Research Article

PAPR and the Density of Information Bearing Signals in OFDM

Holger Boche and Brendan Farrell

Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, Arcisstraße 21, 80333 München, Germany

Correspondence should be addressed to Holger Boche, boche@tum.de

Received 9 November 2010; Accepted 6 February 2011

Academic Editor: Simon Litsyn

Copyright © 2011 H. Boche and B. Farrell. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We address the peak-to-average power ratio (PAPR) of transmission signals in OFDM and consider the performance of tone reservation for reduction of the PAPR. Tone reservation is unique among methods for reducing PAPR, because it does not affect information bearing coefficients and involves no additional coordination of transmitter and receiver. It is shown that if the OFDM system always satisfies a given peak-to-average power ratio constraint, then the efficiency of the system, defined as the ratio of the number of tones used for information to the entire number of tones used, must converge to zero as the total number of tones increases. More generally, we investigate and provide insight into a tradeoff between optimal signal and information properties for OFDM systems and show that it is necessary to use very small subsets of the available signals to achieve PAPR reduction using tone reservation.

1. Introduction

OFDM is one of today’s most widely used and promising information transmission schemes. One of the main disadvantages of OFDM, however, is the large peak-to-average power ratio (PAPR) of the transmit signals. Reducing the PAPR, which we will call the PAPR reduction problem, has been an area of extensive research over the last ten years, and various techniques have been proposed. These include, among others, clipping and filtering, selected mapping, active constellation extension, and tone reservation. See [1] for an overview.

High PAPR is a problem because most amplifiers perform poorly at higher amplitudes. Reduction techniques are most commonly investigated for their effects on bit-error rates, capacity or power consumption, and in these contexts, a probabilistic approach is appropriate. However, when a signal’s peak is cut off, out-of-band radiation is caused, and this is an interference for devices operating in neighboring spectra. If the peak threshold is violated with a certain probability, then other devices will be subjected to this interference the corresponding percent of the time. In many instances, it may be inadmissible to cause such interference to other devices. We explain that when peak cutoff occurs, only a strict PAPR criterion prevents out-of-band radiation. As such, our approach may be viewed as either addressing a strict PAPR criterion or a strict criterion of no out-of-band radiation. Such an approach is necessarily nonprobabilistic.

The probabilistic model is currently the accepted model among communications engineers working with classical wireless communications systems, such as cellular networks or WLAN. In these areas, one is primarily concerned with the effects that PAPR has for communication within the frequency band that one is working. Developments related to the digital dividend (i.e., the reallocation of frequencies made available by converting radio and television communication from analog to digital) have shown, however, that statistical models are insufficient for various user-oriented communications applications and that there is strong resistance in these communities to redistribute frequency bands. In particular, there is concern that new communications networks will be introduced in neighboring frequency bands or that frequency sharing will be introduced, as both of these present the possibility for interference. Examples of such applications where these issues are most important are those where safety plays an important role, such as in wireless automation or car-to-car communication. A further example is wireless microphones, which are particularly important in the entertainment industry. Statistical models can be
permitted in these areas only in very restricted cases. What distinguishes these applications from classical networks is the type of error that may be tolerated. In particular, a certain probability that communication fails is inadmissible.

The starting point for the PAPR reduction schemes mentioned above is a set of coefficients to be transmitted to the receiver. In order to reduce the signal peak, one may adjust these coefficients in some way or add new coefficients on frequencies that have not been used. If coefficients are manipulated, then the receiver must convert the received coefficients back to the original coefficients; however, if coefficients are added on frequencies that do not carry information, the received information bearing coefficients do not have to be converted. Tone reservation, which was introduced in [2], and is one of the popular techniques to mitigate against high PAPR, takes the latter approach.

In tone reservation, the set of available tones is divided into two subsets. One set is used to carry information, while the other is used to reduce the peak value. We will call these two sets the information set and the compensation set. Given a set of coefficients for the tones in the information set, coefficients are chosen for tones in the compensation set, so that the peak value of the combined signal is reduced. The location of these two sets remains fixed for all codewords and over all uses of the channel.

Of the handful of methods to reduce PAPR, tone reservation is particularly robust and canonical. This is because the only information that the receiver requires is the location of the information set. The receiver may simply ignore whatever arrives on the entries of the compensation set. With other schemes, such as active constellation extension [3] or selected mapping [4, 5], not only does the receiver have to be informed of the modifications made to each possible set of coefficients, but the receiver also has to convert the received coefficients back to their original values. Thus, there is additional overhead involved in setting up and then performing the information transmission. Both of these are avoided in tone reservation.

