A Framework for Efficient Rate-Power Allocation for OFDM in a Composite-Fading Environment

Kamau Prince, Brian Krongold and Subhrakanti Dey
ARC Special Research Centre for Ultra-Broadband Information Networks
Affiliated Program of National ICT Australia
University of Melbourne, Australia
Email: k.prince@ee.mu.oz.au, bsk@unimelb.edu.au, s.dey@ee.mu.oz.au

Abstract—We propose a framework to optimally allocate transmission resources for a digital single-user multichannel transmission link operating in a dynamically-fading wireless environment. Beginning with the statistical properties of fading phenomena observed in such channels, we demonstrate how convex optimisation theory may be harnessed for this purpose. Examples are presented to illustrate how service quality criteria may be used to dynamically identify operating points for the allocation algorithm.

I. INTRODUCTION

Multitone communications is being intensely researched due to its potential for supporting increased bandwidth demands required in new generations of wireless networks. Contemporary implementations of such technologies (e.g. IEEE 802.11b WLAN, High-Definition television DVB) allocate power equally across sub-carriers for each transmitted symbol. This may be appropriate in broadcast scenarios, but in point-to-point OFDM transmission, it may be possible to allocate power and rate according to the channel conditions on the link. We propose a framework in this paper to improve the overall efficiency of such multitone systems by exploiting the statistical nature of the wireless fading environments in order to optimize the dynamic allocation of transmission resources.

We structure the rate-power allocation as the solution to a convex optimization problem, having constraints derived from pre-determined quality of service (QoS) metrics. We then apply this framework to illustrate how to optimize rate-power allocation for transmission over a multichannel link operating in a fading environment, subject to service quality parameters such as error probability or outage probability.

The paper opens with a development of the structure supporting the framework. We characterize the symbol transmit power required to fulfill average received-SNR criteria that supports single-user communication as a particular QoS, such as symbol (or bit) error probability (SEP/BEP) or symbol outage probability (SOP). The scope is then widened to a single-user wireless OFDM communication link to illustrate how this framework is applied to optimally allocate rate-power in a dynamically-fading environment. Results are then presented for the operation of this link in a sample Rayleigh fading environment, in which each of the sub-symbols is selected from a closed set of pre-defined modulation depths. Outage probability is used as a QoS constraint and the resulting allocation is compared to the outage probability for a flat-power allocation which uses the same total power.

II. LOADING PROBLEM

A. Background

We model a multitone channel as an aggregation of N narrowband transmission sub-channels, each receiving a signal from a transmitting node and multiplied by a complex-valued scalar , . Hence, the composite channel modulates X by α := {αi} : i ∈ {1, · · · , N}, where α may be time varying. Every transmitted symbol, X ∈ {CN} may be represented by an N-length vector, comprised of the sub-symbols transmitted in each subchannel. We also assume that the total transmission system noise is known at the receiver.

We assume that the system of interest operates in a frequency-selective environment in which transmission effects result in unequal values of αi across subchannels. This phenomenon creates the loading problem [1] [2] [3] where OFDM system performance can be optimized by intelligently allocating (or loading) rate and power to each subchannel. We further assume that the correlation between fade coefficients over two given subchannels, E[αi, αj] becomes negligible once the frequency separation between the jth and jth subchannels exceeds the coherence bandwidth of the channel [4]. In this sense, there is more frequency selectivity which can be exploited through an intelligent loading algorithm to achieve rate increases and/or power savings in multitone links. Several attempts have been made to effectively allocate rate and power resources to maximize OFDM performance, but almost all are suboptimal in some sense (e.g. rate rounding [3], SNR-gap approximation [2]), while many are not very computationally efficient and rely on such things as sorting of all subchannel gains or require too many multiplications and even divides. In this paper we base our new framework on a previous work [1], [5]–[6] by one of the authors which is both theoretically optimal and extremely efficient to implement.

