Computationally Efficient Optimal Power Allocation Algorithms for Multicarrier Communication Systems

Brian S. Krongold, Kannan Ramchandran, *Member, IEEE,* and Douglas L. Jones, *Senior Member, IEEE*

***Abstract*—In this paper, we present an optimal, computationally efficient, integer-bit power allocation algorithm for discrete multitone modulation. Using efficient lookup table searches and a Lagrange-multiplier bisection search, our algorithm converges faster to the optimal solution than existing techniques and can replace the use of suboptimal methods because of its low computational complexity. Fast algorithms are developed for the data rate and performance margin maximization problems.**

***Index Terms*—Discrete multitone modulation, loading algorithm, multicarrier communication systems, power allocation.**

I. Introduction

RESEARCH in multicarrier modulation has grown tremendously in recent years due to the demand for high-speed data transmission over twisted-pair copper wiring, an environment where severe intersymbol interference (ISI) can occur [1], [2]. Instead of employing single-carrier modulation with a very complex adaptive equalizer, the channel is divided into N subchannels that are essentially ISI-free independent additive white Gaussian noise (AWGN) channels, provided N is sufficiently large.

Although multicarrier modulation eliminates the need for an expensive equalizer, it creates a new problem: given some power budget, how should power and bits be allocated to each subchannel in order to maximize performance? Many algorithms for allocating power among subchannels exist; however, these methods are either suboptimal and computationally efficient [3]–[6] or optimal but slow to obtain the power allocation [7]. In this paper, we present practical and efficient discrete multitone modulation (DMT) loading algorithms that are guaranteed to converge to the optimal power allocation solution. The algorithms use efficient lookup tables and a fast Lagrange bisection search that is popular in the image compression community [8].

Paper approved by Y. Li, the Editor for Wireless Communications Theory of the IEEE Communications Society. Manuscript received March 18, 1998; revised December 30, 1998 and July 17, 1999. This work was supported in part by the Office of Naval Research under the Young Investigator Award N00014-97-1-0864 and Contract N00014-95-1-0674, and the Joint Services Electronics Program (JSEP) under Award N00014-96-1-0129. This paper was presented in part at ICC '98, Atlanta, GA, June 1998.

B. S. Krongold and D. L. Jones are with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: krongold@ifp.uiuc.edu; dl-jones@uiuc.edu).

K. Ramchandran was with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. He is now with the Department of Electrical Engineering and Computer Science, University of California at Berkeley, Berkeley, CA 94720 USA (e-mail: kannanr@eecs.berkeley.edu).

Publisher Item Identifier S 0090-6778(00)00497-9.

II. Optimal Power Allocation Problem

Maximizing channel capacity in a spectrally-shaped Gaussian channel is achieved by the well-known waterpouring distribution [9]. However, this distribution is not well suited for practical data transmission because it assumes noninteger-bit constellations, does not obey a given probability of error, and is difficult to compute. Instead, the data throughput optimization problem [3] is of more practical importance

$$\max \sum_{i=1}^{N} R_i \quad \text{subject to} \quad \sum_{i=1}^{N} P_i = P_{\text{budget}} \quad \text{and} \quad \mathcal{P}_{e,\,i} \le \mathcal{P}_{\text{error}} \qquad \forall\, i \qquad (2.1)$$

where R_i, P_i, and $\mathcal{P}_{e,\,i}$ are the rate (in bits/symbol), allocated power, and error probability, respectively, of the ith subchannel, $\mathcal{P}_{\text{error}}$ is a fixed error-probability constraint, and P_{budget} is a total power constraint. An additional constraint (of significant practical importance) is restricting R_i to be an integer number of bits/symbol. We will enforce this condition later and assume for now that R_i can be any nonnegative real number.

A. Lagrange Solution

The optimization problem in (2.1) can by reformulated as an unconstrainted optimization problem[1] by merging rate and power through the Lagrange multiplier λ

$$\min J(\lambda) = -\sum_{i=1}^{N} R_i + \lambda \sum_{i=1}^{N} P_i \qquad (2.2)$$

where $J(\lambda)$ is the Lagrange cost and $\lambda \ge 0$. Each minimum Lagrange cost for a fixed λ corresponds to the optimal power allocation for some total power budget.