However, tone reservation exhibits a tradeoff between the best attainable PAPR and the number of tones in the information set. The main result presented here is that if the OFDM system satisfies a strict bound on the peak-to-average ratio, then as the number of total tones used increases, at some point the proportion of tones used to carry information must decrease and eventually converge to zero. Equivalently, we find a scaling law: if the size of the information set and the total set increase proportionally, signals with larger peaks can be constructed that cannot be compensated for by any compensation signal.

The result presented here certainly does not state that tone reservation does not deliver strong improvements in PAPR. An efficient algorithm for computing compensation coefficients is given in [6], and the reductions it delivers in PAPR are significant. Much experimentation has been done, in particular in searching for subsets with good performance, but the structure of good sets is still not understood. There has also been little theoretical work on the performance bounds of tone reservation. The authors are unaware, for example, of any work that addresses the behavior of tone reservation as the number of tones increases. This paper provides insight to this scheme as the number of tones involved becomes large.

This paper provides an initial investigation into the relationship between the relative size of a subset of transmission signals and several signal processing properties of the subset. In Section 2, we prove our main result on tone reservation for OFDM systems with a finite set of tones. In Section 3, we prove the same type of result for systems with an infinite number of tones. The discrete Fourier case is addressed in Section 4, and a short conclusion is given in Section 5. There we also give a short discussion of transmission schemes for fifth-generation cellular networks, as well as nonorthogonal transmission systems and their PAPR behavior.

2. The Finite Set OFDM Case

We first define our signals: an OFDM signal has the form

\[ s(t) = \sum_{k=-N}^{N} a_k e^{ikt}, \]

where the coefficients \( a_k \) either carry information or, in the tone reservation scheme, are used to reduce the peak value of \( s(t) \). An amplifier generally only has a cutoff or distorting effect at high amplitudes, and the signal is left undisturbed where it is magnitude lies below a threshold. In this case, if \( |s(t)| \) exceeds the threshold, say \( M \), in some regions, a new function \( s_1 \) results, such that \( s(t) = s_1(t) \) for all \( t \) where \( |s(t)| \leq M \). Then, \( s - s_1 \) has compact support and cannot be band-limited. Since \( s \) is band-limited, this means \( s_1 \) cannot be band-limited, and thus out-of-band radiation has been caused. This motivates the investigation of strict PAPR constraints. The PAPR for the vector \( a \) is

\[ \text{PAPR}(a) = \sup_{t \in [0,2\pi]} \frac{\sum_{k=-N}^{N} |a_k e^{ikt}|^2}{\|a\|^2_{L^2}}. \]

While this is the standard definition of PAPR, to make notation easier, we look at \( |\sum_{k=-N}^{N} a_k e^{ikt}|^2 \) rather than \( |\sum_{k=-N}^{N} a_k e^{ikt}|^2 \). Also, we could work with just the nonnegative indices rather than with all the frequencies \( \{-N, \ldots, N\} \). Instead of working with \( L^1(T) \) (defined below), we would then work with the Hardy space \( H^1 \). The results, however, remain the same. Before we formally state the problem, we define our spaces.

**Definition 1.** \( L_N^p \) denotes \( \mathbb{C}^N \) viewed as a linear space and \( \|x\|_N^p = (\sum_{k=-N}^{N} |x_k|^p)^{1/p} \). If \( A \) is a subset of \( \{-N, \ldots, N\} \), \( L^p(A) \) denotes vectors in \( L_N^p \) supported on \( A \). \( T \) denotes the torus. \( L^p(T) \) denotes \( p \)-integrable functions defined on \( T \) with norm

\[ \|f\|_{L^p(T)} = \left( \frac{1}{2\pi} \int_T |f(t)|^p \, dt \right)^{1/p}, \]

for \( 1 \leq p < \infty \) and \( \|f\|_{L^1(T)} = \text{ess.sup}_{t \in T} |f(t)| \). \( L^2(A) \) denotes the subspace of \( L^2(T) \) spanned by \( \{e^{ikt}\}_{k \in A} \).
Tone reservation works as follows: we split \(-N, \ldots, N\) into two subsets \(I_N\), an information set, and \(R_N = \{-N, \ldots, N\} \setminus I_N\), a compensation set. We call the ratio of the number of tones in the information set to the total number of tones available the efficiency of the system. Using \(|A|\) to denote the number of elements in the set \(A\), the efficiency is \(|I_N|/(2N + 1)\). Given a set of information coefficients \(a \in \ell^2(I_N)\), one seeks a vector \(b \in \ell^2(\{-N, \ldots, N\} \setminus I_N)\) satisfying \(\|b\|_{\ell^2(I_N)} \leq C_{EX}\|a\|_{\ell^2(I_N)}\), such that