A key property required for optimality is SNR being a positive, convex function of rate. In contemporary digital modulation schemes operating in a narrowband AWGN environment, this is almost always the case. The resulting resource allocation problem evaluates the convex relationship between required transmit symbol power and associated rate under QoS...
constraints, and optimizes an objective function under any additional resource constraints.

B. Loading Algorithm

We wish to develop a practical resource allocation algorithm based upon [1] for operation in a time-varying multi-tone environment. Consider a single-user digital multitone (DMT) system with a broadband channel divided into \( N \) narrowband sub-channels. Each transmitted symbol, \( \mathbf{X} \in \mathbb{C}^{N} \) is an \( N \)-length vector, comprised of the sub-symbols transmitted in each subchannel \( i \in \{1, \ldots, N\} \). For now, we further assume that the value of the fade coefficient, \( \alpha_k \) during transmission of the \( k^{th} \) symbol (\( \mathbf{X}_k \)), is known at the transmitter. (We pose that use of an estimate, \( \hat{\alpha}_k \), would also be feasible.) The resource allocation algorithm is to assign a transmission rate, \( R_i \), and transmit power, \( P_i \) for all \( i \) and the given transmitted symbol, \( \mathbf{X} \). The rate \( (R_i) \) selected for each channel will affect the modulation depth of the \( i_{th} \) sub-symbol, and hence the modulation scheme used to transmit this sub-symbol.

We assume the algorithm assigns rate \( (R_i) \) from a closed set of cardinality \( L \) and power \( (P_i) \) such that for all \( i \),

\[
P_i \geq 0, \quad R_i \in \{0, r_1, r_2, \ldots, r_{L-1}\}
\]

The total rate of each symbol, \( R_T \) is defined by (2), and the total power, \( P_T \) of each symbol is given in (2).

\[
R_T = \sum_{i=1}^{N} R_i; \quad P_T = \sum_{i=1}^{N} P_i
\]

The algorithm optimizes resources for each transmitted symbol using the channel data (\( \alpha_k \)'s) and provides a set of channel rate/power parameters, \( \{R_i, P_i\} \) (\( i \in 1, \ldots, N \) ) to satisfy additional constraints, say e.g.:  
- Service Quality threshold (\( \Omega_{th} \)), in addition to
- Maximum Transmitted Power (\( P_{MAX} \)), and/or
- Target Symbol Rate (\( R_{TARGET} \))

In [1], symbol (or bit) error probability was considered for the scenario where each subchannel experiences AWGN and a fixed, flat fade by a complex scalar. In this paper, we consider dynamically fading environments and additional QoS constraints may be necessary in addition to symbol/bit error probabilities (averaged over the fading distribution). In the presence of fast-fading, outage probability is another noteworthy QoS measure that will be considered later in the paper.

We first outline the algorithm developed in [1] which forms the basis for the extended wireless framework. In addition to its optimality properties, the approach is very efficient computationally with \( O(N) \) running time. Two important (dual) loading problems were considered: Rate Maximization and Margin Maximization.

Rate maximization maximizes the overall data rate per symbol, \( R_T \) subject to constraints on the total power, \( P_T \) and some error probability constraint on each subchannel. Hence,

\[
\max \quad R_T \quad \text{subject to} \quad P_T \leq P \quad \text{and} \quad \mathcal{P}_{i,k} \leq \mathcal{P}_{th}; \forall i
\]

where \( \mathcal{P}_{i,k} \) is the error probability (bit or symbol) for the \( i^{th} \) subchannel of \( \mathbf{X}_k \), and \( \mathcal{P}_{th} \) is the target maximum error constraint. In general, we try to meet error probability constraints with strict equality in order to not waste power. Margin maximization has a target total rate \( (R_{DES}) \) that must be met, as well as a total power constraint \( (P_{MAX}) \). Upon achieving \( R_{DES} \) with some power \( P_T \), any remaining power is used to add margin \( \varphi \), or scale, the \( P_i \)'s so that power \( P_{MAX} \) is used. Maximizing this margin is equivalent to minimizing \( P_T \) required to achieve \( R_{DES} \), resulting in the following dual problem to (3):

\[
\min P_T \quad \text{subject to} \quad R_T = R_{DES} \quad \text{and} \quad \mathcal{P}_{i,k} \leq \mathcal{P}_{th}; \forall i
\]

The extra power can then be scaled onto each subchannel by the factor \( \varphi = P_{MAX}/P_T \).

