Joint Optimization of Bit and Power Allocation for Multicarrier Systems with Average BER Constraint

Ebrahim Bedeer, Octavia A. Dobre, Mohamed H. Ahmed, and Kareem E. Baddour [†]

Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL, Canada
[†] Communications Research Centre, Ottawa, ON, Canada
Email: {e.bedeer, odobre, mhahmed}@mun.ca, kareem.baddour@crc.ca

Abstract—**This paper proposes a novel joint bit and power allocation algorithm for multicarrier systems operating in fading environments. The algorithm jointly maximizes the throughput and minimizes the transmitted power, while guaranteeing a target average bit error rate (BER). Simulation results are described and they illustrate the performance of the proposed scheme and demonstrate its superiority with respect to existing schemes.**

I. INTRODUCTION

Multicarrier modulation is recognized as a robust and efficient transmission technique, as evidenced by its consideration for diverse communication systems and adoption by several wireless standards [1], [2]. The performance of multicarrier communication systems can be significantly improved by dynamically adapting the transmission parameters, such as power, constellation size, symbol rate, coding rate/scheme, or any combination of these. This can be achieved according to the channel conditions or the wireless standard specifications [3]–[6].

Generally speaking, the problem of optimally loading bits and power per subcarrier can be categorized into two main classes: *rate maximization* (RM) and *margin maximization* (MM) (see, e.g., [3]–[6] and references therein). For the former, the objective is to maximize the achievable data rate [3], while for the latter the objective is to maximize the achievable system margin [4] (i.e., minimizing the total transmit power given a target data rate). Most of the work done in the literature focused on maximizing either the RM or the MM problem separately. In [3], Wyglinski *et al.* proposed an incremental bit loading algorithm with uniform power in order to maximize the throughput while guaranteeing a target BER. Liu and Tang [4] proposed a low complexity power loading algorithm with uniform bit loading that aims to minimize the transmit power while guaranteeing a target BER. In [5], Goldfeld *et al.* proposed a quasi-optimal power loading algorithm that requires no iterations in order to minimize the overall BER with fixed constellation size across all subcarriers.

In emerging wireless communication systems, different requirements are needed. For example, minimizing the transmit power represents a priority when operating in the interference-prone unlicensed bands, or in proximity to other adjacent users. While maximizing the throughput is favoured if sufficient guard bands exist. This motivates us to jointly optimize the RM and MM problems by introducing a weighting factor that reflects the importance of the competing throughput and power objectives. Accordingly, in this paper we propose an optimal algorithm that jointly maximizes the throughput and minimizes the total transmit power, subject to a constraint on the average BER. Simulation results show that the proposed algorithm outperforms existing bit and power loading schemes in the literature.

The remainder of this paper is organized as follows. Section II introduces the proposed joint bit and power allocation algorithm. Simulation results are presented in Section III, while conclusions are drawn in Section IV.

Throughout this paper we will use bold-faced upper case letters to denote matrices, e.g., $\mathbf{X}$, bold-faced lower case letters for vectors, e.g., $\mathbf{x}$, and light-faced letters for scalar quantities, e.g., x. $\mathbf{I}$ represents the identity matrix, $[.]^T$ denotes the transpose operation, ∇ represents the gradient, and $\lfloor x \rfloor$ is the largest integer not greater than x.

II. PROPOSED ALGORITHM

A. Optimization Problem Formulation

A multicarrier communication system decomposes the signal bandwidth into a set of N orthogonal narrowband subcarriers of equal bandwidth. Each subcarrier i transmits b_i bits using power $\mathcal{P}_i$, $i = 1, ..., N$. A delay- and error-free feedback channel is assumed to exist between the transmitter and receiver for reporting channel state information.