\[
\sup_{t \in [0, 2\pi]} \left| \sum_{k \in I_N} a_k e^{ikt} + \sum_{k \in R_N} b_k e^{ikt} \right| \leq C_{EX}\|a\|_{\ell^2(I_N)},
\]

for some constant \(C_{EX}\). One would like the infimum of (1) over all possible \(b\) supported on \(R_N\). The condition \(\|b\|_{\ell^2(I_N)} \leq C_{EX}\|a\|_{\ell^2(I_N)}\) is, therefore, imposed to make this well-defined. We note, though, that any vector \(b\) that satisfies (1) must have this property. To see this, we observe that

\[
\left(\|a\|^2_2 + \|b\|^2_2\right)^{1/2} = \left\| \sum_{k \in I_N} a_k e^{ikt} + \sum_{k \in R_N} b_k e^{ikt} \right\|_2,
\]

so that if (1) holds, then the condition \(\|b\|_{\ell^2(I_N)} \leq C_{EX}\|a\|_{\ell^2(I_N)}\) is also satisfied. Certainly, for finite \(N\), a constant \(C\) can always be found that satisfies inequality (4). Here, however, we address the relationship between \(N\), the size of \(I_N\), and the best possible constant \(C_{EX}\).

To express this relationship, we introduce the extension operator, which we denote \(E_{I_N}\). This operator maps from \(\ell^2(I_N)\) to \(L^2(\{-N, \ldots, N\})\), given by

\[
E_{I_N}a = \sum_{k \in I_N} a_k e^{ikt} + \sum_{k \in R_N} b_k e^{ikt}.
\]

An initial formulation of solvability is that the PAPR reduction problem is solvable for the subset \(I_N \subset \{-N, \ldots, N\}\) with bound \(C_{EX}\) if there exists an operator \(E_{I_N} : \ell^2(I_N) \rightarrow L^\infty(\{-N, \ldots, N\})\) such that for all \(a \in \ell^2(I_N)\) satisfying \(\|a\|_{\ell^2(I_N)} \leq 1\),

\[
\|E_{I_N}a\|_{L^\infty(T)} \leq C_{EX}.
\]

Note that we are interested in an operator that satisfies (7); uniqueness is not part of the discussion. Such an operator will, in general, be nonlinear, since cases where \(E_{I_N}(a+b) \neq E_{I_N}a + E_{I_N}b\) may exist. However, any such operator scales sublinearly. That is, suppose that the PAPR reduction problem as just defined is solvable for \(I_N\) with constant \(C_{EX}\). If \(\|a\|_{\ell^2(I_N)} > 1\), we define \(a' = a/\|a\|_{\ell^2(I_N)}\). Then,

\[
\|E_{I_N}a'\|_{L^\infty(T)} \leq C_{EX},
\]

and we may simply rescale \(E_{I_N}a'\) by \(\|a\|_{\ell^2(I_N)}\) to determine an extension for \(a\) with bound \(C_{EX}\|a\|_{\ell^2(I_N)}\). Note that here the placement of the coefficients is unchanged, and they are all simply multiplied by the same appropriate scaling. Thus, solvability on all of \(\ell^2(I_N)\) and solvability on the unit ball of \(\ell^2(I_N)\) are equivalent, though the best constant in the latter case may be smaller.

**Definition 2.** The PAPR reduction problem is solvable for \(I_N\) with bound \(C_{EX}\) if there exists an operator \(E_{I_N} : \ell^2(I_N) \rightarrow L^\infty(\{-N, \ldots, N\})\) such that

\[
\|E_{I_N}a\|_{L^\infty(T)} \leq C_{EX}\|a\|_{\ell^2(I_N)},
\]

for every \(a \in \ell^2(I_N)\).

Now we proceed as follows: in Theorem 1 we give a necessary condition for solvability. With Theorem 2 we show that if it is required that the peak-to-average power ratio remains bounded, then the efficiency of the OFDM system converges to zero as the system size increases. That is, if the PAPR reduction problem remains solvable with the same bound for a sequence of sets \(\{I_N\}\) as \(N \rightarrow \infty\), then the relative density of the sets, that is, the ratio of information bearing signals to total signals must converge to zero.

