Average Power Reduction for MSM Optical Signals via Sparsity and Uncertainty Principle

Jovana Ilic, *Student Member, IEEE,* and Thomas Strohmer

Abstract

Multiple subcarrier modulation is an appealing scheme for high-data rate optical communication. However a major drawback is its low average power efficiency. While subcarrier reservation is a common approach to combat this problem, little is known about the performance of algorithms that utilize subcarrier reservation. By combining properties of sparse signals with an abstract form of the Uncertainty Principle related to multiple subcarrier signals, we design an effective iterative method for constructing average-power-efficient multicarrier signals. Unlike most existing subcarrier reservation methods, our method provides a guaranteed bound for the achievable average power reduction as well as guaranteed rates of convergence. Numerical simulations demonstrate the performance of the proposed method.

I. Introduction

In optical and wireless communications systems, Multiple Subcarrier Modulation (MSM) is used to combat intersymbol interference in frequency selective channels. In MSM multiple subcarriers are modulated at a lower rate with a conventional modulation scheme, while the overall bit rate remains the same as if a single carrier was used [1].

In optical fiber communications multiple digital and/or analog information sources are modulated onto different electrical subcarriers, which are then modulated onto a single optical carrier [2]. The MSM electrical signal can be modulated onto the optical carrier using intensity, frequency of phase modulation, however most current MSM systems use intensity modulation [2].

One of the main drawbacks of Intensity Modulated (IM) MSM is average optical power inefficiency. Since optical intensity needs to be non-negative, a large dc bias must be added to the electrical MSM signal so that it becomes non-negative. This increases the average power of the system which is not only power inefficient, but can also trigger non-linear effects and signal distortion.

In [2], two techniques have been proposed which reduce the average optical power in IM MSM systems: use of reserved subcarriers with block coding and a variable dc bias that depends on the current symbol. In [3] a lower bound for the maximum average power of the signal with reserved subcarriers placed contiguously, at the end of the transmission band, is obtained.

Both authors were supported by NSF project DMS 0811169, and T.S. was supported by AFOSR project 5-36230.5710

Subcarrier reservation is a method where a certain set of subcarriers, *reserved subcarriers*, is not used for data transmission. Reserved subcarriers are used to increase the average power efficiency of the system [3]. When reserved subcarriers are used, two questions arise: how should reserved subcarriers be positioned and what is the optimal amplitude that should be transmitted on those subcarriers, that would result in smaller dc bias that is necessary.

In this paper we give answer to both of these questions. We propose the use of Kashin's representation [4] of a signal with reserved subcarriers in order to obtain a more flat signal in the electrical domain. Lyubarskii and Vershynin have shown that it is possible to obtain a Kashin's representation of a vector, if the transformation matrix used satisfies the Uncertainty Principle for matrices, given in [4]. They present an efficient algorithm for computing Kashin's representation in connection with vector quantization. When adapting their approach to our problem, we need to make some key modifications, which stem from the fact that the setting in [4] is for discrete, finite-dimensional vectors, while we have do deal with continuous-time signals. Once Kashin's representation is obtained, the information is more evenly distributed over all subcarriers - large peaks are eliminated, hence required bias is reduced, which results in the average power reduction (APR). The data subcarriers are positioned only on odd subcarriers, since it has been shown in [5] that in this case the clipping noise falls on the even subcarriers. We further provide a robust and numerically cheap algorithm to obtain a Kashin's representations for IM MSM signals. We note that in terms of APR reduction, the method of reserved subcarriers is not necessarily the best one, however, in case reserved subcarriers are used, Kashin's representation provides a good performance and a simple algorithm for finding the amplitudes of reserved subcarrriers.

The rest of the paper is organized as follows. In section II we define the problem and give a brief overview of the system model and on the reserved subcarrier approach. In Section III, which contains our main results, we introduce a key concept of our approach, Kashin's representation, as well as our main theorem. Section IV gives a detailed description of our APR algorithm, while Section V contains numerical simulations. Concluding remarks are given in Section VI.

II. Problem Statement

We consider the following baseband model for real-valued MSM optical signals with BPSK-modulation[1]

$$\mathbf{s}(t) = \left(\sum_{l=1}^{N} c_l \cos \omega_l t \right) g(t) \qquad 0 \leq t \leq T,$$

where $c_l \in \{-1, 1\}$, $g(t)$ is a rectangular transmit pulse shape, T is the symbol duration and ω_n are the subacarrier frequencies given by [3]

$$\omega_l = \frac{\pi l}{T}.$$

Without loss of generality we may assume that $T = 1$.

[1]We note that the results derived in this paper can be easily extended to MSM signals with different trigonometric representations.

As mentioned in the introduction, we need to ensure that $\mathbf{s}$ is not negative so that intenstity-based detection at the receiver can work properly. In order to make $\mathbf{s}(t)$ non-negative, a (potentially large) dc bias, $\mathbf{s}_{dc}$, must be added to the signal

$$\mathbf{s}_1(t) = \mathbf{s}(t) + \mathbf{s}_{dc}.$$

It is obvious that the minimum possible value for $\mathbf{s}_{dc}$ that will result in $\mathbf{s}_1(t) \geq 0$ is given by

$$\mathbf{s}_{dc} = |\min_{t \in [0,1)} \mathbf{s}(t)|.$$

Since the average power of $\mathbf{s}_1(t)$ is proportional to $\mathbf{s}_{dc}$, the problem of reducing the average power comes down to reducing the negative peaks of $\mathbf{s}(t)$.

Remark 1: We note that in multiple subcarrier modulation a fraction of the time domain symbol from the end is inserted at the start of the same time domain symbol (cyclic prefix) in order to prevent inter-sybmol intereference. Since cyclic prefix insertion is done after the peak reduction, it in no way affects the amount of dc bias that needs to be added to the signal to ensure its non-negativity.

A. Average power reduction via reserved subcarriers

The method of reserved subcarriers was first introduced in [6], and has become a standard approach to combat the peak-to-average power problem in OFDM [7]. It is thus not surprising that it has also been proposed in connection with the average-power problem in optical communications. We note here that a convex optimization approach using reserved subcarriers appeared in [8]. While the performance of this approach is attractive, its computational complexity is still too large in most applications.