In order to minimize the total transmit power and maximize the throughput subject to average BER constraint, the optimization problem is formulated as

$$\underset{\mathcal{P}_i}{\text{Minimize}} \quad \mathcal{P}_T = \sum_{i=1}^{N} \mathcal{P}_i \quad \text{and} \quad \underset{b_i}{\text{Maximize}} \quad b_T = \sum_{i=1}^{N} b_i,$$
$$\text{subject to} \quad \text{BER}_{av} = \frac{\sum_{i=1}^{N} b_i\, \text{BER}_i}{\sum_{i=1}^{N} b_i} \leq \text{BER}_{th}, \quad (1)$$

where $\mathcal{P}_T$ and b_T are the total transmit power and throughput, respectively, and BER_{av}, BER_{th}, and BER_i are the average BER, threshold value of BER, and the BER per subcarrier i, $i = 1, ..., N$, respectively. An approximate expression for the BER per subcarrier i in the case of M-ary QAM is given by[1] [4], [6]

$$\text{BER}_i \approx 0.2 \exp\left(-1.6 \frac{\mathcal{P}_i}{2^{b_i} - 1} \frac{|\mathcal{H}_i|^2}{\sigma_n^2}\right), \quad (2)$$

---

[1]This expression is tight within 1 dB for BER $\leq 10^{-3}$ [6].

where $\mathcal{H}_i$ is the channel gain of subcarrier i and σ_n^2 is the variance of the additive white Gaussian noise (AWGN).

The multi-objective optimization function in (1) can be rewritten as a linear combination of multiple objective functions as follows

$$\underset{\mathcal{P}_i, b_i}{\text{Minimize}} \quad \mathcal{F}(\mathbf{p}, \mathbf{b}) = \left\{ \alpha \sum_{i=1}^{N} \mathcal{P}_i - (1-\alpha) \sum_{i=1}^{N} b_i \right\},$$

$$\text{subject to} \quad g(\mathbf{p}, \mathbf{b}) = 0.2 \sum_{i=1}^{N} b_i \exp\left(\frac{-1.6\, \mathcal{C}_i \mathcal{P}_i}{2^{b_i}-1}\right)$$

$$- \text{BER}_{th} \sum_{i=1}^{N} b_i \leq 0, \quad i = 1, ..., N, \quad (3)$$

where α $(0 < \alpha < 1)$ is a constant whose value indicates the relative importance of one objective function relative to the other, $\mathcal{C}_i = \frac{|\mathcal{H}_i|^2}{\sigma_n^2}$ is the channel-to-noise ratio for subcarrier i, and $\mathbf{p} = [\mathcal{P}_1, ..., \mathcal{P}_N]^T$ and $\mathbf{b} = [b_1, ..., b_N]^T$ are the N-dimensional power and bit distribution vectors, respectively.

B. Optimal Bit and Power Distributions

The problem in (3) can be solved by applying the method of Lagrange multipliers. Accordingly, the inequality constraint is transformed to equality constraint by adding non-negative slack variable, $\mathcal{Y}^2$ [7]. Hence, the constraint is rewritten as

$$\mathcal{G}(\mathbf{p}, \mathbf{b}, \mathcal{Y}) = g(\mathbf{p}, \mathbf{b}) + \mathcal{Y}^2 = 0, \quad i = 1, ..., N, \quad (4)$$

and the Lagrange function $\mathcal{L}$ is then expressed as

$$\mathcal{L}(\mathbf{p}, \mathbf{b}, \mathcal{Y}, \lambda) = \mathcal{F}(\mathbf{p}, \mathbf{b}) + \lambda\, \mathcal{G}(\mathbf{p}, \mathbf{b}, \mathcal{Y}),$$

$$= \alpha \sum_{i=1}^{N} \mathcal{P}_i - (1-\alpha) \sum_{i=1}^{N} b_i$$

$$+ \lambda \left[0.2 \sum_{i=1}^{N} b_i \exp\left(\frac{-1.6\, \mathcal{C}_i \mathcal{P}_i}{2^{b_i}-1}\right) \right.$$

$$\left. - \text{BER}_{th} \sum_{i=1}^{N} b_i + \mathcal{Y}^2 \right], \quad (5)$$

where λ is the Lagrange multiplier. A stationary point can be found when $\nabla \mathcal{L}(\mathbf{p}, \mathbf{b}, \mathcal{Y}, \lambda) = 0$, which yields

$$\frac{\partial \mathcal{L}}{\partial \mathcal{P}_i} = \alpha - 0.2 \times 1.6\, \lambda \frac{b_i\, \mathcal{C}_i}{2^{b_i}-1} \exp\left(\frac{-1.6\, \mathcal{C}_i \mathcal{P}_i}{2^{b_i}-1}\right) = 0, \quad (6)$$