**Definition 3.** For a subset \(I_N \subset \{-N, \ldots, N\}\) we define

\[
\mathcal{F}(I_N) = \left\{ f \in L^1(\mathbb{T}), f(t) = \sum_{k \in I_N} a_k e^{ikt} \right\}.
\]

**Theorem 1.** If the PAPR problem is solvable for the subset \(I_N\) with extension norm \(C_{EX}\), then

\[
\|f\|_{L^1(T)} \leq C_{EX}\|f\|_{L^1(\mathbb{T})},
\]

for all \(f \in \mathcal{F}(I_N)\).

**Proof.** By assumption, for all \(s(t) = \sum_{k \in I_N} a_k e^{ikt}\), \(\|a\|_{\ell^2(I_N)} \leq 1\),

\[
\|E_{I_N}a\|_{\ell^2(I_N)} \leq C_{EX}\|a\|_{\ell^2(I_N)} \leq C_{EX}.
\]

Again, by assumption,

\[
(E_{I_N}a)(t) = \sum_{k \in I_N} a_k e^{ikt} + \sum_{k \in R_N} b_k e^{ikt}.
\]

Let \(f \in \mathcal{F}(I_N)\), \(f(t) = \sum_{k \in I_N} c_k e^{ikt}\), be arbitrary. Then

\[
\left| \sum_{k \in I_N} a_k \overline{c_k} \right| = \left| \sum_{k \in I_N} a_k \overline{c_k} + \sum_{k \in R_N} b_k \overline{c_k} \right| = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \overline{E_{I_N}s(t)} dt \right| \leq \|f\|_{L^1(\mathbb{T})} \|E_{I_N}s\|_{L^\infty(T)} \leq C_{EX}\|f\|_{L^1(\mathbb{T})}.
\]

Set

\[
a_k = \begin{cases} \frac{c_k}{\|c\|_{\ell^2(I_N)}}, & c_k \neq 0, \\ 0, & \end{cases}
\]
Then
\[ \|f\|_{L^2(T)} = \|c\|_{\ell^2_{N+1}} = \left| \sum_{k \in I_N} a_k e_k \right| \leq C_{Ex} \|f\|_{L^2(T)}. \] (16)

The following definition gives the efficiency of the best subset selection for which the PAPR reduction problem is solvable for a given bound.

Definition 4 (Optimal subset size—OFDM).
\[ \mathcal{E}_N(C_{Ex}) = \max \{|I_N|; I_N \subset \{-N, \ldots, N\}, \text{ such that PAPR is solvable for } I_N \} \] (17)
with constant \( C_{Ex} \).

Now, we may state the following theorem.

Theorem 2. For all \( 0 < C_{Ex} < \infty \), the following limit holds:
\[ \lim_{N \to \infty} \frac{\mathcal{E}_N(C_{Ex})}{2N + 1} = 0. \] (18)

In other words, the theorem states that if a PAPR bound is always satisfied, then the system efficiency converges to 0 as the total size increases. Thus, the number of tones that may be used to carry information does not scale with \( N \). Theorem 2 gives a limiting value as the dimension approaches infinity, but it should not be read as a strictly asymptotic statement. Because of the convergence to zero, (18) rules out the existence of any arbitrarily large dimensions \( N \) for which solvability occurs or a certain parameter pair \( C_{Ex} \) and \( |I_N|/n \). In other words, given a constant \( C_{Ex} \) and a relative density, there is a limit to how large the dimension can be and satisfy both the extension constant \( C_{Ex} \) and the prescribed relative density.

The proof will use arithmetic progressions and Szemerédi’s Theorem, Theorem 3.

Definition 5. An arithmetic progression of length \( k \) is a subset of \( \mathbb{Z} \) that has the form \( \{a, a + d, a + 2d, \ldots, a + (k - 1)d\} \) for some integer \( a \) and some positive integer \( d \).

Theorem 3 (Theorem 1.2 in [7]). For any integer \( k \geq 1 \) and any \( 0 < \delta \leq 1 \), there exists an integer \( N_{SZ}(k, \delta) \) such that for every \( N \geq N_{SZ}(k, \delta) \), every set \( A \subset \{1, \ldots, N\} \) of cardinality \( |A| \geq \delta N \) contains at least one arithmetic progression of length \( k \).