Let $\mathcal{I}_{\text{sub}} = \{0, 1, \dots, N-1\}$ be the index set of available subcarriers. In the reserved subcarrier approach, we split $\mathcal{I}_{\text{sub}}$ into two sets $\mathcal{I}_{\text{data}}$ and $\mathcal{I}_{\text{res}}$, $\mathcal{I}_{\text{sub}} = \mathcal{I}_{\text{data}} \cup \mathcal{I}_{\text{res}}$, where we denote $|\mathcal{I}_{\text{data}}| := n$. The subcarriers associated with $\mathcal{I}_{\text{data}}$ are used for sending data symbols $\{x_k\}_{k=1}^{n}$, while the coefficients $\{r_k\}_{k=1}^{N-n}$ sent on the *reserved subcarriers* associated with $\mathcal{I}_{\text{res}}$ are chosen such that $|\min s(t)|$ is minimized, where

$$\mathbf{s}(t) = \sum_{k \in \mathcal{I}_{\text{sub}}} c_k \cos \pi k t, \text{ with } c_k = \begin{cases} x_k & \text{if } k \in \mathcal{I}_{\text{data}}, \\ r_k & \text{if } k \in \mathcal{I}_{\text{res}}. \end{cases} \tag{1}$$

Since the coefficients r_k carry no information and are thus disregarded at the receiver, this scheme reduces the data rate by a factor n/N. This is the main drawback of any reserved subcarrier APR algorithm. When system resources, such as bandwidth in our case, are used to reduce the APR, the system capacity is increased. On the other hand, the sacrifice of bandwidth results in capacity reduction. There is a trade-off, so the real measure of performance of any APR reduction method based on subcarrier reservation is achievable capacity. Here, we focus solely on the APR reduction - the analysis of achievable capacity when reserved subcarriers are used is beyond the scope of this paper.

While the set of reserved subcarriers is chosen in advance and cannot change from signal to signal (otherwise the resulting overhead of informing the receiver about the new set of reserved subcarriers would render such a scheme useless), the coefficients r_k are optimized depending on the data vector to be transmitted.

This raises the following crucial questions:

1) How shall the index set of reserved subcarriers be chosen so that it provides optimal performance for all signals?

2) Given a properly chosen set $\mathcal{I}_{\text{res}}$, how can we efficiently compute the coefficients r_k?

3) Can we give a bound on the maximum dc bias that has to be added for a given set $\mathcal{I}_{\text{res}}$?

B. Oversampling and average power reduction

It should be noted that all previously stated equations are dealing with continuous-time signals. However, in practical systems, the signal is digitally generated and processed, before it is converted to a continuous-time signal. Essentially all existing peak reduction methods operate on discrete signals, the samples of an MSM signal. It is well known that the peak value of the samples can *only* be related to the peak value of the continuous signal, if the signal is oversampled, sampling at Nyquist rate does not suffice. Therefore, any peak reduction method must operate on sufficiently oversampled discrete signals to produce meaningful results. In [9] an upper bound is derived on the peak value of real and complex bandlimited signals, given the peak value of the samples and the oversampling rate. To achieve this oversampling rate, it is customary to zero-pad the signal in the frequency domain with $(L-1)N$ zeros, where N is the number of subcarriers that the MSM signal occupies. An oversampling factor of $L=4$ has been shown to be sufficient for most practical purposes [10].

In the design of our algorithm, we will focus on minimizing the dc-bias for sufficiently oversampled discrete-time signals, as defined in (2). By invoking a theorem by Boche and Wunder [9], we will then be able to infer an accurate statement about the required dc-bias for the associated continuous-time signals.

Therefore we consider the reserved subcarrier approach for sufficiently oversampled discrete-time signals. We introduce a discrete-time signal, $s(l)$, associated with (1), as follows. For a given oversampling factor $L \geq 1$ we define $\mathcal{I}_{\text{zero}} = \{N, N+1, \ldots, NL-1\}$ and set $\mathcal{I} := \mathcal{I}_{\text{sub}} \cup \mathcal{I}_{\text{zero}}$ (the first N indices of $\mathcal{I}$ coincide with $\mathcal{I}_{\text{sub}}$). Denoting $M := LN$, we define $s(l) = s(t/NL) = s(t/M)$ for $l = 0, \ldots, M-1$ as:

$$s(l) = \sum_{k \in \mathcal{I}} c_k \cos \pi k l / M, \text{ where } c_k = \begin{cases} x_k & \text{if } k \in \mathcal{I}_{\text{data}}, \\ r_k & \text{if } k \in \mathcal{I}_{\text{res}}, \\ 0 & k \in \mathcal{I}_{\text{zero}}. \end{cases} \tag{2}$$

For the purpose of designing an APR algorithm, using reserved subcarriers, the introduction of zero-padding in (2) can also be interpreted as having two types of data subcarriers, $\mathcal{I}_{\text{data}}$, where we send the actual desired information, and the zero-padded "subcarriers" $\mathcal{I}_{\text{zero}}$, where we send zeros. We emphasize that this viewpoint only makes sense while designing the APR algorithm, in practice one does, of course, not send zeros on the zero-padded subcarriers.

C. Choice of reserved subcarriers

In [5] a clipping method has been proposed in order to achieve a more power efficient optical OFDM. The authors propose using no bias - all negative values are clipped and forced to zero. The method of the authors is based on the very nice observation that if the data in the unclipped MSM signal in frequency are positioned only on the odd subcarriers, after clipping the *whole negative portion* of the signal then all clipping noise falls on the even subcarriers. Unfortunately this observation holds only when no oversampling is used (i.e., $L = 1$). Thus, in a real case scenario this method would fail since the clipped signal will have a much larger bandwidth due to the clipping (it is easy to see that if the oversampled signal is clipped, then the subcarriers that correspond to zero-pading would not be zero after the clipping). Why not simply filter the clipped signal? This however introduces new, possibly quite large, negative peaks into the signal. Surely, the reader might think, we can simply iterate this process of clipping and filtering and achieve our goal. Unfortunately it is not that easy. Indeed, if one did clip the whole negative part of the signal, one can show that iterating between clipping and filtering would not result in a convergent algorithm. Using zero-padding instead of filtering would produce equally useless results in this case.

Still, this algebraic property of frequency distribution of clipping noise that was observed in [5] is very appealing and can be effectively used in combination with our algorithm. Therefore, in one version of our algorithm the reserved subcarriers will be positioned only on even subcarriers. The set of reserved positions in that case is given by

$$\mathcal{I}_{\text{res}} = \{0, p, 2p, \ldots, (N/p-1)p\},$$

where we assume that $2 \leq p$ divides N, and that p is even.