$$\frac{\partial \mathcal{L}}{\partial b_i} = -(1-\alpha) + \lambda \left[0.2 \exp\left(\frac{-1.6\, \mathcal{C}_i \mathcal{P}_i}{2^{b_i}-1}\right) \right.$$

$$\left. \times \left[1 + 1.6 \times \ln(2) \frac{\mathcal{C}_i \mathcal{P}_i b_i 2^{b_i}}{(2^{b_i}-1)^2} \right] - \text{BER}_{th} \right] = 0, \quad (7)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0.2 \sum_{i=1}^{N} b_i \exp\left(\frac{-1.6\, \mathcal{C}_i \mathcal{P}_i}{2^{b_i}-1}\right) - \text{BER}_{th} \sum_{i=1}^{N} b_i$$

$$+ \mathcal{Y}^2 = 0, (8)$$

$$\frac{\partial \mathcal{L}}{\partial \mathcal{Y}} = 2\lambda \mathcal{Y} = 0. \quad (9)$$

It can be seen that (6) to (9) represent $2N+2$ equations in the $2N+2$ unknowns $\mathbf{p}, \mathbf{b}, \mathcal{Y}$, and λ. Equation (9) implies

that either $\lambda = 0$ or $\mathcal{Y} = 0$; hence, two possible solutions exist, and we are going to investigate each case separately.

— *Case 1* ($\lambda = 0$): choosing $\lambda = 0$ results in an underdetermined system of one equation in $2N+1$ unknowns, and, hence, no unique solution can be reached.

— *Case 2* ($\mathcal{Y} = 0$): by choosing $\mathcal{Y} = 0$, we obtain a system $\mathcal{S}(\mathbf{x})$ of $2N+1$ equations in the $2N+1$ unknowns $\mathbf{x}$, where $\mathbf{x} = [\mathbf{p}, \mathbf{b}, \lambda]$, that cannot be solved analytically. Hence, we resort to solve such system numerically. Various numerical methods are available in the literature, e.g., the steepest descent, the Gauss-Newton, and the Levenberg-Marquardt (LM) methods [7], [8]. The steepest descent method is efficient when $\mathbf{x}$ is away from the optimal solution $\mathbf{x}_{op}$. On the other hand, the Gauss-Newton method converges fast when $\mathbf{x}$ is close to $\mathbf{x}_{op}$. The LM method takes advantage of both methods by introducing a positive damping factor μ_k to control the step size at every iteration k depending on the closeness to $\mathbf{x}_{op}$.

The LM algorithm is briefly discussed here for completeness of the presentation; however, further details can be found in [7], [8]. We start from an initial point $\mathbf{x}_0$ and initial step $\mathbf{d}_0$, then a series of points $\mathbf{x}_1$, $\mathbf{x}_2$, ... is obtained that converges towards the solution $\mathbf{x}_{op}$; hence, at iteration k one can write $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$, where $\mathbf{d}_k$ is the LM step given by [7], [8]

$$\mathbf{d}_k = -\left[\mathbf{J}(\mathbf{x}_k)^T\, \mathbf{J}(\mathbf{x}_k) + \mu_k \mathbf{I} \right]^{-1} \mathbf{J}(\mathbf{x}_k)^T \mathcal{S}(\mathbf{x}_k), \quad (10)$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{J}(\mathbf{x}_k)$ is the Jacobian matrix of the system $\mathcal{S}(\mathbf{x}_k)$, defined earlier, both at point $\mathbf{x}_k$. The damping parameter μ_k has several advantages. First, for all $\mu_k > 0$, the matrix $\mathbf{J}(\mathbf{x}_k)^T\, \mathbf{J}(\mathbf{x}_k) + \mu_k \mathbf{I}$ is positive definite, which insures that $\mathbf{d}_k$ is a descent direction and that the system $\mathcal{S}(\mathbf{x})$ has a unique solution. Second, if μ_k is large, the step value is given by $\mathbf{d}_k \simeq -\frac{1}{\mu_k} \mathbf{J}(\mathbf{x}_k)^T \mathcal{S}(\mathbf{x}_k)$ representing a short step in the steepest descent direction, and is preferred if the current iteration is far from $\mathbf{x}_{op}$. On the other hand, if μ_k is very small, then $\mathbf{d}_k$ equals the Gauss-Newton step which is suitable in the final stages of the iterations, i.e., when $\mathbf{x}_k$ is close to $\mathbf{x}_{op}$. Third, it prevents the step $\mathbf{d}_k$ from being too large when $\mathbf{J}(\mathbf{x}_k)^T\, \mathbf{J}(\mathbf{x}_k)$ is nearly singular. Furthermore, it guarantees that the step is defined when $\mathbf{J}(\mathbf{x}_k)^T\, \mathbf{J}(\mathbf{x}_k)$ is singular, in contrast to the Gauss-Newton method where the step is undefined.