Proof of Theorem 2. Assume that the claim is not true. Then, there exists a subsequence \( \{N_{k}\}_{k=1}^{\infty} \subset \mathbb{N} \) and a constant \( G(C_{Ex}) > 0 \) such that
\[ \frac{\mathcal{E}_{N_k}(C_{Ex}, \mathcal{F})}{2N_k + 1} \geq G(C_{Ex}) \] (19)
for all \( k = 1, 2, \ldots \). Now, we set \( \delta = G(C_{Ex})/2 \) and apply Szemerédi’s Theorem, Theorem 3. Thus, for any \( m \), there exists some large \( N \in \{N_k\}_{k=1}^{\infty} \) such that \( I_N \) contains an arithmetic progression of length \( m \). Denote this progression \( \{a + dl\}_{l=0}^{m-1} \). Now, note that
\[ \left\| \frac{1}{\sqrt{m}} \sum_{l=0}^{m-1} e^{i(a+dl)} \right\|_{L^2(T)} = 1, \] (20)
while
\[ \left\| \frac{1}{\sqrt{m}} \sum_{l=0}^{m-1} e^{i(a+dl)} \right\|_{L^1(T)} \leq \frac{\log(m/2)}{\sqrt{m}}. \] (21)
(This is the usual bound for the Dirichlet kernel.) Applying Theorem 1, for any fixed constant \( C_{Ex} \), (20) and (21) lead to the contradiction
\[ 1 \leq \left\| \frac{1}{\sqrt{m}} \sum_{l=0}^{m-1} e^{i(a+dl)} \right\|_{L^1(T)} \leq C_{Ex} \left\| \frac{1}{\sqrt{m}} \sum_{l=0}^{m-1} e^{i(a+dl)} \right\|_{L^2(T)} \leq C_{Ex} \frac{\log(m/2)}{\sqrt{m}} \] (22)
when \( m \) is large.

Thus, if a bound on the peak of all transmission signals is given as one increases the number of total tones available, at some size \( N \) the proportion of tones allocated to carry information must decrease in order to satisfy the PAPR constraint.

From this theorem, we also see that when tone reservation is used, the subsets chosen as information and compensation sets are very important. In particular, the information set should not have any long arithmetic progressions; however, determining subsets with little additive structure is a very challenging problem. As an indication of this, consider that what is now known as Szemerédi’s theorem was an open question for length 3 for decades before Roth proved it in 1952 [8], for which he was awarded the Fields Medal in 1958. Szemerédi proved the result for length 4 in 1969 [9] and his final result in 1975 [10]. Sets with only short arithmetic progressions is certainly a nearly equivalent problem and thus also very difficult. For a taste of this area, one may see Chapter 2 of [11].

The following theorem shows that if the PAPR reduction problem is solvable for a finite information set when we allow the compensation set to be the entire rest of the integers; then, using a projection onto a finite set, we obtain solvability for the finite compensation set. However, the extension norm increases depending on the size of the compensation set.

Theorem 4. Suppose that \( I_N \) is a subset of \( \{-N, \ldots, N\} \) and that for every \( a \in \mathbb{F}(\mathbb{Z}) \) supported on \( I_N \) the PAPR reduction problem is solvable with an extension sequence supported on \( \mathbb{Z} \setminus I_N \) with extension bound \( C_{Ex} \). Assume that \( \lambda > 1 \) and that \( \lambda N \) is an integer. Then the PAPR reduction problem is also solvable with an extension sequence supported on \( \{-\lambda N, \ldots, \lambda N\} \setminus I_N \) with extension constant \( (2/(\lambda - 1))C_{Ex} \).
Proof. Let $f(t) = \sum_{n \in I_N} a_n e^{int}$, and let

$$f(t) + g(t) = \sum_{n \in I_N} a_n e^{int} + \sum_{n \in \mathbb{Z} \setminus I_N} a_n e^{int}$$

be its extension, such that $\|f + g\|_{L^2(T)} \leq C_{\text{Ex}} \|a\|_{F(\mathbb{Z})}$. We simply project $f + g$ onto span $\{e^{int} \}_{n = -N_{-1}N}$ using a Fejér kernel. We define the following set of kernels:

$$\mathcal{K}_{N, \lambda} = \left\{ K(t) = \sum_{n = -N}^{N} d_n e^{int} + \sum_{n = N+1}^{\infty} d_n e^{int}, \right\}$$

where $d_n = d_{-n}$ for $n = N + 1, \ldots, \lambda N$.

