 lead to the same subchannel SNR ordering. Let

$$\begin{aligned} f(r_i, \mathrm{snr}_i) &= \frac{2^{r_i+\beta}-1}{\mathrm{snr}_i},\\ g(r_i, \mathrm{snr}_i) &= (r_i+\beta)\mathrm{ber}_i(r_i+\beta).\end{aligned} \tag{D.1}$$

We then have to prove that

$$f(r_i, \mathrm{snr}_i) \le f\left(r_j, \mathrm{snr}_j\right) \Longleftrightarrow g(r_i, \mathrm{snr}_i) \le g\left(r_j, \mathrm{snr}_j\right). \tag{D.2}$$

It is straightforward that

$$f(r_i, \mathrm{snr}_i) \le f\left(r_j, \mathrm{snr}_j\right) \Longleftrightarrow \frac{\mathrm{snr}_j}{\mathrm{snr}_i} \le \frac{2^{r_j+\beta}-1}{2^{r_i+\beta}-1}. \tag{D.3}$$

With square QAM, in high SNR regime and using (26)

$$g(r_i, \mathrm{snr}_i) = 2\left(1 - \frac{1}{\sqrt{2^{r_i+\beta}}}\right)\mathrm{erfc}\left(\sqrt{\frac{3}{2(2^{r_i+\beta}-1)}\mathrm{snr}_i}\right) \tag{D.4}$$

and it can be approximated by the following valid expression

$$g(r_i, \mathrm{snr}_i) = 2\,\mathrm{erfc}\left(\sqrt{\frac{3}{2(2^{r_i+\beta}-1)}\mathrm{snr}_i}\right). \tag{D.5}$$

Then,

$$g(r_i, \mathrm{snr}_i) \le g\left(r_j, \mathrm{snr}_j\right) \Longleftrightarrow \frac{\mathrm{snr}_j}{\mathrm{snr}_i} \le \frac{2^{r_j+\beta}-1}{2^{r_i+\beta}-1} \tag{D.6}$$

which is also given by the first inequality. In high SNR regime and with square QAM, that is, $\beta = 2$, $f(\cdot)$ and $g(\cdot)$ lead to the same subchannel SNR ordering and then

$$\arg\min_i f(r_i, \mathrm{snr}_i) = \arg\min_i g(r_i, \mathrm{snr}_i). \tag{D.7}$$

This last equation does not hold in low SNR regime (the BER approximation is not valid) or when the modulations are not square, that is, when r_i is odd. Note that (D.5) is not only a good approximation in high SNR regime, it can also be used with high modulation orders with moderate SNR regime defined in Appendix C.

E. Proof of Theorem 6

As the infinite norm is not differentiable, we use the k norm with

$$\lim_{k\to+\infty}\left(\sum_{i\in\mathcal{I}'}\gamma_i^{-k}\right)^{1/k} = \max_{i\in\mathcal{I}'}(\gamma_i^{-1}). \tag{E.1}$$

In the subset $\mathcal{I}'$, the Lagrangian of (18) for all k is

$$L_k(\{r_i\}_{i\in\mathcal{I}'}, \lambda) = \left(\sum_{i\in\mathcal{I}'}\frac{(2^{r_i}-1)^k}{\mathrm{snr}_i^k}\right)^{1/k} + \lambda\left(R' - \sum_{i\in\mathcal{I}'} r_i\right). \tag{E.2}$$

Let λ' such as

$$\lambda' = \left(\sum_{i\in\mathcal{I}'}\frac{(2^{r_i}-1)^k}{\mathrm{snr}_i^k}\right)^{(k-1)/k}\frac{\lambda}{\log 2}. \tag{E.3}$$

The optimal condition yields

$$2^{r_i}(2^{r_i}-1)^{k-1} = \mathrm{snr}_i^k\lambda'. \tag{E.4}$$

In asymptotic regime, $r_i \gg 1$ and then $2^{r_i}-1 \simeq 2^{r_i}$. The equation of the optimal condition can be simplified and

$$r_i = \log_2(\mathrm{snr}_i) + \frac{1}{k}\log_2\lambda'. \tag{E.5}$$

The Lagrange multiplier is to identify using the bitrate constraint, and replacing λ' in the above equation leads to the solution

$$r_i = \frac{R'}{|\mathcal{I}'|} + \frac{1}{|\mathcal{I}'|}\sum_{j\in\mathcal{I}'}\log_2\frac{\mathrm{snr}_i}{\mathrm{snr}_j}. \tag{E.6}$$

Note that we do not need to calculate the convergence of the solution with $k \to +\infty$ to obtain the result for the infinite norm. The result holds for all values of k in asymptotic regime.