C. Proposed Joint Bit and Power Loading Algorithm

To solve the problem defined in (3), we propose the following algorithm. Given an initial point $\mathbf{x}_0$, the value of $\mathcal{S}(\mathbf{x}_0)$ is calculated, and the initial step $\mathbf{d}_0$ is determined according to (10), then we set $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{d}_0$, and the process repeats. At each iteration k, if μ_k is large, i.e., small $\mathbf{d}_k$ step, then μ_{k+1} is decreased to approximate the Gauss-Newton step and converges faster to $\mathbf{x}_{op}$; otherwise μ_{k+1} is increased to approximate a steepest descent step. The algorithm converges to the optimal solution $\mathbf{x}_{op}$ at iteration k if *both* $\mathcal{S}(\mathbf{x}_k)$ and $\mathbf{d}_k$ are less than the thresholds ϵ and ε, respectively[2]. To avoid infinite loop, we set the maximum allowed number of

---

[2]If either $\mathcal{S}(\mathbf{x}_0) < \epsilon$ *or* $\mathbf{d}_0 < \varepsilon$, the algorithm stops without convergence.

iterations to k_{max} (if the number of iterations reach k_{max}, this means that the algorithm could not converge to the optimal solution $\mathbf{x}_{op}$). Once $\mathbf{x}_{op}$ is reached, $\mathbf{p}_{op}$ and $\mathbf{b}_{op}$ are obtained and the final bit and power distributions are calculated by rounding down the non-integer $\mathbf{b}_{op}$, while keeping the power distribution the same, i.e., $\mathbf{b}_{final} = \lfloor \mathbf{b}_{op} \rfloor$ and $\mathbf{P}_{final} = \mathbf{P}_{op}$. The proposed algorithm can be formally stated as follows.

Proposed Algorithm

1: **INPUT** The AWGN variance σ_n^2, channel gain per subcarrier i ($\mathcal{H}_i$), target average BER (BER$_{th}$), weighting factor α, ν_1 ($0 < \nu_1 < 1$), ν_2 ($\nu_2 > 1$), and tolerance errors ε, and ϵ.
2: Set the iteration number k to 0.
3: Pick an initial solution $\mathbf{x}_0$ and initial damping parameter μ_0.
4: **while** $\mathcal{S}(\mathbf{x}_k) > \epsilon$ and $\mathbf{d}_k > \varepsilon$ and $k < k_{max}$ **do**
5:      $k = k + 1$
6:      $\mathbf{d}_k = -\left[\mathbf{J}(\mathbf{x}_k)^T \mathbf{J}(\mathbf{x}_k) + \mu_k \mathbf{I} \right]^{-1} \mathbf{J}(\mathbf{x}_k)^T \mathcal{S}(\mathbf{x}_k)$
7:      $\mathbf{x}_{op} = \mathbf{x}_k + \mathbf{d}_k$
8:      **if** $\mu_k > \mu_{th}$[3] **then**
9:          $\mathbf{x}_{k+1} = \mathbf{x}_{op}$
10:         $\mu_{k+1} = \nu_1 \mu_k$
11:      **else**
12:         $\mu_{k+1} = \nu_2 \mu_k$
13:      **end if**
14: **end while**
15: Given $\mathbf{x} = [\mathbf{p}, \mathbf{b}, \lambda]$, find the values of $\mathbf{p}_{op}$ and $\mathbf{b}_{op}$ corresponding to $\mathbf{x}_{op}$.
16: **for** $i = 1, ..., N$ **do**
17:      **if** $b_{i,op} \geq 2$ **then**
18:         $b_{i,final} = \lfloor b_{i,op} \rfloor$ and $\mathcal{P}_{i,final} = \mathcal{P}_{i,op}$
19:      **else**
20:         $b_{i,final}$ and $\mathcal{P}_{i,final} = 0$
21:      **end if**
22: **end for**
23: **OUTPUT** $b_{i,final}$ and $\mathcal{P}_{i,final}$, $i = 1, ..., N$.