For any kernel $K \in \mathcal{K}_{N, \lambda}$, we have

$$f(t) = \frac{1}{2N + 1} \sum_{n = -N}^{N} f\left(\frac{2\pi t}{2N + 1}\right) K\left(t - \frac{2\pi t}{2N + 1}\right),$$

and we may define $f + g_{\lambda N}$ to be the convolution of $f + g$ with $K$. The Fourier expansion of $f + g_{\lambda N}$ is supported on $\{-\lambda N, \ldots, \lambda N\}$ and agrees with $a$ on $I_N$. Using $P_K$ to denote the projection given by $K$,

$$\|f + g_{\lambda N}\|_{L^2(T)} \leq \|P_K f + \|g\|_{L^2(T)} \leq C_{\text{Ex}} \|a\|_{F(\mathbb{Z})} \|P_K\|.$$

(26)

The norm $\|P_K\|$ is the $\| \cdot \|_{L^2(T)}$-norm of $K$. We will construct $K$ using two Fejér kernels. We recall that the Dirichlet kernel is defined by

$$D_n(t) = \sum_{k = -n}^{n} e^{ikt},$$

and the Fejér kernel by

$$F_n(t) = \frac{2n}{n+1} \sum_{k=0}^{n} D_k = \left(\frac{\sin(nt/2)}{\sin(t/2)}\right)^2.$$

(28)

Thus, for any $m > l$,

$$\sum_{k=0}^{m} D_k - \sum_{k=0}^{l} D_k = (m-l) \sum_{k=0}^{2l-m} e^{ikt} + e^{-ikt} + \sum_{k=1}^{2(m-l)} (m-l-k) e^{i(2l-m+k)t} + e^{-i(2l-m+k)t}.$$

(29)

By setting $d_k = (\lambda N - N - (k/2))$ for $k = N + 1, \ldots, \lambda N$ and $K(t) = 1/(\lambda N - N) (\sum_{k=0}^{N} D_k(t) - \sum_{k=0}^{N} D_k(t))$ and using the positivity given in (29), we obtain

$$\|K\|_{L^1([0,1])} = \int_0^{1/\lambda N - N} \left| \frac{1}{\lambda N - N} \left( \sum_{k=0}^{N} D_k(t) - \sum_{k=0}^{N} D_k(t) \right) \right| dt$$

$$\leq \frac{1}{\lambda N - N} \int_0^{1/\lambda N - N} \sum_{k=0}^{N} D_k(t) + \sum_{k=0}^{N} D_k(t) dt$$

$$\leq \frac{2\lambda N}{\lambda N - N} = \frac{2}{\lambda - 1}.$$ 

(30)

Returning to (26), we have

$$\|f + g_{\lambda N}\|_{L^2(T)} \leq \frac{2}{\lambda - 1} C_{\text{Ex}} \|a\|_{F(\mathbb{Z})},$$

(31)

where the Fourier expansion of $g_{\lambda N}$ is supported on $\{-\lambda N, \ldots, \lambda N\} \setminus I_N$.

\[ \square \]

3. The Infinite Set Case

The infinite set case is particularly important because of the insight it brings to the mathematical structure of tone reservation. By using the projections discussed in Theorem 4 though, the infinite set case also has practical implications. Our first step en route to proving the infinite-dimensional form of Theorem 2 is to prove an equivalence between the PAPR reduction problem and a norm equivalence. Recall that Theorem 1 stated that solvability implies a norm equivalence in the finite set case. In the infinite set case, we show that solvability holds if and only if the norm equivalence holds. By constructing functions that violate the norm relation, we will show when solvability cannot hold (Theorem 6). However, using a special case when the norm relation holds, we will use the if and only if statement to identify sets where solvability does hold (Theorem 7).

The equivalence statement holds for arbitrary orthonormal systems, so we state it in that generality. This is Theorem 5. In Theorem 6, we prove that the PAPR problem is not solvable in the OFDM setting at positive efficiencies for sets of infinite cardinality. Let $\{\psi_k\}_{k \in \mathbb{Z}}$ be an orthonormal basis for $L^2(\mathbb{T})$. Let $K$ be a subset of $\mathbb{Z}$ and define

$$X := \left\{ f \in L^1(\mathbb{T}) : f(t) = \sum_{k \in K} a_k \psi_k(t) \right\}.$$ 

(32)

Given a function $s \in X$, we are interested in finding a compensation function $g$, $g(t) = \sum_{k \in K} b_k \psi_k$, such that $\|s + g\|_{C_k} \leq C_{\text{Ex}} \|s\|_{C_k}$. Here, we may view the nonlinear operator as a map from $L^2(\mathbb{T})$ to $L^2(\mathbb{T})$, so that $E_k s = s + g$.