III. Uncertainty Principle and Kashin's Representations

In this section we introduce Kashin's representation, a key concept in the design of our algorithm. Kashin's representation has been proposed by Lyubarskii and Vershynin for the construction of "spread" frame representations by linking it to an abstract form of the Uncertainty Principle. We refer the reader to [4] for more details on the material presented in the next few paragraphs. For background on frame theory we recommend [11]. In this paper we only need a very specific type of frame, so-called *Parseval frames*, which are a natural way to extend many of the powerful properties of orthonormal bases to a redundant (i.e., linear dependent) collection of vectors. To be precise, a sequence $\{u_i\}_{i=1}^M \subset \mathbb{C}^m$ is called a Parseval frame if

$$\|x\|_2^2 = \sum_{i=1}^{M} |\langle x, u_i\rangle|^2,$$

for all $x \in \mathbb{C}^m$. A frame $\{u_i\}_{i=1}^M \subset \mathbb{C}^m$ can be identified with its frame matrix $\mathbf{U}$ of size $m \times M$ whose columns are u_i. Given a Parseval frame, every vector $x \in \mathbb{C}^m$ can be represented as

$$x = \sum_{i=1}^{M} b_i u_i = \mathbf{U}b, \tag{3}$$

with frame coefficients $\{b_i\}_{i=1}^M$, where $b_i = \langle x, u_i\rangle = (\mathbf{U}^* x)_i$.

When $M > m$, (Parseval) frames are linearly dependent systems, therefore the representation (3) is not unique. In fact, there exist inifinitely many different choices for the frame coefficients b_i (as opposed to orthonormal bases where the coefficients are unique). The choice $b_i = \langle x, u_i \rangle$ is special in the sense that the resulting sequence of coefficients $\{b_i\}_{i=1}^M$ has minimal ℓ_2-norm among all possible choice. However, sometimes it is more desirable to have representation coefficients that satisfy other properties, as is the case in this paper. For instance we may be interested in "spreading" the information contained in x "as uniform as possible" across the frame coefficients. To achieve this goal, Lyubarskii and Vershynin propose to use *Kashin's representation*, which we define now.

Definition 3.1 (Kashin's representation): Consider a sequence $\{u_i\}_{i=1}^M \subset \mathbb{C}^m$, where $M > m$. We say that the expansion

$$x = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} a_i u_i \tag{4}$$

is a Kashin's representation with level K of a vector $x \in \mathbb{C}^m$ if

$$|a_i| \leq K \|x\|_2.$$

We call the coefficients a_i Kashin's coefficients.

It is easy to see that if all frame coefficients b_i have the same magnitude, their magnitude is equal to $1/\sqrt{M}$. Thus, owing to the normalization by $1/\sqrt{M}$ in (4), the smallest value for K is 1 (which however is often not achievable). We emphasize that K is independent of x.

The existence of Kashin's representation is related to an abstract form of the Uncertainty Principle, which we recall in a moment. Whenever a frame $\{u_i\}_{i=1}^M$ satisfies this Uncertainty Principle, one can effectively transform every frame representations into a Kashin's representations.

Recall that the classical Uncertainty Principle states that a function cannot be localized in both time and frequency. A generalized version of the Uncertainty Principle has been introduced for the discrete setting by Lyubarskii and Vershynin, however we consider a special case of their definition. Before we proceed we need to define a way of measuring "localization" or sparsity of a signal in the discrete, finite setting, since the usual notion of localization of a signal via its second-order moments does not yield anything useful in finite dimensions. To that end, for $a \in \mathbb{C}^m$ let its quasi-norm[2] $\|a\|_0$ denote the number of non-zero entries of a. In other words, if $\|a\|_0$ is small, then a is a sparse signal. As in [4], but with an additional restriction, we introduce the following:

Definition 3.2 (Uncertainty Principle for matrices): An $m \times M$ matrix $\mathbf{U}$ satisfies the Uncertainty Principles with parameters $\eta, \delta \in (0, 1)$ if

$$\|\mathbf{U}a\|_2 \leq \eta \|a\|_2, \tag{5}$$

for all $a \in \mathbb{C}^m$ such that $\|a\|_0 \leq \delta M$, and $a_i \leq 0$, where $i = 0, \ldots m - 1$.

To see why this is indeed a form of Uncertainty Principle, consider the case where $\mathbf{U}$ consists of m random rows of the $M \times M$ Fourier matrix. Then inequality (5) says that the energy of the Fourier transform of x cannot be entirely concentrated on the frequencies associated with the row indices of $\mathbf{U}$ (otherwise we would have $\|\mathbf{U}x\|_2 = \|x\|_2$).

[2]The "zero-norm" $\|\cdot\|_0$ is not really a norm, since the property $\|\alpha a\|_0 = |\alpha| \|a\|_0$ is not satisfied.

We now introduce a negative-sided truncation (we are only interested in eliminating peaks on the negative side of our signal):

$$\tau_{-\Delta} = \max(z, -\Delta)$$

where $\Delta > 0$. Let

$$x = \sum_{i=0}^{M-1} b_i u_i,$$

be the frame representation of $x \in \mathbb{C}^m$ and u_i are the columns of transformation matrix $\mathbf{U}$ that satisfies the Uncertainty Principle with parameters η and δ. Then we can define a truncation operator, $\mathcal{T}$ as:

$$\mathcal{T}x = \sum_{i=0}^{M-1} \hat{b}_i u_i, \qquad \hat{b}_i = \tau_{-\Delta}(b_i), \qquad \Delta = \frac{\|x\|_2}{\sqrt{\delta M}}$$

One can show that the residual of this one sided truncation is small, due to the fact that $\mathbf{U}$ satisfies the Uncertainty Principle. Specifically we have [4]

Lemma 3.3:

$$\|x - \mathcal{T}x\|_2 \leq \eta \|x\|_2 \tag{6}$$

Proof: While this lemma is essentially identical to Lemma 3.6 in [4], we include the short proof nevertheless since it highlights in a nice way the crucial role of the correct choice of the threshold.

We consider a subset T of $\{0, ..., M-1\}$ defined as $T = \{i : b_i \neq \hat{b}_i\} = \{i : \ b_i < -\Delta\}$.