C. Optimal Loading

The optimal loading approach in [1] was first written by assuming that power is a continuous, strictly increasing convex function of rate for a fixed error probability constraint met with equality. The constrained optimization problems are converted into unconstrained ones by introducing a Lagrange multiplier, and in the margin maximization problem, the new problem is

\[
\min J(\lambda) = \sum_{i=1}^{N} P_i + \lambda \left( R_{DES} - \sum_{i=1}^{N} R_i \right)
\]

where \( J(\lambda) \) is the Lagrange cost and \( \lambda \geq 0 \). Each minimum Lagrange cost for a fixed \( \lambda \) corresponds to the minimum power for some total rate. The goal then is to find an optimal \( \lambda^* \) which achieves \( R_{DES} \). For a given \( \lambda \), the minimum \( J(\lambda) \) is achieved with a same-slope solution \( \frac{\partial J}{\partial R_i} = \lambda \) for each subchannel [1]. If the \( \lambda \) is less than the derivate at zero, then no resources are allocated to that subchannel.

In practical systems, \( R_i \) is chosen from a discrete set, and the rate-power function is discretized. Derivatives are relaxed into differentials between operating points \( (R_i, P_i) \), and the Lagrangian approach is still optimal. For a given \( \lambda \), the optimal operating point in a given subchannel is where a line of slope \( \lambda \) is tangent to the convex hull. The discretized nature of the problem makes it simpler to solve as each operating point has a continuous interval of \( \lambda \) values associated with it, i.e.,

\[
R_i = \begin{cases} \quad r_j, & d_i(r_j) \leq \lambda < d_i(r_{j+1}) \\ \lambda \leq d_i(r_j) \end{cases}
\]

where \( d_i(r_j) \) are the differentials (\( \Delta \)power)/(\( \Delta \)rate) between adjacent operating points at rates \( r_{j-1} \) and \( r_j \). This motivates the use of lookup tables to accelerate the loading, but requires also requires computing a table for each subchannel.

We can also consider a required SNR\( r_{req}(r_j) \), to allow a specific subchannel rate \( r_j \) under the error probability constraint. Defining SNR as

\[
SNR_i = \frac{P_i \left| H_i \right|^2}{2\sigma_i^2} = P_i \cdot CNR_i
\]

we can see that a required \( P_i \) to meet \( SNR_{req}(r_j) \) is inversely proportional to the channel-to-noise ratio, \( CNR_i \). We can define SNR-rate operating points as well, and a lookup table with differentials \( \beta(r_j) = d_i(r_j) \cdot CNR_i \) to facilitate loading.
No \(i\) subscript is needed because all subchannels are assumed to have the same SNR-rate characteristics. The lookup table approach can be simplified to [5]:

\[
R_i = \begin{cases} 
  r_j, & \frac{\lambda}{\beta(r_j+1)} \leq \frac{1}{\text{CNR}_i} < \frac{\lambda}{\beta(r_j)} \\
  0, & \frac{1}{\text{CNR}_i} \leq \frac{\lambda}{\beta(r_j)}
\end{cases}
\]  

(8)

The lookup table boundaries are no longer subchannel dependent, and a huge computational reduction results. With each new tested \(\lambda, L - 1\) lookup table boundaries are computed and \(1/\text{CNR}_i\) is used to find the allocation.