III. NUMERICAL RESULTS

We investigate the performance of the proposed algorithm in terms of achieved average throughput and average transmit power, and compare it with various bit and power loading algorithms presented in the literature.

A. Simulation Setup

As an example of a multicarrier systems, we consider an orthogonal frequency division multiplexing (OFDM) system with a total of $N = 128$ subcarriers. The channel impulse response $h(n)$ of length $N_{ch} = 5$ is modeled as independent complex Gaussian random variables with zero mean and exponential power delay profile [9]

$$\mathbb{E}\{|h(n)|^2\} = \sigma_h^2 \, e^{-n\Xi}, \qquad n = 0, 1, ..., N_{ch} - 1, \quad (11)$$

[3]For more details on the choice of μ_{th} we refer the reader to [7], [8] due to space limitations.

where σ_h^2 is a constant chosen such that the average energy per subcarrier is normalized to unity, i.e., $\mathbb{E}\{|\mathcal{H}_i|^2\} = 1$, and Ξ represents the decay factor, $\Xi = \frac{1}{5}$. Representative results are presented in this section, which were obtained through Monte Carlo trials for 10^4 channel realizations with BER$_{th} = 10^{-4}$. The LM algorithm parameters are as follows: $\mu_0 = 10^5$, $\nu_1 = 0.5$, $\nu_2 = 2$, $\varepsilon = \epsilon = 10^{-6}$, and $k_{max} = 10^4$.

B. Performance of the Proposed Algorithm

Fig. 1 illustrates the allocated bits and power, and the obtained BER for an example channel realization with SNR = 26 dB (this SNR is calculated for this channel realization only) and $\alpha = 0.5$. It can be seen from the plots in Fig. 1 that when the channel-to-noise ratio per subcarrier $\mathcal{C}_i$ is low, the corresponding subcarrier is nulled; when this is high, the optimal solution yields a non-integer number of allocated bits per subcarrier, which is not suitable for practical implementations. This value is rounded down to the nearest integer, as shown in Fig. 1 (b). The improvement in the BER, as shown in Fig. 1 (d), is due to the fact that the allocated bits are rounded down, while the transmit power is kept the same.

Fig. 2 depicts the average throughput and average transmit power when $\alpha = 0.5$ as a function of average SNR. For an average SNR ≤ 21 dB, one finds that both the average throughput and the average transmit power increase as the SNR increases, while for an average SNR ≥ 21 dB, the transmit power saturates while the throughput continues to increase. This observation can be explained as follows. For lower values of the average SNR, many subcarriers are nulled. By increasing the average SNR, the number of used subcarriers increases, resulting in a noticeable increase in the throughput and power. Apparently, for average SNR ≥ 21 dB, all subcarriers are used and our proposed algorithm essentially minimizes the average transmit power by keeping it constant, while increasing the average throughput.

In Fig. 3, the average throughput and average transmit power are plotted as a function of the weighting factor α for $\sigma_n^2 = 10^{-3}$ μW. By increasing α, more weight is given in our problem formulation to minimizing the transmit power over maximizing the throughput. In this case, the corresponding reduction in the minimum transmit power is accompanied by a reduction in the maximum throughput. This behaviour emphasizes the importance of introducing such weighting factor in our problem formulation to tune for various levels of average throughput and transmit power as needed by the wireless communication system.