Then we have that the following holds:

$$\|x\|_2^2 = \sum_{i=0}^{M-1} |b_i|^2 > |T|\Delta^2.$$

Thus,

$$|T| \leq \frac{\|x\|_2^2}{\Delta^2} = \delta M.$$

We can then estimate the residual that results from the truncation:

$$x - \mathcal{T}x = x = \sum_{i \in T} (b_i - \hat{b_i}) u_i = \mathbf{U}(b - \hat{b}).$$

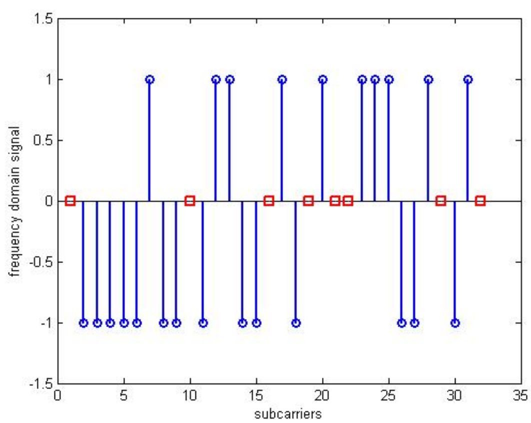
Then,

$$\begin{aligned}
\|x - \mathcal{T}x\|_2 &\leq \eta (\sum_{i \in T} |b_i - \hat{b_i}|^2)^{1/2} \\
&\leq \eta (\sum_{i \in T} |b_i|^2)^{1/2} \\
&\leq \eta (\sum_{i=0}^{M-1} |b_i|^2)^{1/2} = \eta \|x\|_2
\end{aligned}$$

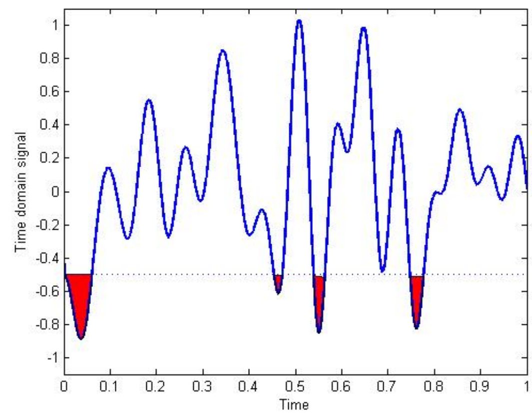
where we have used the Uncertainty Principle for matrices. ■

In figure 1 we have illustrated how the Uncertainty Principle allows for the average power reduction in multicarrier signals. In figure 1(a) we have a frequency domain representation of a multicarrier signal with reserved subcarriers.

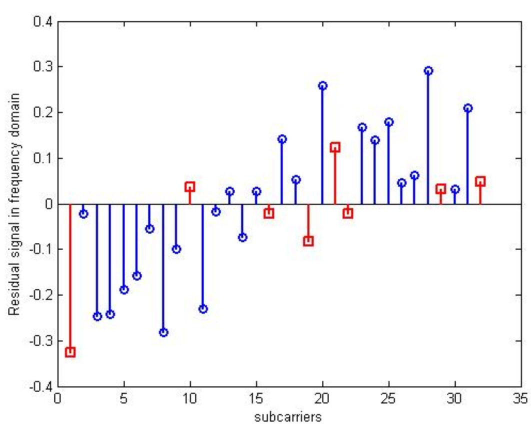
The data subcarriers of the multicarrier signal are represented with circles, while the reserved subcarriers, set to zero, are represented with squares. In figure 1(b), we can see the corresponding time domain signal of the MSM signal from figure 1(a), which in lemma 3.3 is denoted by coefficients b_i. If we now truncate the negative peaks of the time domain signal (the shaded area in figure 1(b)), we obtain the residual signal, $b - \hat{b}$ which is shown in figure 1(d). We can see that this signal is sparse in time. If we now look at the corresponding frequency domain of the residual signal, $x - \mathcal{T}x = \mathbf{U}(b - \hat{b})$ from lemma 3.3, we can see that the residual signal is not sparse in the frequency domain, due to the Uncertainty Principle. The reserved subcarriers are no longer equal to zero. What this means in terms of the average power reduction is that we can clip the negative peaks and use the reserved subcarriers to represent the clipping noise, in order to preserve the orignal data. This distortionless clipping is the basic idea behind our algorithm which is presented in section IV.



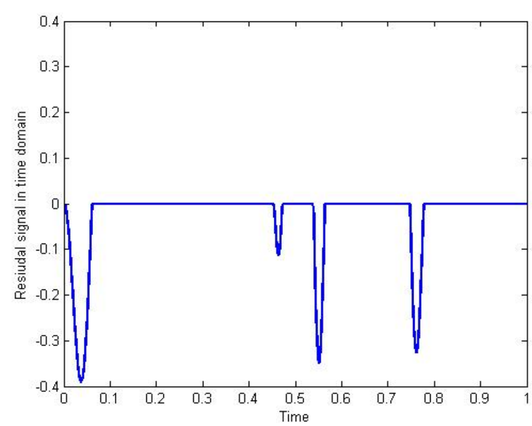
(a) Frequency domain: Data subcarriers and reserved subcarriers



(b) Time domain: Corresponding signal in time with clipping of the negative peaks



(c) Frequency domain: Residual signal after clipping



(d) Time domain: Sparse residual signal

Fig. 1. Illustration of the Uncertainty Principle on the MSM signal with reserved subcarriers

It is shown in section 4 of [4] that certain random orthogonal matrices satisfy the Uncertainty Principle with high probability. In particular, it is proven in Theorem 4.3 that the random partial Fourier $m \times M$ matrix (constructed by randomly selecting m rows of the $M \times M$ Fourier matrix for properly chosen m) satisfies the Uncertainty Principle for matrices with high probability.

Unfortunately we cannot simply resort to this appealing result. The reason, as mentioned earlier, is that we have to oversample our signal in order to get an accurate estimate of its peak properties. This oversampling means that even if we did select the reserved subcarriers randomly, the resulting matrix we are dealing with contains a block of $(L-1)N$ (deterministically selected) rows that correspond to the zero-padded subcarriers $\mathcal{I}_{\text{zero}}$, as explained in Subsection II-B. Thus in our case the matrix $\mathbf{U}$ will never be entirely random.

In order to use Kashin's representation, a crucial step is therefore to demonstrate that the transform matrix described in the lemma below does indeed satisfy the Uncertainty Principle for matrices.