The key to this simplification is the **scalability** across the power-rate operating points between subchannels for the case of fixed, flat-faded subchannels corrupted by AWGN. This scalability property can be generalized to the following:

**Definition 1:** A loading problem has scalability across subchannels if the QoS can be written as a function of only \((R_i = r_j)\) and \((\zeta_i \cdot P_i)\) for all \(i, j\), where \(\zeta_i\) is a known, given parameter which completely characterizes the exact or statistical nature of each subchannel.

For a deterministic channel, \(\zeta_i\) is simply \(\text{CNR}_i\) as above. In a dynamically-fading wireless environment, this property may still hold with \(\zeta_i\) representing the statistical state of a subchannel (perhaps in terms of an average SNR).

Significantly more detail of the fast, optimal loading algorithm can be found in [5]. The important aspects of it to this paper are its optimality, basic structure, as well as its computational efficiency and use of the subchannel scalability property. We have not yet mentioned coding within the loading structure, which is an important part to OFDM wireless transmission. The effect of a code and its gain may be incorporated into the power-rate operating points and the allocations.

**III. WIRELESS CHANNELS**

We consider a wideband channel supporting \(N\) narrowband channels, over which some composite fading/shadowing effects are observed [7] [8], producing frequency-selective multipath fading. Hence, the channel-induced gain (\(\alpha_i\)) imposed on the \(i^{th}\) subchannel of a given symbol transmission may be described by:

\[
\alpha_i = F_i \cdot G_i \cdot H_i
\]  

(9)

where \(F_i\) represents the rapidly-varying fading coefficient over the channel, \(G_i\) indicates the more slowly-varying variations in path gain, and \(H_i\) denotes the path-length dependent loss. We assume that \(F_i\) and \(G_i\) take values from independent random processes, where all \(G_i^{'}s\) vary on a time scale which allows for accurate tracking or prediction, and that the values assumed by \(F_i\) vary unpredictably. Without loss of generality, we collapse \(H_i\) coefficients into a composite slow-varying \(G_i\).

The effects of fluctuations in channel gain as observed by the various sub-channels is such that the transmitter must vary the sub-symbol power assignments \((P_i)\), in order to produce a consistent receiver symbol-to-noise power ratio (SNR), which is related to the service quality delivered over that channel. This is consistent with the scalability property (and its associated complexity reduction) described in the previous section, and can be applied to the loading problem in a composite-fading environment. Our approach recognizes treatment of mobile channels as falling into two main categories, as shown below.

1. Slowly-Developed Fading in which the \(G\) terms dominate, and we assume without loss of generality that the mean value of the fast fading components is unity \((E[F_i] = 1\) for all \(i\)). Depending on the transmission factors, such an environment will produce highly time-correlated values of \(\alpha\), which may vary little over several symbol intervals. Channel estimation may be implemented, if the fading conditions vary slowly enough. For less slowly-developed fading, well-defined models exist for the time-developing behavior in such environments [4]. These models have been developed to provide reasonable short-to medium-term prediction of future values of \(\alpha\) [9], and which have been integrated with predictive channel loading [10].

In such cases, the algorithm is given information on the instantaneous channel conditions, and is able to optimally allocate resources for the channel. Error probability \((P_E)\) will likely be used as indicate the QoS for such systems.

2. Rapid Fading, with significant \(F_i\) terms which we assume to be identically distributed with some fading distribution. In this environment, there is additional uncertainty about the values assumed by the instances of \(\alpha\). Due to the decreased time autocorrelation of \(\alpha\), the time delay between channel measurement, feedback, estimation, and loading optimization becomes significant, and reduces the certainty with which the transmission channel estimate is presented. In this case, channel prediction/estimation techniques are less reliable. We propose “outage” as an additional channel service quality metric for such channels and define an “outage” to occur whenever the receiver SNR falls below the value required to fulfill a desired \(P_E\). The target of the algorithm in such cases is to allocate transmission resources to satisfy power, rate and \(P_o\) constraints.