If a map exists so that such a $g$ can be found for every $s \in X$, then we say that the PAPR reduction problem is solvable for the pair $K$ and $\{\psi_k\}_{k \in \mathbb{Z}}$ with extension norm $C_{\text{Ex}}$. 


Theorem 5. The PAPR problem is solvable for the pair $K$ and \{\psi_k\}_{k \in \mathbb{Z}}$ with extension norm $E_K$ if and only if
\[
\|f\|_{L^2(T)} \leq C_{E_K}\|f\|_{L^1(T)},
\] (33)
for all $f \in X$.

Remark 1. We give a short discussion of the geometric interpretation of this theorem in our conclusion in Section 5.

When (33) holds, we will say that $X$ has the norm equivalence property. Note that in contrast to the finite set case of Theorem 1, here we have a necessary and sufficient condition for solvability.

Before giving the proof of Theorem 5 we emphasize several points. If $C_{E_K}$ is the smallest constant for which the reduction problem is solvable, then it is also the smallest constant for which (33) holds and vice versa. To see this, suppose that $C_1$ is the smallest constant for which solvability holds, but the norm equivalence holds for $C_1 < C_1$. By the equivalence statement of Theorem 5, solvability also holds for $C_2$, which contradicts that $C_1$ is the smallest such constant. The same argument applies for the opposite implication as well. The statement that the best constant for the norm relation is also the best constant for solvability is the more significant, since here, this property follows from the Hahn-Banach Theorem. The Hahn-Banach Theorem, however, requires the axiom of choice to repeatedly extend the functional's domain by one dimension. Indeed, here the axiom of choice is used to justify this because in most settings it cannot be done constructively. Therefore, constructing an extension with the optimal constant is equivalent to realizing what the axiom of choice states exists, but for which a method of construction does not exist. This shows how mathematically complex an optimal implementation of tone reservation is.

Proof. (i) Assume that the PAPR problem is solvable. Then, for all $s(t) = \sum_{k \in K} a_k \psi_k(t)$, $\|a\|_{E^0(\mathbb{Z})} \leq 1$,
\[
\|E_K s\|_{L^\infty(T)} \leq C_{E_K} \|s\|_{L^2(T)} \leq C_E^{10}.
\] (34)
Since $L^\infty(T) \subset L^2(T)$,
\[
E_K s = \sum_{k \in K} a_k \psi_k + \sum_{k \in \mathbb{Z} \setminus K} b_k \psi_k.
\] (35)
Let $f \in X$, $f(t) = \sum_{k \in K} c_k \psi_k(t)$, be arbitrary. Then,
\[
\left| \sum_{k \in K} a_k c_k \right| = \left| \sum_{k \in K} a_k \bar{\psi}_k + \sum_{k \in \mathbb{Z} \setminus K} b_k \bar{\psi}_k \right| = \left| \frac{1}{2\pi} \int_T f(t) \overline{E_K s(t)} dt \right| \leq \|f\|_{L^2(T)} \|E_K s\|_{L^\infty(T)} \leq C_{E_K} \|f\|_{L^2(T)}.
\] (36)

Set
\[
a_k = \begin{cases} \frac{c_k}{\|c\|_P} & c_k \neq 0, \\ 0 & c_k = 0. \end{cases}
\] (37)
Then, $\|f\|_{L^2(T)} = \|c\|_P = \|\sum_{k \in K} a_k \bar{\psi}_k\| \leq C_{E_K} \|f\|_{L^2(T)}$.

(ii) Assume $\|f\|_{L^1(T)} \leq C_{E_K} \|f\|_{L^1(T)}$ for all $f \in X$. Let $a \in \ell^2(\mathbb{Z})$ be a sequence supported in $K$ with only finitely many nonzero terms satisfying $\|a\|_{\ell^2(\mathbb{Z})} \leq 1$. Set $s(t) = \sum_{k \in K} a_k \psi_k(t)$. For $f \in X$, $f(t) = \sum_{k \in K} c_k \psi_k(t)$, define the functional $\Psi_a$ by
\[
\Psi_a f = \sum_{k \in K} a_k c_k.
\] (38)
Since
\[
\|\Psi_a f\| \leq \|a\|_{\ell^2(\mathbb{Z})}\|c\|_{\ell^2(\mathbb{Z})} \leq \|f\|_{L^1(T)} \leq C_{E_K} \|f\|_{L^1(T)},
\] (39)
$\Psi_a$ is continuous on $X$. Since $X$ is a closed subspace of $L^1(\mathbb{Z})$, by the Hahn-Banach Theorem [12], the functional $\Psi_a$ has the extension $\Psi_a$ to all of $L^1(\mathbb{Z})$, where $\|\Psi_a\| = \|\Psi_a\|_{\ell^2(\mathbb{Z})}$. The dual of $L^1(\mathbb{Z})$ is $L^\infty(\mathbb{Z})$. Thus, for some $r \in L^\infty(\mathbb{Z})$,
\[
\Psi_a f = \langle f, r \rangle,
\] (40)
for all $f \in L^1(\mathbb{Z})$, so that $\|\Psi_a\| = \|r\|_{L^\infty(\mathbb{Z})}$. Since $L^\infty(\mathbb{Z}) \subset L^2(\mathbb{Z})$, $r$ possesses the unique expansion
\[
r(t) = \sum_{k \in \mathbb{Z}} d_k \psi_k(t),
\] (41)
for some $d \in \ell^2(\mathbb{Z})$. The sequences $d$ and $a$ agree on $K$, and we define $E_K s := r$. $\Box$