Let $\mathbf{D}$ be an $M \times M$ discrete cosine transform matrix given by

$$\mathbf{D} = \begin{bmatrix} T_0(0) & T_0(1) & \dots & T_0(M-1) \\ T_1(0) & T_1(1) & \cdots & T_1(M-1) \\ \vdots & \vdots & \ddots & \vdots \\ T_{M-1}(0) & T_{M-1}(1) & \cdots & T_{M-1}(M-1) \end{bmatrix}, \tag{7}$$

where

$$T_k(i) = \begin{cases} \sqrt{\frac{1}{M}} & k = 0, \\ \sqrt{\frac{2}{M}} \cos \frac{(2i+1)k\pi}{2M} & \text{for } k = 1, 2, \dots, M-1. \end{cases}$$

Lemma 3.4: Let $\mathbf{D}$ be an $M \times M$ discrete cosine transform matrix where $M = LN$. We divide the index set $\mathcal{I} = \{0, \dots, M-1\}$ into three disjoint subsets $\mathcal{I}_{\text{data}}$, $\mathcal{I}_{\text{res}}$, $\mathcal{I}_{\text{zero}}$, where $0 \in \mathcal{I}_{\text{res}}$, $|\mathcal{I}_{\text{res}}| = q$, $\mathcal{I}_{\text{zero}} = \{N, N+1, \dots, M-1\}$ and $|\mathcal{I} \setminus \mathcal{I}_{\text{res}}| = m$, so that $M = m + q$. Let $\mathbf{U}$ be the $m \times M$ submatrix of $\mathbf{D}$ that consists of the rows of $\mathbf{D}$ indexed by $\mathcal{I} \setminus \mathcal{I}_{\text{res}}$. Then there exist $(\eta, \delta) \in (0, 1)$ such that $\mathbf{U}$ satisfies the Uncerainty Principle with parameters (δ, η)

$$\|\mathbf{U}a\|_2 \leq \eta \|a\|_2,$$

for all $a \in \mathbb{C}^m$ such that $\|a\|_0 \leq \delta M$, $a_i \leq 0$, for $i = 0, \dots m-1$, and $\eta = 1 - \frac{1}{\sqrt{M}} \|a_T\|_1$, where a_T is a vector of all non zero elements in $a \in \mathbb{C}^m$.

Proof: Let $\mathbf{V}$ be the $q \times M$ submatrix of $\mathbf{D}$ that consists of the of the rows of $\mathbf{D}$ indexed by $\mathcal{I}_{\text{res}}$.

Since $\mathbf{D}$ is a unitary matrix, we have

$$\|\mathbf{D}a\|_2 = \|a\|_2 \tag{8}$$

Then

$$\|\mathbf{D}a\|_2^2 = \|\mathbf{U}a\|_2^2 + \|\mathbf{V}a\|_2^2,$$

and

$$\|\mathbf{U}a\|_2^2 = \|\mathbf{D}a\|_2^2 - \|\mathbf{V}a\|_2^2. \tag{9}$$

Since we want to show that

$$\|\mathbf{U}a\|_2 < \eta\|a\|_2$$

for all nonpositive a such that $\|a\|_0 \leq \delta M$ from (8) and (9) we can see that it is sufficient to show that there exists a number $\tilde{\eta} \in (0,1)$ such that

$$\|\mathbf{V}a\|_2 \geq \tilde{\eta} > 0, \tag{10}$$

for all a with $\|a\|_0 \leq \delta M$, $a_i \leq 0, i = 0, \ldots, M-1$.

Let T be any index set from $\{0, \ldots, M-1\}$ of non-zero elements of the vector a such that $|T| < \delta M$. Then, since $\|\mathbf{V}a\|_2 = \|\mathbf{V}_T a_T\|_2$, we have that (10) is equivalent to

$$\|\mathbf{V}_T a_T\|_2 \geq \tilde{\eta} > 0 \tag{11}$$

for all a_T such that $(a_T)_i \neq 0$, for $i = 0, \ldots, |T|-1$; where a_T is a vector of non-zero elements of vector a, and $\mathbf{V}_T$ is constructed by selecting columns of $\mathbf{V}$ that correspond to positions of non-zero elements in vector a. We emphasize that we have no control over which columns of $\mathbf{V}$ are selected, we only know that the number of selected columns is bounded by $|T|$.

Considering (7), we can look at the 0^{th} element of the product of $\mathbf{V}_T a_T$:

$$(\mathbf{V}_T a_T)_0 = \frac{1}{\sqrt{M}} \sum_{k=0}^{q-1} a_T(k) = -\frac{1}{\sqrt{M}} \|a_T\|_1,$$

since all the entries of a are nonpositive. Hence for elements of a_T we have that $a_{Ti} < 0$. From here, it is obvious that

$$(\mathbf{V}_T a_T)_0 < 0. \tag{12}$$

We can write $\|\mathbf{V}_T a_T\|_2^2$ as:

$$\|\mathbf{V}_T a_T\|_2^2 = |(\mathbf{V}_T a_T)_0|^2 + \sum_{k=1}^{q-1} |(\mathbf{V}_T a_T)_k|^2.$$

Therefore, due to (12) there holds

$$\|\mathbf{V}_T a_T\|_2 > 0.$$

and

$$\tilde{\eta} := \frac{1}{\sqrt{M}} \|a_T\|_1 > 0.$$

Setting

$$\eta = 1 - \tilde{\eta}$$

we note that $\eta < 1$, which concludes our proof. ∎

Remark 2: Due to the one sided truncation, the non-zero elements of the truncated signal have indeed all the same (negative) sign.

Remark 3: It is clear that the proof is independent of the choice of $\mathcal{I}_{\text{res}}$, as long as $0 \in \mathcal{I}_{\text{res}}$. However, that does not mean that all choices of $\mathcal{I}_{\text{res}}$ will perform equally well. The choice of $\mathcal{I}_{\text{res}}$ that minimizes η and maximizes δ will give the optimal performance (note that η and δ are not independent). In this paper we consider two choices for $\mathcal{I}_{\text{res}}$ - one presented in subsection II-C, and one where elements of $\mathcal{I}_{\text{res}}$ are randomly chosen (although we do ensure that 0 is always included). We furthermore note that the estimate for η in our proof is far from sharp.