With $k = 1$, the problem is a sum SNR-gap maximization problem under peak-power constraint, and it can be solved without asymptotic regime condition. Note that this sum SNR-gap maximization problem, or sum inverse SNR-gap minimization problem, under peak-power and bitrate constraints is

$$\min \sum_{i\in\mathcal{I}'} \gamma_i^{-1} = \min_{\{r_i\}_{i\in\mathcal{I}'}} \sum_{i\in\mathcal{I}'} (2^{r_i}-1)\frac{\sigma_{W_i}^2}{|h_i|^2 P p_i} \tag{E.7}$$

and is very similar to power minimization problem under bitrate and SNR-gap constraints exchanging p_i with γ_i^{-1}

$$\min \sum_{i\in\mathcal{I}'} p_i = \min_{\{r_i\}_{i\in\mathcal{I}'}} \sum_{i\in\mathcal{I}'} (2^{r_i}-1)\frac{\sigma_{W_i}^2}{|h_i|^2 P \gamma_i^{-1}}. \tag{E.8}$$

Both problems are identical if $p_i\gamma_i = \alpha$ as it is stated by Lemma 1.

F. Proof of Theorem 7

To prove this theorem, variables I_i and J_i are used instead of r_i with

$$I_i = 2^{\lfloor r_i/2\rfloor}, \qquad J_i = 2^{\lceil r_i/2\rceil}, \tag{F.1}$$

and the bitrate constraint is

$$R = \sum_{i=1}^{n} \log_2(I_i J_i). \tag{F.2}$$

In the subset $\mathcal{I}'$, the Lagrangian of (25) is then

$$\begin{aligned} L(\{I_i,J_i\}_{i\in\mathcal{I}'},\lambda) = {} & \frac{1}{R'}\sum_{i\in\mathcal{I}'}\left(2-\frac{1}{I_i}-\frac{1}{J_i}\right) \\ & \times \operatorname{erfc}\left(\sqrt{\frac{3}{I_i^2+J_i^2-2}\operatorname{snr}_i}\right) \\ & + \lambda\left(R' - \sum_{i\in\mathcal{I}'}\log_2(I_iJ_i)\right). \end{aligned} \tag{F.3}$$

Let $X_i \in \{I_i, J_i\}$, then

$$\frac{\partial L}{\partial X_i} = X_i f(I_i,J_i) + \frac{1}{X_i^2} g(I_i,J_i) - \frac{1}{X_i}\lambda, \tag{F.4}$$

with

$$\begin{aligned} f(I_i,J_i) &= \frac{1}{R'}\left(2-\frac{1}{I_i}-\frac{1}{J_i}\right)\frac{2\sqrt{3\operatorname{snr}_i}\times e^{-3\operatorname{snr}_i/(I_i^2+J_i^2-2)}}{\sqrt{\pi}(I_i^2+J_i^2-2)^{3/2}}, \\ g(I_i,J_i) &= \frac{1}{R'}\operatorname{erfc}\left(\sqrt{\frac{3\operatorname{snr}_i}{I_i^2+J_i^2-2}}\right). \end{aligned} \tag{F.5}$$

The optimality condition yields

$$(I_i^2 - J_i^2) I_i J_i f(I_i,J_i) = (I_i - J_i) g(I_i,J_i) \ \forall i. \tag{F.6}$$