C. Performance Comparison with Algorithms in [3]–[5]

In Fig. 4, the throughput achieved by the proposed algorithm is compared to that obtained by Wyglinski *et al.*'s algorithm [3] for the same operating conditions. To make a fair comparison, the uniform power allocation used by Wyglinski's algorithm in [3] is computed by dividing the average transmit power allocated by our algorithm by the total number of subcarriers. As it can be seen from Fig. 4, the proposed algorithm provides a significantly higher throughput than the scheme in [3] for
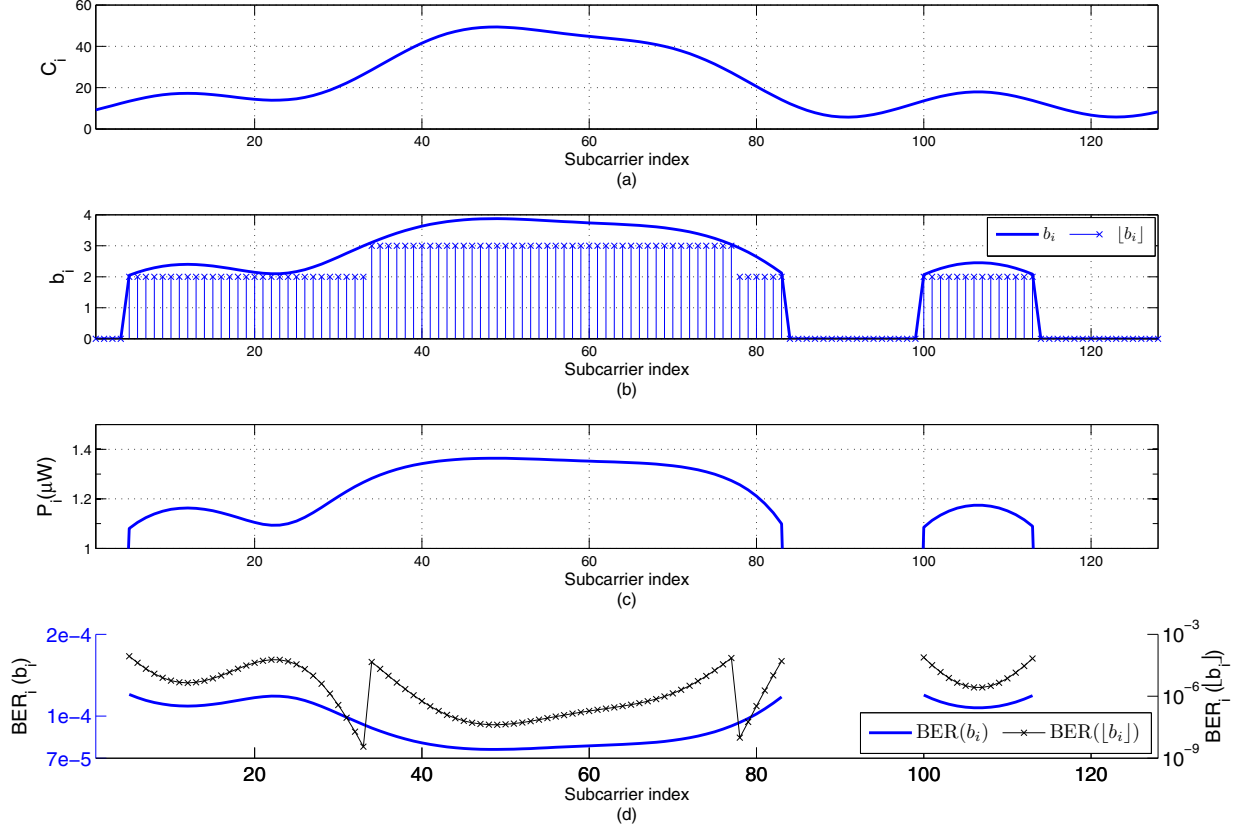
Fig. 1: Example of allocated bits and power per subcarrier for a typical channel realization with SNR = 26 dB, $\alpha = 0.5$, and $\text{BER}_{th} = 10^{-4}$.
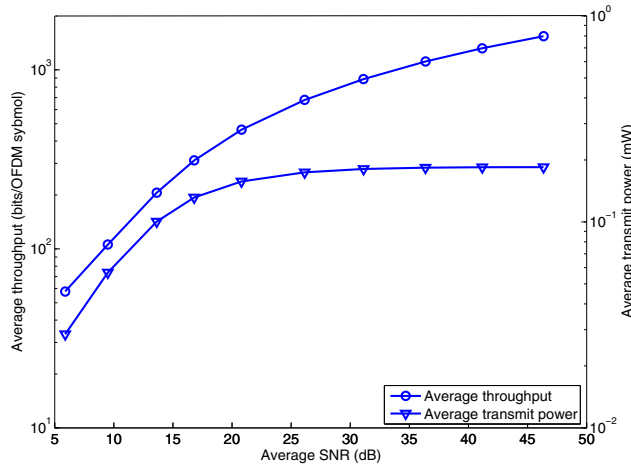

Fig. 2: Average throughput and average transmit power as a function of average SNR when $\alpha = 0.5$.