For a fixed λ, the Lagrange cost is minimized when $\partial J(\lambda)/\partial P_i = 0$, for all i, or more precisely

$$\frac{\partial R_i}{\partial P_i} = \lambda, \qquad \text{for } i = 1, 2, 3, \cdots, N \qquad (2.3)$$

where R_i is a function of P_i that satisfies the error-probability constraint with equality.[2] Thus, the cost is minimized when the rates and powers for each subchannel are chosen to correspond to the point on the rate-versus-power curve with slope λ. The total power allocated for a fixed λ is obtained by simply summing the power allocated to the subchannels. The goal is to find

[1] This reformulation is equivalent provided that rate is a convex function of power, which is the case in virtually every standard class of signal constellations, including quadrature amplitude modulation (QAM). When this is not true, the solution is only optimal to within a convex-hull approximation.

[2] It can be shown that meeting the error-probability constraint with equality is optimal.

the optimal λ^* such that the total power allocated equals the given value of P_{budget} in the problem.

An additional formulation can be derived by defining the signal-to-noise ratio (SNR) of the ith subchannel to be $\text{SNR}_i = (P_i T |H_i|^2 / 2\sigma_i^2)$ and the channel-to-noise ratio (CNR) to be $\text{CNR}_i = (T|H_i|^2/2\sigma_i^2)$, where T is the symbol period, $|H_i|^2$ is the subchannel power gain, and σ_i^2 is the one-dimensional subchannel noise power. Applying the chain rule to (2.3), the following cost minimization criterion is obtained, which will be used to develop efficient loading algorithms proposed later

$$\frac{\partial R_i(\text{SNR}_i)}{\partial(\text{SNR}_i)} \text{CNR}_i = \lambda, \qquad \text{for } i = 1, 2, 3, \cdots, N. \quad (2.4)$$

B. Integer-Bit Restriction

Enforcing the restriction of integer-bit constellations, we obtain a sampled version of the continuous rate-power curves at *operating points*, which constitute the only admissible rate-power combinations. The optimal operating point for the ith subchannel for a given λ can be shown to be the point which is first "impinged upon" by a "plane-wave" of slope λ, as shown in Fig. 1(a) [8]. Rather than a unique value of λ, the discrete nature of the problem results in each operating point having a continuous range of optimal λ values associated with it as shown in Fig. 1(b).

The combination of all possible subchannel rate-power combinations summed together gives the composite rate-power function, and an example of this is shown in Fig. 2. The upper-leftmost operating points are the set of all possible optimal operating points, and the lines connecting them form the convex hull of the composite function. For a given P_{budget}, the optimal operating point is the one on the convex hull with power closest to, without exceeding, the power budget. Furthermore, the Lagrange solution will always obtain the convex hull solution and, hence, the optimal operating point.

III. Fast Algorithm for Power Allocation

We now develop the fast integer-bit loading algorithm for data-rate maximization. Because computing all the composite rate-power operating points is much too expensive, a more efficient approach is to iteratively search for a λ^*. This can be done by evaluating a chosen λ for its corresponding total power, followed by an update to get closer to an optimal solution.

A. Fast Power Allocation via Table Lookup

Each λ encountered during the search must be evaluated to determine the total power associated with it and requires computing the optimal operating point for each subchannel on the rate-power function using (2.3), summing the power allocated to the subchannels and comparing the result to P_{budget}.

Direct computation of the optimal operating point for each subchannel can be avoided by using the slope nonuniqueness property, shown in Fig. 1(b), and precomputing lookup tables of operating point slope bounds. Evaluating a given λ can easily be done for each subchannel by finding which slope range λ falls in and assigning the corresponding rate and power. The lookup tables can be generated from the rate-SNR characteristics of the channel, which are invariant to the channel conditions.