**IV. OFDM LOADING IN WIRELESS CHANNELS**

We now apply the loading framework to a wireless single-user OFDM link. Consider a wideband channel, divided into \(N\) subchannels, operating in the presence of frequency-selective fading. We model each of the subchannels of the transmission link as having an associated (multiplicative) time-varying gain amplitude \((\alpha_i)\) as defined above. Each \(\alpha\) parameter is defined by some statistical density function, and therefore, manipulation of this underlying function leads to an approach that is optimal in an ergodic sense. In this way, the system design no longer has to be dependent upon the subtleties of the phenomena producing the fade effects, but could rather be constructed utilizing a set of long-term coefficients. We additionally assume that information regarding the transmission noise process is available at the receiver. We customize the
framework for application in the OFDM environment, subject to the conditions assumed for the channel variations.

The time-varying nature of the channel gains motivates the development of an algorithm to perform the resource allocation in the minimum possible time. Additionally, since the typical fading scenario experienced in a narrowband wireless channel imposes a practical limit on the transmission rate (arising from limitations on feasible transmit power levels), then the possible \{r_j, P_j, \text{MIN} \mid j \in 1, \cdots, L \} will be limited to a few elements. This result encourages the use of such an algorithm, since an effective one requiring little overhead computation time would provide an operating point more appropriate for the channel estimate provided, or which would reduce the time horizon required for the channel prediction.

A. Slowly-Developing Fading

It is assumed that the trackable components \( (G_{i \in \{1, \cdots, N \}} \)'s) of the channel gain are dominant (i.e. \( F_{i \in \{1, \cdots, N \}} \approx 1 \)), and that possibly \( \alpha_i \neq \alpha_j \neq \alpha_i \) for a given symbol interval. The relatively slow development of the channel gain \( \alpha \) allows the transmitter to obtain a reliable estimate to use for resource allocation. (This estimated gain vector may be directly obtained by channel measurement, or indirectly provided through channel prediction). Furthermore, for sufficiently slowly-varying conditions, the algorithm may not need to be run for every transmitted symbol provided that sufficiently high autocorrelation of the channel-varying conditions are observed.

B. Composite Fading with Fast and Slow Components

In this environment, we assume the presence of a variable slowly-fading channel gain coefficient \( (G) \), the instantaneous values of which may be reliably estimated/predicted, and which therefore provide a deterministic distortion of the transmitted symbols. However, this effect is accompanied by other, more rapidly-varying processes, modeled as a fast-fading gain coefficient \( (F) \), which reduces the correlation of the overall gain factors between successive transmission intervals. We assume that the realizations of \( F_i \) coefficients are drawn from a known statistical distribution identical for all \( i \). The random nature of \( F_i \) increases the uncertainty associated with each estimate \( \hat{\alpha} \), and motivates the inclusion of a more conservative approach to resource allocation.

Performance over the previous (slowly-fading) channel was parameterized in terms of the \( \mathcal{P}_E \). For this model, the randomness introduced by the faster fades prevents the derivation of an exact error probability. Two possible schemes exist for specifying the QoS delivered in such environments. The first uses of an average error probability statistic \( \mathcal{P}_E \), which may be evaluated by taking the relationship between the receiver SNR and the \( \mathcal{P}_E \), and integrating it over the range of the fading PDF. The second is to evaluate an outage probability statistic \( \mathcal{P}_0 \), which indicates the proportion of time for which the SNR exceeds the minimum value required to satisfy a certain \( \mathcal{P}_E \).

In each subchannel, the average error probability \( \mathcal{P}_E \) for any fixed \( P_i \) is given by

\[
\mathcal{P}_E = \int_{F} [\mathcal{P}_E | F] \cdot p_F(F) \, dF
\]  

If each possible \( F_i \) scales all symbols uniformly, regardless of \( R_{j \in 1, \cdots, L} \), then averaging the resulting error probability equations and evaluating this function for a fixed \( P_E \) becomes the sum of an infinite number of convex functions relating SNR to rate for the given \( P_E \). The result is a convex SNR-rate relationship (for any given \( P_E \)), which must be optimized for resource allocation. The Lagrange methods are therefore optimal for performing this allocation.