Theorem 3.5: Let $\mathcal{I}_{\text{sub}} = \{0, 1, \ldots, N-1\}$ be the set of available subcarriers. Let $\mathcal{I}_{\text{res}}$ be the set of reserved subcarriers such that $0 \in \mathcal{I}_{\text{res}}$ and $|\mathcal{I}_{\text{res}}| = q$. Let $\{x_k\}_{k=1}^n$ be a vector with $c_k \in \{-1, +1\}, k = 1, \ldots, n$, transmitted on the subcarriers $\mathcal{I}_{\text{data}} = \mathcal{I}_{\text{sub}} \setminus \mathcal{I}_{\text{res}}$. Assume we construct the Kashin representation of $\{c_k\}_{k=1}^n$ with respect to $\mathcal{I}_{\text{data}}$, $\mathcal{I}_{\text{res}}$ with oversampling factor L, where $\{r_k\}_{k=1}^q$ are the coefficients transmitted on the reserved subcarriers. Then the continuous-time signal

$$\mathbf{s}(t) = \sum_{k \in \mathcal{I}_{\text{sub}}} c_k \cos \pi k t, \text{ with } c_k = \begin{cases} x_k & \text{if } k \in \mathcal{I}_{\text{data}}, \\ r_k & \text{if } k \in \mathcal{I}_{\text{res}}. \end{cases} \tag{13}$$

satisfies

$$\min_t s(t) \geq -(K + N(C_L - 1)), \tag{14}$$

where $C_L = (\cos \frac{\pi}{2L})^{-1}$ and K is the Kashin constant given by $K = (1-\eta)^{-1}\delta^{-1/2}$, where η and δ are as defined in Lemma 3.4.

Proof:

In section V of [9] it has been shown that for a complex trigonometric polynomial (multicarrier signal) given as:

$$s(t) = \sum_{k=0}^{N-1} c_k e^{jk2\pi t}, \quad 0 \leq t \leq 1 \tag{15}$$

the following is true:

$$\left| \max_{t \in [0,1]} |s(t)| - \max_{0 \leq l < LN} \left| s\left(\frac{l}{LN}\right) \right| \right| \leq N(C_L - 1) \tag{16}$$

where N is the number of subcarriers, c_k, $k = 0, ..., N-1$ are the information bearing coefficients and L is the oversampling factor. It is not difficult to see that this result also applies to our setting.

Equation (16) gives the bound on the error that is made when we use an oversampling grid to estimate the peaks of $s(t)$. We can see that with choosing the appropriate L, we can achieve any desired accuracy [9]. If we are only interested in estimating the peaks on the negative side, (16) becomes:

$$\min_{t \in [0,1]} s(t) - \min_{0 \leq l < LN} s\left(\frac{l}{LN}\right) \geq -N(C_L - 1)$$

By Lemma 3.4 the matrix $\mathbf{U}$ satisfies the Uncertainty Principle with parameters η and δ. Then each vector $x \in \mathbb{C}^m$ admits Kashin's representation of level $K = (1-\eta)^{-1}\delta^{-1/2}$.

Let $s = [s_0, \ldots, s_l, \ldots, s_{LN-1}]$, where $s_l = s\left(\frac{l}{LN}\right)$, be a Kashin's representation of some vector $x \in \mathbb{C}^m$. Then s_l are Kashin's coefficients, so we know that:

$$|s_l| \leq K\|x\|_2, \quad l = 0, \ldots, LN-1$$

We can assume that $\|x\|_2 = 1$ without loss of generality. For the negative peaks of s we have that:

$$\min_{0 \le l < LN} s_l \ge -K \tag{17}$$

Equation (17) gives a lower bound on the negative peaks of the discrete, L times oversampled signal. It is now obvious, that using (16) we can obtain the lower bound on the negative peaks for the continuous-time signal as:

$$\min_{t \in [0,1]} s(t) \ge -(K + N(C_L - 1)).$$

■

We note that for $K >> N(C_L - 1)$, which is always true if L is large enough, we have that

$$\min_{t \in [0,T_s]} s(t) \gtrsim -K$$

So, when using a Kashin's representation we can give a guarantee on what the maximum value of the dc bias that needs to be added to the signal will be: it is equal to Kashin's constant K.

Remark : Along the same lines as in Theorem 3.5 one can derive a bound on the dc bias in case the reserved subcarriers are placed at the end of the transmission band. This setup coincides then with the setup in [3].

So far, we have established that in our setup a Kashin's representation of any $x \in \mathbb{C}^m$ exists, and we have shown that we can relate a Kashin's representation to the continuous time signal that needs to be transmitted. We now introduce the algorithm for obtaining a Kashin's representation of any vector $x \in \mathbb{C}^m$.

As stated before, choosing optimal positions of reserved subcarriers is a crucial part of an APR algorithm, and it is a problem that has yet to be solved. Although we do not give an optimal solution for this problem, we provide an insight on how such optimal solution might be obtained: the optimal set of reserved subcarriers is the one that minimizes η.

IV. Average power reduction algorithm via Kashin's representation

Here we show how our APR algorithm can be successfully implemented.

Let N be a number of subcarriers in one MSM symbol and $\mathcal{I}_{\text{sub}}$, $\mathcal{I}_{\text{res}}$, and $\mathcal{I}_{\text{data}}$ index sets as defined in section II. Let q be the number of reserved subcarriers, and n the number of data subcarriers, such that $n = N - q$. If L is the oversampling factor, then the size of the oversampled signal, $\tilde{x}$ is $M = LN$, and the index set that corresponds to the zero-padded portion of the the signal is $\mathcal{I}_{\text{zero}}$. We now define an index set $\mathcal{I}_{\text{interest}}$ as $\mathcal{I}_{\text{interest}} = \mathcal{I}_{\text{data}} \cup \mathcal{I}_{\text{zero}}$. The portion of vector $\tilde{x}$ that corresponds to positions in $\mathcal{I}_{\text{interest}}$ is denoted as $x = \tilde{x}(\mathcal{I}_{\text{interest}})$.

We define the matrix $\mathbf{U}$ as a partial discrete cosine transformation (DCT) matrix which is constructed by selecting the rows of a full $M \times M$ DCT matrix that correspond to positions stored in $\mathcal{I}_{\text{interest}}$. $\mathbf{U}$ is of size $m \times M$, where $m = M - q$. We can compute now the frame representation of x as:

$$y = U^{-1}x$$

We now introduce the algorithm for obtaining Kashin's representation of x:

Algorithm 1: **Input:** $\tilde{x}$, $\mathcal{I}_{\text{interest}}$, η, δ,(the choice for η and δ is discussed in section V) ε - desired accuracy in reconstructing x, and/or r - maximum number of iterations. It should be noted here that when constructing $\tilde{x}$ we set $\tilde{x}(I_{\text{res}}) = 0$.