A trivial solution is $I_i = J_i$, and the other solution must verify

$$(I_i + J_i) I_i J_i f(I_i,J_i) - g(I_i,J_i) = 0. \tag{F.7}$$

To find the root of (F.7), let

$$h(x,y) = x\sqrt{y}e^{-y} - \operatorname{erfc}\left(\sqrt{y}\right) \tag{F.8}$$

with

$$\begin{aligned} x &= \frac{2}{\sqrt{\pi}}\frac{(I_i+J_i)I_iJ_i}{I_i^2+J_i^2-2}\left(2-\frac{1}{I_i}-\frac{1}{J_i}\right), \\ y &= \frac{3\operatorname{snr}_i}{I_i^2+J_i^2-2}. \end{aligned} \tag{F.9}$$

We will prove that this function is positive in a specific domain. Consider that

(1) $\sqrt{y}e^{-y} > \operatorname{erfc}(\sqrt{y})$ for $y \geq 0.334$, then for BER lower than 10^{-1}.

(2) $\sqrt{\pi}/2x > 1$ for $\{I_i, J_i\} \in [1,+\infty)^2$ and $I_i \neq 1$ or $J_i \neq 1$, and $\lim_{I_i,J_i\to 1}\sqrt{\pi}/2x = 1^+$.

Then, in the domain defined by

$$\{I_i, J_i\} \in [1,+\infty)^2 \wedge \operatorname{ber}_i \leq 0.1, \tag{F.10}$$

$h(x,y)$ is positive, and (F.7) has no solution. Thus, the only one solution of (F.6) with (F.10) is $I_i = J_i$. As we will see later the domain of (F.10) is less restrictive than the asymptotic one.

The problem is now to allocate bits with square QAM. The following upper bound is used:

$$\operatorname{ber}(r_i) = \frac{2}{r_i}\operatorname{erfc}\left(\sqrt{\frac{3\operatorname{snr}_i}{2(2^{r_i}-1)}}\right). \tag{F.11}$$

Note that this upper bound is a tight approximation with high SNR and with high modulation orders. The Lagrangian is that

$$\begin{aligned} L(\{r_i\}_{i\in\mathcal{I}'},\lambda) = {} & \frac{2}{R'}\sum_{i\in\mathcal{I}'}\operatorname{erfc}\left(\sqrt{\frac{3\operatorname{snr}_i}{2(2^{r_i}-1)}}\right) \\ & + \lambda\left(R' - \sum_{i\in\mathcal{I}'} r_i\right). \end{aligned} \tag{F.12}$$

And its derivative is

$$\frac{\partial L}{\partial r_i} = \frac{\ln 2}{\sqrt{\pi}}\frac{2^{r_i}}{2^{r_i}-1}\sqrt{\frac{3\operatorname{snr}_i}{2(2^{r_i}-1)}}e^{-3\operatorname{snr}_i/2(2^{r_i}-1)} - \lambda. \tag{F.13}$$

Let $r_i \gg 1$ for all i, then $2^{r_i} - 1 \simeq 2^{r_i}$, and the optimality condition yields

$$-\frac{3\operatorname{snr}_i}{2^{r_i}}e^{-3\operatorname{snr}_i/2^{r_i}} = -\frac{2\lambda^2\pi}{\ln^2 2}. \tag{F.14}$$

With reliable communication over the subchannel i, the Shannon's relation states that $r_i \leq \log_2(1+\operatorname{snr}_i)$ and $3\operatorname{snr}_i/2^{r_i} \geq 3/2$ because $r_i \geq 1$. The relation between r_i and

λ is then bijective, and the real branch W_{-1} of the Lambert function [40] can be used with no possibility for confusion

$$r_i = \log_2(3\mathrm{snr}_i) - \log_2\left(-W_{-1}\left(-\frac{2\lambda^2\pi}{\ln^2 2}\right)\right). \tag{F.15}$$

With the bitrate constraint $R' = \sum_{i\in\mathcal{I}'} r_i$, we can write

$$-\log_2\left(-W_{-1}\left(-\frac{2\lambda^2\pi}{\ln^2 2}\right)\right) = \frac{R'}{|\mathcal{I}'|} - \frac{1}{|\mathcal{I}'|}\sum_{i=1}^{n}\log_2(3\mathrm{snr}_i) \tag{F.16}$$

and with (F.15)

$$r_i = \frac{R'}{|\mathcal{I}'|} + \frac{1}{|\mathcal{I}'|}\sum_{j\in\mathcal{I}'}^{n}\log_2\frac{\mathrm{snr}_i}{\mathrm{snr}_j}. \tag{F.17}$$

This result is obtained with square QAM in asymptotic regime (high modulation orders and high SNR) which is a more restrictive domain than that of (F.10).