Fig. 3: Average throughput and average transmit power as a function of weighting factor α, with and without power constraint, at $\sigma_n^2 = 10^{-3} \ \mu W$.

low average SNR values. This result demonstrates that optimal allocation of transmit power is crucial for low power budgets. Furthermore, for increasing average SNR values, the average transmit power is constant as seen in Fig. 2 for values ≥ 21 dB, which in turn results in a saturating throughput for Wyglinski's algorithm. In contrast, the proposed algorithm provides an
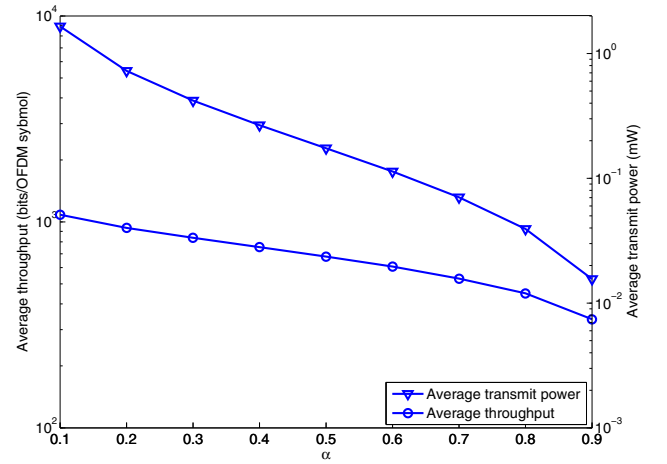
increasing throughput for the same range of SNR values.

Fig. 5 compares the average transmit power obtained by the proposed algorithm with the optimum power allocation of Liu and Tang [4], as well as with a variation called E-BER [4] that assumes an equal BER per subcarrier. To ensure that
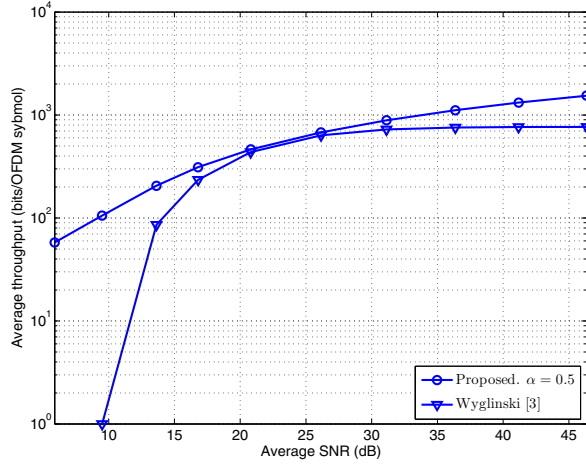
Fig. 4: Average throughput as a function of average SNR for allocation schemes that maximize the throughput and constrain the average BER.
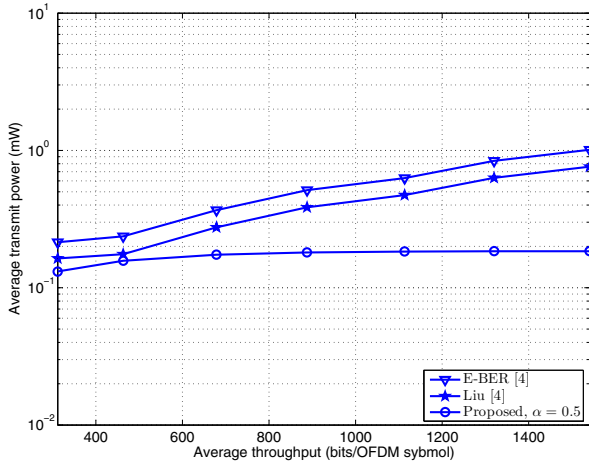

Fig. 5: Average transmit power as a function of average throughput for various allocation techniques that minimize the allocated power and constrain the BER.

the same operating conditions are considered, the fixed bit allocation per subcarrier for Liu's algorithm is set by dividing the average throughput of the proposed algorithm by the total number of subcarriers. As it can be seen from Fig. 5, the proposed allocation scheme assigns significantly less average power than both schemes in [4] to achieve the same average BER and average throughput.