Output: vector a of length M - Kashin's representation of x.

Initialitzation: Initialize vector a: $a_i = 0, i = 1 : M$, calculate the truncation level for the first iteration as $\Delta = \frac{\|x\|_2}{\sqrt{\delta M}}$

Repeat r times:

- Calculate y: the inverse DCT of x
- Truncate y at level $\Delta^{(k)}$: obtain $y_{trunc} = t_{-\Delta^{(k)}}(y)$ where k is the current iteration
- $a^{(k)} = a^{(k-1)} + y_{trunc}$
- Calculate the DCT of $a^{(k)}, \tilde{x}^{(k)} = dct(a^{(k)})$
- Calculate the residual in the following way:
  $res^{(k)} = \tilde{x}^{(k)}(\mathcal{I}_{\text{interest}}) - \tilde{x}(\mathcal{I}_{\text{interest}})$
- Set $\tilde{x}^{(k)}(I_{\text{res}}) = 0$
- Calculate new truncation level as $\Delta^{(k+1)} = \eta\Delta^{(k)}$
- If $\frac{\|res^{(k)}\|_2}{\|\tilde{x}(\mathcal{I}_{\text{interest}})\|_2} \leq \varepsilon$ - break

We can see that by treating the zero-paded portion of x as data subcarriers we enforce that amplitudes on those subcarriers remain 0 which ensures that our signal has the same bandwidth as the original signal, which is, as previously discussed, not the case in [5].

We can give a bound on the complexity of our algorithm: in each iteration we do 2 DCTs (which has the same complexity as FFT, $\mathcal{O}(NL \log NL)$. So the algorithm complexity is $r\mathcal{O}(NL \log NL)$, where r, for a given η and ε, can be obtained from $\varepsilon = \eta^r$.

V. Simulation Results

In this section we evaluate the performance of our algorithm. Our measure of performance is complementary cumulative distribution function (CCDF) of the added dc bias, which denotes the probability that the required dc bias for one IM MSM signal exceeds a given threshold. We plot the CCDF in case our algorithm is used and in case when no average power reduction method is used.

We use an MSM signal with 128 subcarriers. We explore the performance of our algorithm with different number of reserved subcarriers: 64, 32, 16 and 8, positioned on every other, every fourth, every eighth and every sixteenth, even subcarrier, respectively. The oversampling factor is $L = 4$, so the oversampled signal in each case is of length 512. The modulation used is BPSK, and in each case we have averaged over 10^5 randomly generated symbols. We have also set $\|x\|_2 = 1$ in order to ensure fairness of comparison of our curves for different number of reserved subcarriers.

It is important to note that there are no clear guidelines on choice of η and δ , although the right choice is critical for the performance of the algorithm. In [4] it is shown that from Kashin's representation,x can be reconstructed

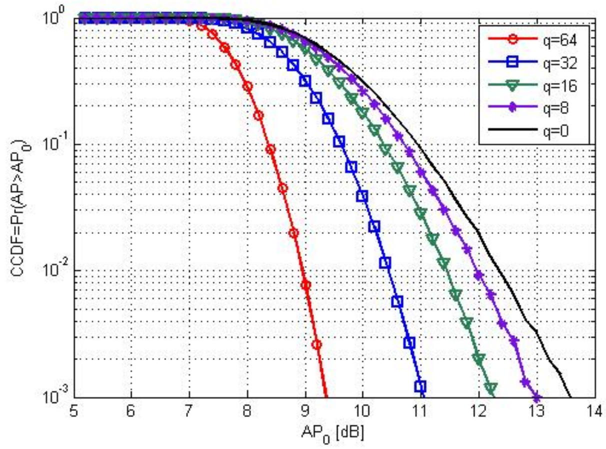


Fig. 2. CCDF of the probability that the minimum dc bias for a 128 IM MSM signal exceeds a given value on the x-axis, in case 64, 32, 16 and 8 reserved subcarriers positioned on every 2^{nd}, 4^{th}, 8^{th} and 16^{th} even subcarrier, respectively, are used, and in case when no reserved subcarriers are used

with accuracy of $\eta^r\|x\|_2$. So for a given η we can calculate the maximum number of iterations in which the algorithm should converge. Obviously, for a smaller η we can get more average power reduction, but if η is too small, $\tilde{x}$ will not be reconstructed with the desired accuracy. So for a set $\delta = 0.8$ and $\epsilon = 10^{-4}$ we experimentally found the minimum possible η that satisfies this ϵ . The results are given in Table I. We note here that for our choice of η and δ, $\mathbf{U}$ does not always satisfy the Uncertainty Principle. However, it does satisfy it with very high probability, while with such a choice we can achieve much better power efficiency.

q	η
64	0.6125
32	0.6865
16	0.7425
8	0.775

TABLE I

USED η VALUES FOR DIFFERENT NUMBERS OF RESERVED SUBCARRIERS

In Figure 2 the results of our simulations are shown. As it can be seen from the plots, when 64 reserved subcarriers are used we get an improvement of more than 4dB. In case when 8 reserved subcarriers are used the improvement is around 0.5dB.

As stated before in all our simulations the reserved positions are chosen as every p^{th} even subcarrier. We have also explored the performance of our algorithm in case the reserved positions are chosen randomly, with exception that 0^th subcarrier is always included in the set of reserved subcarriers. We plot this result in figure 3 where we can see that we still achieve a slightly better performance in case when we position reserved subcarriers on even

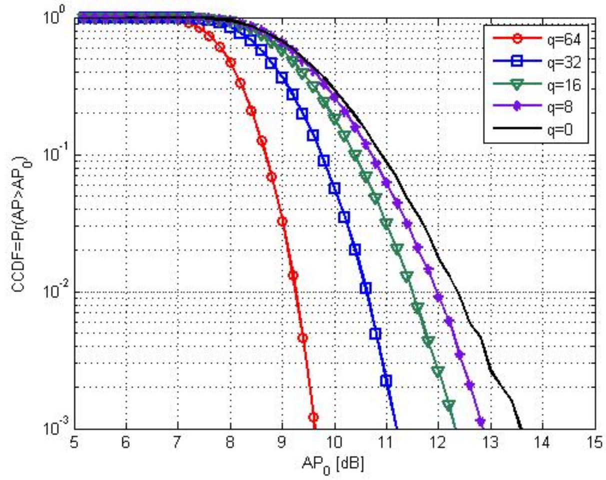


Fig. 3. CCDF of the probability that the minimum dc bias for a 128 IM MSM signal exceeds a given value on the x-axis, in case 64, 32, 16 and 8 randomly positioned reserved subcarriers are used, and in case no reserved subcarriers are used

subcarriers when the number of reserved subcarriers is 64 or 32. When the number of reserved subcarriers is 16 both positioning schemes result in similar performance, while in the case when 8 reserved subacarriers are used, the random positioning gives a slightly better performance.