Acknowledgments

The research leading to these results has received partial funding from the European Community's Seventh Framework Program FP7/2007-2013 under grand agreement no 213311 also referred to as OMEGA.

References

[1] A. N. Akansu, P. Duhamel, X. Lin, and M. De Courville, "Orthogonal transmultiplexers in communication: a review," *IEEE Transactions on Signal Processing*, vol. 46, no. 4, pp. 979–995, 1998.

[2] J. M. Cioffi, "A multicarrier primer," Tech. Rep. ANSI T1E1.4/91–157, Committee Contribution, Washington, DC, USA, 1991.

[3] G. D. Forney and M. V. Eyuboglu, "Combined equalization and coding using precoding," *IEEE Communications Magazine*, vol. 29, no. 12, pp. 25–34, 1991.

[4] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the Institute of Radio Ingineers*, vol. 37, pp. 10–21, 1949.

[5] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel gaussian channels with arbitrary input distributions," *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 3033–3051, 2006.

[6] T. Antonakopoulos and N. Papandreou, "Bit and power allocation in constrained multicarrier systems: the single-user case," *Eurasip Journal on Advances in Signal Processing*, vol. 2008, Article ID 643081, 14 pages, 2008.

[7] S. T. Chung and A. J. Goldsmith, "Degrees of freedom in adaptive modulation: a unified view," *IEEE Transactions on Communications*, vol. 49, no. 9, pp. 1561–1571, 2001.

[8] A. Fasano and G. di Blasio, "The duality between margin maximization and rate maximization discrete loading problems," in *Proceedings of the IEEE Workshop on Signal Processing Advances in Wireless Communications*, pp. 621–625, July 2004.

[9] D. P. Palomar and J. R. Fonollosa, "Practical algorithms for a family of waterfilling Solutions," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 686–695, 2005.

[10] I. Kim, I. S. Park, and Y. H. Lee, "Use of linear programming for dynamic subcarrier and bit allocation in multiuser OFDM," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 4, pp. 1195–1207, 2006.

[11] D. Hughes-Hartogs, *Ensemble Modem Structure for Imperfect Transmission Media*, Telebit Corporation, Cupertino, Calif, USA, 1987, US Patent 4,679,227.

[12] B. S. Krongold, K. Ramchandran, and D. L. Jones, "Computationally efficient optimal power allocation algorithms for multicarrier communication systems," *IEEE Transactions on Communications*, vol. 48, no. 1, pp. 23–27, 2000.

[13] W. J. Choi, K. W. Cheong, and J. M. Cioffi, "Adaptive modulation with limited peak power for fading channels," in *Proceedings of the 51st Vehicular Technology Conference "Shaping History Through Mobile Technologies" (VTC '00)*, pp. 2568–2572, May 2000.

[14] P. Uthansakul and M. E. Bialkowski, "Performance comparisons between greedy and Lagrange algorithms in adaptive MIMO MC-CDMA systems," in *Proceedings of the Asia-Pacific Conference on Communications*, pp. 163–167, Perth, Australia, October 2005.

[15] A. J. Goldsmith, "Variable-rate variable-power MQAM for fading channels," *IEEE Transactions on Communications*, vol. 45, no. 10, pp. 1218–1230, 1997.

[16] S. Ye, R. S. Blum, and L. J. Cimini Jr, "Adaptive OFDM systems with imperfect channel state information," *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, pp. 3255–3265, 2006.

[17] N. Y. Ermolova and B. Makarevitch, "Practical approaches to adaptive resource allocation in OFDM systems," *Eurasip Journal on Wireless Communications and Networking*, vol. 2008, Article ID 160307, 2008.