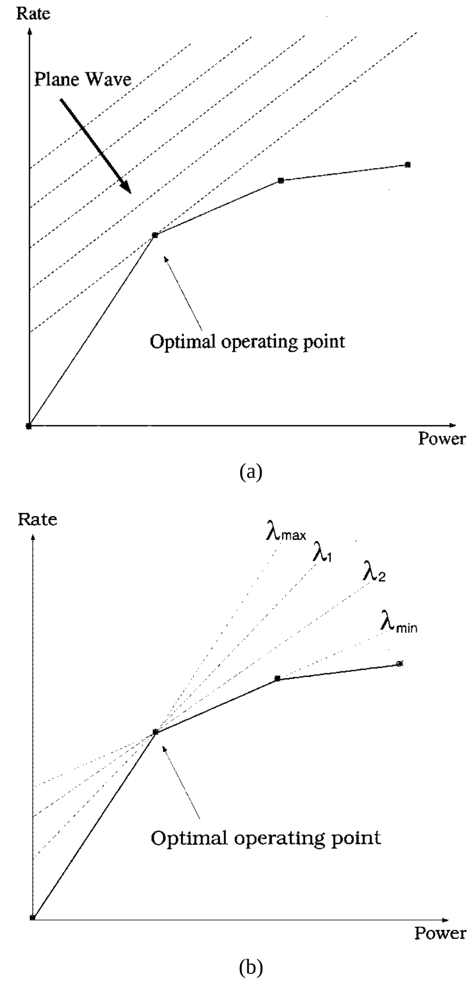

Fig. 1. (a) Depiction of a plane wave of slope λ impinging upon the rate-power convex hull.(b) Illustration of nonunique slopes for rate-power operating points.

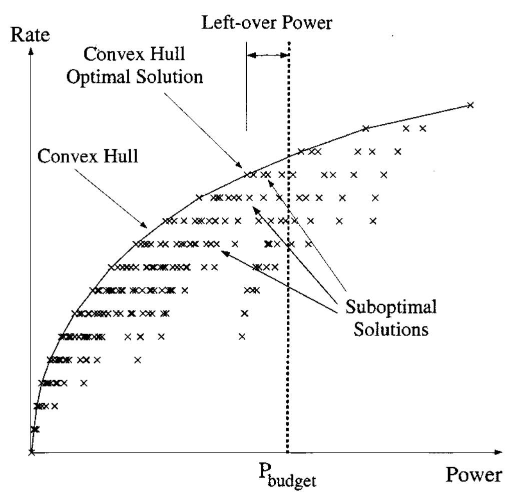

Fig. 2. Composite rate-power curve for three subchannels with up to 5 bits/symbol. It can be seen that more than one point can operate at the maximum rate with a total power less than P_{budget}. However, only one point with the maximum rate can be on the convex hull, and it is always the most power-efficient solution.

For a given λ, the Lagrange minimization formulation of (2.4) states that the optimal operating point is found using the slope $\beta_i = (\lambda/\mathrm{CNR}_i)$ of the rate-SNR function of each subchannel. Due to the discrete number of operating points, rate-SNR β ranges can also be precomputed and placed into a lookup table to avoid real-time computation. In practice, all or most of the subchannels have identical rate-SNR operating characteristics (i.e., the available signal constellations and $\mathcal{P}_{\mathrm{error}}$ constraint are the same), and the difference between them are the CNR_i's. Consequently, the β ranges for these channels will be identical, and only one lookup table need be stored,[3] resulting in a significant memory reduction.

Following the computation of β_i for a subchannel and using the lookup table to find the rate-SNR operating point, the allocated power and rates are computed as

$$P_i = \frac{\mathrm{SNR}(\beta_i)}{\mathrm{CNR}_i} \qquad R_i = R(\beta_i)$$
$$P_{\mathrm{total}} = \sum_{i=1}^{N} P_i \qquad R_{\mathrm{total}} = \sum_{i=1}^{N} R_i \tag{3.1}$$

B. Bisection Method for Fast λ Convergence

There are two major problems that can be encountered when trying to find a λ^*: fast convergence and the ability to recognize when a λ^* has been reached. A bisection method (similar to the false position method [10]) solves both of these problems and exploits the monotonic relationship between λ and P_{total} through a binary search-like procedure [8].