Alternatively, if \( P_0 \) is included as a QoS metric, then the \( P_i \) on each subchannel must be optimized to satisfy the outage constraint

\[
\Pr[\mathcal{P}_E \leq \mathcal{P}_E, \text{TAR}] \leq \mathcal{P}_0, \text{TAR}
\]  

Since the fade equally modulates all transmitted symbols, evaluation of the inverse PDF of \( F_i \) returns the receive SNR (and hence, transmit power) scaling factor required to achieve the desired outage probability. Effecting the required scaling linearly transforms the relationship for SNR/\( P_E \) in AWGN for a particular fade intensity, producing a convex objective function. The Lagrange methods are also therefore optimal for allocating transmission resources in this environment.

V. EXAMPLES/SIMULATIONS

The results obtained above were used to evaluate the relative performance of the margin-maximization algorithm (MMA) for dynamic allocating resources in a composite fading channel. It is assumed that the dynamic resource allocation algorithm has access to the development of the slowly-varying gain component and can allocate accordingly. In addition to this component, the model was developed with a faster-varying fading process, which was untrackable, the values of which could be modeled as realizations of a Rayleigh random variable. OFDM transmission was used having 128 subcarriers with a target \( R_T \) of 256 bits per symbol.

The minimum power was calculated to satisfy the error and outage constraints, thus providing unity margin \( (\varphi) \) for the MMA. The flat-power allocation was done by uniformly dividing equal total transmit power across subchannels. Additionally, both systems were provided with similar average target error and outage probability criteria. The channel model used for these analyses is presented in Figure 1. The results obtained are presented graphically in Figure 2. It was observed that the average outage probability obtained with the MMA algorithm was able to meet the constraints, whereas the outage probability for the flat-power allocation was slightly higher than target at 6.1%, indicating inferior performance. More significantly, it was observed that the MMA was able to equalize outage probabilities across the subchannels, indicating uniform performance. In contrast, the flat-power allocation produced

\(^1\)Varying the \( P_i \) scales the relationship linearly, but does not otherwise affect the nature of the optimization problem.
significantly variable results for outage probability for the sub-symbols transmitted along different subcarriers, for the same transmit power $P_T$. These results indicate that the MMA algorithm is superior to flat-power allocation in optimizing resources to achieve a stable communication link in a fading transmission environment. The rate allocation is presented in Figure 3.

VI. CONCLUSION

The framework proposed in this paper extends our previous loading algorithm to deal with wireless systems, where a time-varying channel and latency in the channel-state feedback and loading time render the allocation to be suboptimal. Using a composite fading model, if the slowly-fading component is be trackable, a loading can be based upon this parameter and a fade distribution on the fast-fading component. Two important cases of QoS are considered: the typical average error probability (bit or symbol) or outage probability. In both QoS cases, convexity and scalability properties were shown to hold, and the result is an optimal, fast loading algorithm for real-time point-to-point wireless systems.

Further extensions of this framework are ongoing and include the use of linear or Kalman prediction to estimate future, instantaneous subchannel gains. By predicting enough ahead, this may overcome some or all of the latency of channel-state feedback and loading time, and thereby result in increased system performance. Such linear-prediction OFDM-allocation ideas have been considered in [11]–[12].

Other extensions to follow include the optimal loading in MIMO systems with space-time coding, and the resource allocation in multiuser systems [12]–[14], which results in $K$-dimensional Lagrange optimization (i.e., there is a separate Lagrange multiplier for each of $K$ users). Subchannels remain separable within the same user, but a single subchannel allocation is constrained by multiple users, and we are looking at the scenarios of subchannel sharing as well as unique assignment. These problems are more difficult and again require an extremely efficient algorithm along with the optimal single-user wireless framework contained in this paper.

REFERENCES