We have discussed in section II-C that the method for reducing the dc bias proposed in [5] fails because the authors did not consider oversampled signal. In figure 4 we demonstrate what happens when their method is implemented with oversampling. We plot the absolute value of the 4 times oversampled 128-MSM signal in the frequency domain. The x-axis represents the subcarrier index. We note that subcarriers $129-512$ correspond to the zero-padding, and should therefore be equal to 0 after any peak reduction algorithm so that the signal's bandwidth is not increased. As it can be seen from figure 4. this is not the case when all negative values of the MSM time domain signal are clipped, as proposed in [5]. We can also see that, when Kashin's representation is used, zero-padding subcarriers remain zero.

VI. Conclusion

A new algorithm for average power reduction in IM MSM signal has been proposed. We use reserved subcarriers (reserved subcarriers) positioned on even subcarriers, and show that in this case the transformation matrix U satisfies the Uncertainty Principle. Then, a Kashin's representation can be obtained which results in a flatter signal, so large peaks are eliminated. We then propose an algorithm that will yield a Kashin's representation of the signal that needs to be transmitted. As a result, large peaks on the negative side are reduced so the added dc component. For 64 reserved subcarriers we get an improvement of more then 4dB in average power reduction. The proposed algorithm is fast, easy to implement, robust, and it gives a guarantee on the maximum average power in the system.

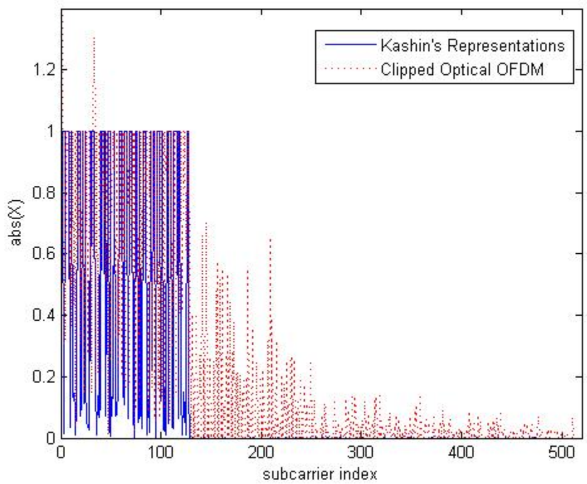


Fig. 4. Absolute value of the frequency domain 128 IM MSM signal with oversampling factor $L = 4$ with 64 reserved tones on even subcarriers obtained via Kashin's representation compared to the clipped optical OFDM. Subcarriers $129 - 512$ correspond to zero-padded portion of the signal

REFERENCES

[1] T. Ohtsuki, "Multiple-Subcarrier Modulation in Optical Wireless Communications," *IEEE Communications Magazine*, vol. 41, no. 3, pp. 74–79, Mar. 2003.

[2] R. You and J. Kahn, "Average Power Reduction Techniques for Multiple Subcarrier Intensity-Modulated Optical Signals," *IEEE Transactions on Communications*, vol. 49, no. 12, pp. 2164–2171, Dec. 2001.

[3] M. Sharif and B. Hassibi, "On the Achievable Average Power Reduction of MSM Optical Signals," *IEEE Communications Letters*, vol. 8, no. 2, pp. 84– 86, Feb. 2004.

[4] Y. Lyubarskii and R. Vershynin, "Uncertainty Principles and Vector Quantization," *ArXiv Mathematics e-prints*, Nov. 2006.

[5] J. Armstrong and A. Lowery, "Power Efficient Optical OFDM," *Electronics Letters*, vol. 42, no. 06, pp. 370–372, Mar. 2006.

[6] J. Tellado and J. Cioffi, "Efficient algorithms for reducing par in multicarrier systems," in *Proc. IEEE International symposium on information theory*, 1998, p. 191.

[7] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Artech House Publishers, 2000.

[8] A. Aggarwal and T. Meng, "A convex interior-point method for optimal OFDM PAPR reduction," in *IEEE International Conference on Communications (ICC)*, Seoul, South Korea, May 2005.

[9] G. Wunder and H. Boche, "Peak Value Estimation of Bandlimited Signals From Their Samples, Noise Enhancement, and a Local Characterization in the Neighborhood of and Extremum," *IEEE Transactions on Signal Processing*, vol. 51, no. 3, pp. 771–780, Mar. 2003.

[10] C. Tellambura, "Computation of the Continuous-Time PAR of and OFDM Signal with BPSK Subcarriers," *IEEE Communications Letters*, vol. 5, no. 5, pp. 185–187, Feb. 2004.

[11] O. Christensen, *An introduction to frames and Riesz bases*. Boston: Birkhäuser, 2003.

Jovana Ilic has received her BS in Electrical Engineering in 2005 from School of Electrical Engineering, University of Belgrade, Serbia. She is currently pursuing the Ph.D. degree in electrical and computer engineering at the University of Califiornia, Davis. Her main research interests are in wireless communications and signal processing.

Thomas Strohmer received his M.S. and Ph.D. in Mathematics in 1991 and 1994 respectively from the University of Vienna, Austria. He was a Research Assistant at the Department of Mathematics, University of Vienna from 1991 to 1997. He spent one year as Erwin-Schroedinger fellow at the Department of Statistics at the Stanford University and then joined the Department of Mathematics at the University of California in Davis in 1998, where he is now Full Professor. His general research interests are in harmonic analysis, numerical analysis, digital signal processing, and wireless communications. He is co-editor of two books and on the editorial board of several journals. He also serves as consultant to the industry in the areas of telecommunications, bioengineering, and medical imaging.