The bisection method uses two previously evaluated slope values λ_{low} and λ_{high} corresponding to total powers P_{low} and P_{high} (which are below and above P_{budget}, respectively) and total rates R_{low} and R_{high}. The bisection method simply lowers the gap between λ_{low} and λ_{high} by computing the following updated slope on the composite rate-power curve:

$$\lambda_{\mathrm{new}} = \frac{P_{\mathrm{high}} - P_{\mathrm{low}}}{R_{\mathrm{high}} - R_{\mathrm{low}}}. \tag{3.2}$$

The total power corresponding to λ_{new} is then evaluated. If P_{new} is greater than P_{budget}, we update λ_{high} with λ_{new} while keeping λ_{low} the same. The opposite update is done if P_{new} is less than P_{budget}. An example of the bisection method slope update is shown in Fig. 3. The slope update procedure is repeated until P_{new} equals either P_{low} or P_{high}, and the power allocation corresponding to λ_{low} is chosen to load the multicarrier system.[4]

C. Performance Margin Optimization

Another important quantity of interest in DMT systems is the performance margin[5] γ_{margin}, which is the amount of noise (in decibels) that a system can tolerate while still operating under the bit-error-probability constraint [4]. The performance margin optimization problem for a given target rate is as follows:

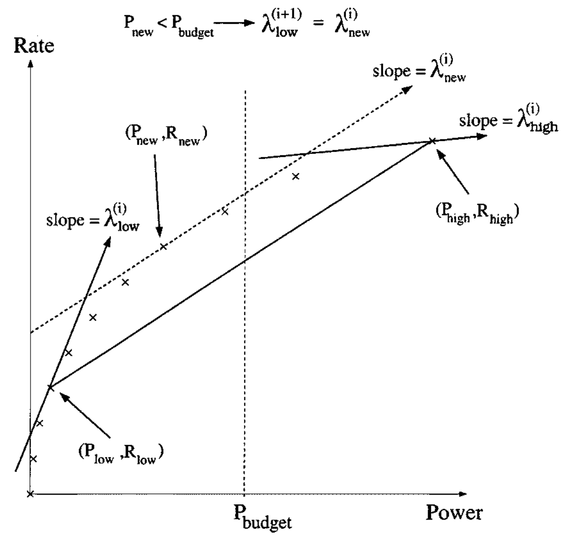

Fig. 3. Illustration of bisection method slope update.

$$\max \gamma_{\mathrm{margin}}$$

subject to

$$\sum_{i=1}^{N} R_i = R_{\mathrm{target}}, \ \sum_{i=1}^{N} P_i = P_{\mathrm{budget}}$$

and

$$\mathcal{P}_{e,\,i}, \leq \mathcal{P}_{\mathrm{error}} \qquad \forall\, i. \tag{3.3}$$

Since performance margin is simply a scaling of the zero-margin allocated power in each subchannel by a constant amount, the minimum power allocation needed to meet the rate target with zero margin can be found, and the resulting powers can be scaled to utilize the total power budget. This new optimization problem is as follows:

$$\min \overline{P}_{\mathrm{min}} = \sum_{i=1}^{N} \overline{P}_i$$

subject to

$$\sum_{i=1}^{N} R_i = R_{\mathrm{target}}$$

and

$$\mathcal{P}_{e,\,i} \leq \mathcal{P}_{\mathrm{error}} \qquad \forall\, i. \tag{3.4}$$

In this case, only the convex hull operating points of the composite rate-power function can be optimal solutions because there is no leftover power as in the rate maximization problem. An algorithm very similar to the rate maximization algorithm can be used to solve (3.4), with the difference that λ is updated so that the total rate converges to R_{target} [11]. Once this has been done, the final power allocation is computed as $P_i = \overline{P}_i \cdot \gamma_{\mathrm{margin}}$, where $\gamma_{\mathrm{margin}} = P_{\mathrm{budget}}/\overline{P}_{\mathrm{min}}$.

[3]For any remaining subchannels with different rate-SNR characteristics, different lookup tables need to be defined.

[4]It is highly improbable that P_{new} will exactly equal P_{budget}, but if this does occur, the algorithm allocates power according to λ_{new}.

[5]For example, if a single channel requires a 12-dB SNR to operate at some data rate with $\mathcal{P}_{\mathrm{error}}$, then providing 18 db of SNR results in a +6-dB performance margin.