The average BER obtained with the proposed algorithm is compared in Fig. 6 to that achieved by Goldfeld *et al.*'s algorithm [5] for the same operating conditions. The allocation scheme in [5] minimizes the BER while assuming that a constant constellation size is used over all subcarriers and that the transmit power is kept below a threshold. The common operating conditions are set by matching the constellation size and transmit power threshold for Goldfeld's algorithm to the average throughput per subcarrier and average transmit power of the proposed algorithm, respectively. Under these conditions, as evidenced in Fig. 6, the proposed algorithm
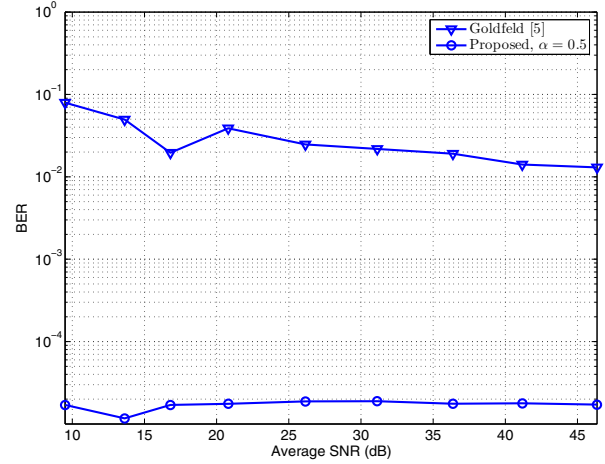

Fig. 6: Comparison of average BER achieved by the proposed allocations and by the scheme in [5] for the throughput and transmit power levels shown in Fig. 2.

was found to significantly outperform Goldfeld's algorithm. It is worthy to mention that the achieved BER by the proposed algorithm is less than the threshold average $BER_{th} = 10^{-4}$, due to rounding down the allocated bits while keeping the allocated power the same, as discussed in Fig. 1.

IV. CONCLUSION

In this paper, we proposed a novel algorithm that jointly maximizes the throughput and minimizes the transmit power given an average BER constraint. Simulation results demonstrated that the proposed algorithm outperforms various allocation schemes from the literature under the same operating conditions.

ACKNOWLEDGMENT

REFERENCES

[1] K. Fazel and S. Kaiser, *Multi-carrier and Spread Spectrum Systems: from OFDM and MC-CDMA to LTE and WiMAX*. John Wiley & Sons Inc, 2008.

[2] H. Mahmoud, T. Yucek, and H. Arslan, "OFDM for cognitive radio: merits and challenges," *IEEE Wireless Commun. Mag.*, vol. 16, no. 2, pp. 6–15, Apr. 2009.

[3] A. Wyglinski, F. Labeau, and P. Kabal, "Bit loading with BER-constraint for multicarrier systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1383–1387, Jul. 2005.

[4] K. Liu and B. Tang, "Adaptive power loading based on unequal-BER strategy for OFDM systems," *IEEE Commun. Lett.*, vol. 13, no. 7, pp. 474–476, Jul. 2009.

[5] L. Goldfeld, V. Lyandres, and D. Wulich, "Minimum BER power loading for OFDM in fading channel," *IEEE Trans. Commun.*, vol. 50, no. 11, pp. 1729–1733, Nov. 2002.

[6] S. Chung and A. Goldsmith, "Degrees of freedom in adaptive modulation: a unified view," *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1561–1571, Sep. 2001.

[7] S. Rao, *Engineering Optimization: Theory and Practice*. Wiley, 2009.

[8] K. Madsen, H. Bruun, and O. Tingleff, *Methods for Non-linear Least Squares Problems*. Informatics and Mathematical Modelling, Technical University of Denmark, 2nd edition, 2004.

[9] M. Morelli, "Timing and frequency synchronization for the uplink of an OFDMA system," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 296–306, Feb. 2004.