[18] E. Baccarelli, A. Fasano, and M. Biagi, "Novel efficient bit-loading algorithms for peak-energy-limited ADSLtype multicarrier systems," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1237–1247, 2002.

[19] A. Fasano, "On the optimal discrete bit loading for multicarrier systems with constraints," in *Proceedings of the 57th IEEE Semiannual Vehicular Technology Conference (VTC '03)*, vol. 2, pp. 915–919, April 2003.

[20] M. A. Khojastepour and B. Aazhang, "The capacity of average and peak power constrained fading channels with channel side information," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, vol. 2, pp. 77–82, Atlanta, Ga, USA, March 2004.

[21] Y. Ding, T. N. Davidson, and K. M. Wong, "On improving the BER performance of rate-adaptive block transceivers, with applications to DMT," in *Proceedings of the IEEE Global Communications Conference*, vol. 3, pp. 1654–1658, San Francisco, Calif, USA, December 2003.

[22] D. P. Palomar, *A unified framework for communications through MIMO channels [Ph.D. thesis]*, Universitat politecnica de Catalunya, Barcelona, Spain, 2003.

[23] A. M. Wyglinski, F. Labeau, and P. Kabal, "Bit loading with BER-constraint for multicarrier systems," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1383–1387, 2005.

[24] G. D. Forney, R. G. Gallager, G. R. Lang, F. M. Longstaff, and S. U. Qureshi, "Efficient modulation for bandlimited channels," *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 5, pp. 632–647, 1984.

[25] J. M. Cioffi, *Digital Communication*, Department of Electrical Engineering, Stanford University, Stanford, Calif, USA, 2007, Course.

[26] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, "Optimum linear joint transmit-receive processing forMIMO channels with QoS constraints," *IEEE Transactions on Signal Processing*, vol. 52, no. 5, pp. 1179–1197, 2004.

[27] J. G. Proakis, *Digital Communications*, Electrical engineering, McGraw-Hill, New York, NY, USA, 3rd edition, 1995.

[28] K. Cho and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations," *IEEE Transactions on Communications*, vol. 50, no. 7, pp. 1074–1080, 2002.

[29] B. Fox, "Discrete optimization via marginal analysis," *Management Science*, vol. 13, no. 3, pp. 210–216, 1966.

[30] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "Practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Transactions on Communications*, vol. 43, no. 2, pp. 773–775, 1995.

[31] J. Campello, "Optimal discrete bit loading for multicarrier modulation systems," in *IEEE International Symposium on Information Theory*, p. 193, IEEE Publishing, Cambridge, Mass, USA, 1998.

[32] H. E. Levin, "A complete and optimal data allocation method for practical discrete multitone systems," in *Proceedings of the IEEE Global Communications Conference*, vol. 1, pp. 369–374, San Antonio, Tex, USA, November 2001.

[33] J. Campello, "Practical bit loading for DMT," in *Proceedings of the IEEE International Conference on Communications*, vol. 2, pp. 801–805, British Columbia, Canada, June 1999.

[34] G.992.3, *Asymmetric Digital Subscriber Line Transceivers*, ITU-T Recommendation, Geneva, Switzerland, 2002.

[35] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, Article ID 1664998, pp. 1426–1438, 2006.

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

[37] A. Pascual-Iserte, *Channel state information and joint transmitter-receiver design in multi-antenna systems [Ph.D. thesis]*, Universitat Politecnica de Catalunya, Barcelona, Spain, 2004.

[38] L.-P. Zhu, Y. Yao, S.-D. Zhou, and S.-W. Dong, "A heuristic optimal discrete bit allocation algorithm for margin maximization in DMT systems," *Eurasip Journal on Advances in Signal Processing*, vol. 2007, Article ID 12140, 7 pages, 2007.

[39] R. J. Wilson, "An introduction to matroid theory," *The American Mathematical Monthly*, vol. 80, no. 5, pp. 500–525, 1973.

[40] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Advances in Computational Mathematics*, vol. 5, no. 4, pp. 329–359, 1996.