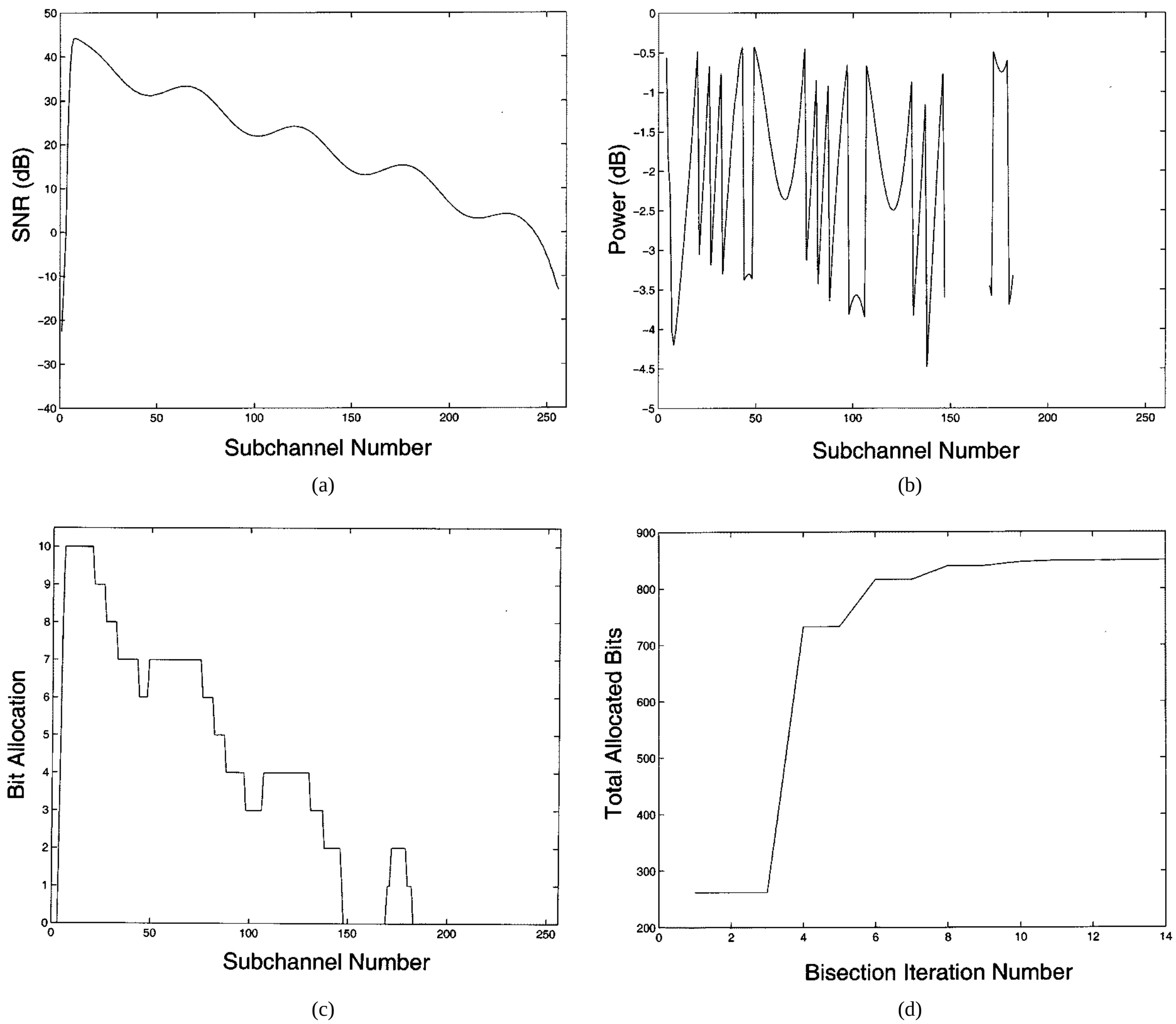

Fig. 4. (a) SNR (in decibels) for 256 subchannels. (b) Optimal power allocation. (c) Optimal bit allocation.(d) Total bits allocated versus bisection iteration number.

TABLE I
ALGORITHM PSEUDOCODE

1. Obtain initial values of λ_{low} and λ_{high}. Use (3.1) to ensure that $P_{low} \leq P_{budget} \leq P_{high}$.

2. Apply the bisection method to obtain $\lambda_{new} = \frac{P_{high}-P_{low}}{R_{high}-R_{low}}$

3. Determine the lookup parameter β_i for $i = 1, \cdots, N$ and obtain rate-power characteristics P^i_{new} and R^i_{new} from the table.
$$\beta_i = \frac{\lambda_{new}}{CNR_i}, \quad R^i_{new} = R(\beta_i), \quad P^i_{new} = \frac{SNR(\beta_i)}{CNR_i}$$

4. Compute total rate $R_{new} = \sum_{i=1}^{N} R^i_{new}$ and total power $P_{new} = \sum_{i=1}^{N} P^i_{new}$ and compare as follows:

   ⇒ IF $P_{new} = P_{low}$, P_{high}, or P_{budget}, go to step 5.

   ⇒ ELSE IF $P_{new} < P_{budget}$, set $P_{low} = P_{new}$ and $R_{low} = R_{new}$. Store P^i_{low} and R^i_{low}. Go to step 2.

   ⇒ ELSE set $P_{high} = P_{new}$, and $R_{high} = R_{new}$. Go to step 2.

5. Power and rate allocation is complete. Assign P^i_{low} and R^i_{low} to the ith subchannel.
In the improbable case of $P_{new} = P_{budget}$, assign P^i_{new} and R^i_{new} to the ith subchannel.

D. Algorithm Implementation

Pseudocode for the rate maximization algorithm using rate-SNR lookup tables is listed in Table I. Initial loading of the system may require initial λ_{low} and λ_{high}, which are far from the optimum values to ensure that the condition in step 1 is met. Simple worst case initializations are the point with zero rate and power and the point with maximum rate and power ($\lambda = 0$). Prior knowledge of typical channels can help the designer choose initial λ values closer to the range of optimal values, which will allow the algorithm to converge faster. Convergence of the bisection method to a λ^* takes approximately $\log_2(NM)$ iterations. Assuming $1/\text{CNR}_i$ is precomputed, each λ iteration requires at most $2N$ additions, 1 division, $2N$ multiplies, and N lookup table evaluations. The resulting computational complexity of the algorithm is $O(N \log N)$. Knowledge of previous iteration lookup results can be used to *drastically* reduce complexity [12] as some subchannels converge rather quickly, and most subchannel rate-power assignments do not change in the final one-third or so iterations.

Fig. 4 shows an example of our algorithm for a test channel where the available signal constellations were 0–10 bits/symbol QAM and a symbol error probability constraint of 10^{-7} was imposed on each subchannel. The algorithm converged very quickly[6] to the optimal solution with 14 bisection search iterations using very conservative initial low and high slope values.

[6]After only eight iterations, 98.8% of the optimal rate is achieved.

For tracking scenarios, when the channel conditions change only slightly, an optimal λ value may not be very different from the previous optimal one. Therefore, initial low and high λ values can be chosen much closer to obtain faster convergence.

IV. Comparisons and Conclusions

The Hughes–Hartog algorithm [7] is an optimal loading algorithm which achieves the solution by adding one bit at a time to the channel requiring the smallest additional power to increase its rate. Whereas this technique can be used to solve both data rate and margin maximization, the algorithm requires an intensive amount of sorting and converges very slowly in practical DMT scenarios [3]. Our algorithms achieve exactly the same solutions and are cheaper to implement.

The algorithm in [4] attempts to maximize margin in a suboptimal fashion that relies on rounding to integer rates. Another disadvantage of this algorithm is its use of the SNR gap approximation [2] to allocate bits to its subchannels. Furthermore, in the final part of the algorithm, it requires a modest amount of sorting to subtract or add bits one at a time to meet the target bit rate. The overall complexity is less than Hughes–Hartog, but is approximately the same or slightly more than the proposed margin algorithm in this paper.

The algorithm in [6] attempts to maximize the subchannel SNR's rather than the margin and again relies on rounding. Whereas this is a different criterion for loading, the resulting allocation should be extremely close if not identical. The results in [6] in show improvement of overall SNR compared to [4] as well as some reduction in complexity. As in [4], it uses a modest amount of sorting to subtract or add bits one at a time, which may be expensive if the initial part of the algorithm is too far from the target rate. The overall complexity of the algorithm is dominated by searches and additions, but the operation count will typically be on the same order as the margin algorithm in this paper.

Using the same channel SNR's shown in Fig. 4(a), optimal margin maximization is compared to the two algorithms described above. In the case of [6], the final rate assignment is used and power is allocated to meet the 10^{-7} symbol-error-probability constraint. The target data rate is 500 bits/symbol which corresponds to 2 Mbp/s for a multicarrier symbol rate of 4 kHz, and QAM signal constellations with a range of 2–10 bits/symbol are employed.[7] Table II shows the resulting margins for the example channel and with AWGN levels of +3, –3, and –6 dB from the original. As can be seen, the method of [6] is very close to the optimal margin, with about a 2% margin differential for all conditions. However, the method of [4] has significant performance variation with different amounts of AWGN added to the channel. This is due to the SNR gap becoming even more suboptimal as subchannel SNR's decrease and the number of assigned bits in a subchannel also decreases.

TABLE II
COMPARISON OF MARGIN MAXIMIZATION ALGORITHMS

Loading Algorithm	Margin in dB			
	+3 dB AWGN	0 dB AWGN	-3 dB AWGN	-6 dB AWGN
Proposed Algorithm	9.4636	12.4636	15.4636	18.4636
Algorithm in [6]	9.3727	12.3797	15.3797	18.3797
Algorithm in [4]	7.3727	11.7114	15.3449	18.4602

The rate maximization algorithms in [3] and [5] are also based upon the SNR gap approximation combined with rounding, and they suffer from its drawbacks as well. Both methods require moderate amounts of sorting, in addition to mathematical computation, and the resulting complexities are probably higher than the proposed optimal algorithm.

Although we developed our algorithms specifically for integer-bit constellations, they can be used for systems containing noninteger-bit constellations as well. In fact, the algorithms will work for any set of discrete points on the rate-SNR curve provided the function is convex, which it almost always will be. Thus, these algorithms should be considered optimal loading algorithms for any discrete set of available signal constellations.

References

[1] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.
[2] J. M. Cioffi, "A multicarrier primer," ANSI T1E1.4 Committee Contribution, Nov. 1991.
[3] P. S. Chow, "Bandwidth optimized digital transmission techniques for spectrally shaped channels with impulsive noise," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1993.
[4] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 43, pp. 773–775, Feb./Mar./Apr. 1995.
[5] A. Leke and J. M. Cioffi, "A maximum rate loading algorithm for discrete multitone modulation systems," in *Proc. IEEE GLOBECOM'97*, Phoenix, AZ, Nov. 1997, pp. 1514–1518.
[6] R. F. H. Fischer and J. B. Huber, "A new loading algorithm for discrete multitone transmission," in *Proc. IEEE GLOBECOM'96*, London, U.K., Nov. 1996, pp. 724–728.
[7] D. Hughes-Hartogs, "Ensemble modem structure for imperfect transmission media," 4 679 227 (July 1987), 4 731 816 (March 1988) and 4 833 796 (May 1989).
[8] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 2, pp. 160–175, Apr. 1993.
[9] R. G. Gallager, *Information Theory and Realiable Communication*. New York: Wiley, 1968.
[10] D. G. Luenberger, *Introduction to Linear and Nonlinear Programming*, 2nd ed. Reading, MA: Addison-Wesley, 1989.
[11] B. S. Krongold, "Power and bandwidth optimization for multicarrier communication systems," M.S. thesis, Univ. Illinois at Urbana–Champaign, 1997.
[12] B. S. Krongold, K. Ramchandran, and D. L. Jones, "Section division operating point slope determination method for multicarrier communication systems," pending.

[7] Both the proposed algorithm and [6] can achieve slight performance gains by allowing 1 bit/symbol as a signaling choice. However, [4] often yields poorer solutions in this case.