Theorem 6. For $K \subset \mathbb{Z}$, let $S(N) = K \cap \{-N, \ldots, 0, \ldots, N\}$. If $\limsup_{N \rightarrow \infty} (|S(N)|/(2N + 1)) > 0$, then the PAPR problem is not solvable for $K$ and the Fourier basis $\{e^{ik}\}_{k \in \mathbb{Z}}$.

In particular, this theorem states that if the ratio of the number of basis functions used for transmission to the total number of basis functions does not tend to zero, then arbitrarily high peaks can be constructed that cannot be sufficiently damped by any compensation function.

Similar questions concerning the sizes of subsets of orthonormal bases that have a norm equivalence have been studied. In [13], Bourgain addresses an $L^2 - L^p$ norm equivalence for $p > 2$. The general technique used here to determine a norm equivalence is well known in the functional analysis and local Banach space community.

Proof. First suppose that the PAPR problem is solvable for $K$ and $\{e^{ik}\}_{k \in \mathbb{Z}}$. We develop a contradiction to the equivalence given in Theorem 5. Suppose that arbitrary subsets $S(N)$ of $\{-N, \ldots, N\}$ are chosen such that
\[
\limsup_{N \rightarrow \infty} \frac{|S(N)|}{2N + 1} = \delta > 0.
\] (42)
For any positive integer \( k \), using Szemerédi’s Theorem (Theorem 3) again, there exists a large integer \( N \) such that \( S(N) \) has an arithmetic progression of length \( k \). Denote the arithmetic progression \( \{a + dl\}_{l=0}^{k-1} \). We again have

\[
\left\| \frac{1}{\sqrt{K}} \sum_{l=0}^{k-1} e^{i(a+dl)} \right\|_{L^2} = 1, \tag{43}
\]

while

\[
\left\| \frac{1}{\sqrt{K}} \sum_{l=0}^{k-1} e^{i(a+dl)} \right\|_{L^1} \leq \log(k/2) \sqrt{K}. \tag{44}
\]

Applying Theorem 5, for any fixed constant \( C_{\text{Ex}} \), for large enough lines (43) and (44) give the contradiction

\[
1 = \left\| \frac{1}{\sqrt{K}} \sum_{l=0}^{k-1} e^{i(a+dl)} \right\|_{L^2} \leq C_{\text{Ex}} \left\| \frac{1}{\sqrt{K}} \sum_{l=0}^{k-1} e^{i(a+dl)} \right\|_{L^1} \leq C_{\text{Ex}} \frac{\log(k/2)}{\sqrt{K}}.
\]

We point out that the sequence of coefficients used to give the contradiction is not at all exotic—it is just a sequence of 1’s placed at the right locations.

If we work with a finite total number of tones and have an extension constant \( C_{\text{Ex}} \), then since the constant \( C_{\text{Ex}} \) in both aspects of Theorem 5 is the same, we can deduce a bound on the longest arithmetic progression in \( I_N \). Namely, denoting by \( k \) the length of the longest arithmetic progression, we have \( 1 \leq C_{\text{Ex}}(\log(k)/\sqrt{K}) \).

To emphasize the role of the density, we contrast Theorem 6 with the following theorem.

**Theorem 7** ([14] and Theorem III.F.6 in [15]). Let \( \lambda > 1 \) be a real number and assume that the subset \( K = \{n_k\}_{k=1}^\infty \subset \mathbb{Z} \) has the property \( |n_{k+1}| \geq \lambda |n_k| \) for all \( k \leq 1 \). Then there exists a constant \( C_K \) such that for all \( a \in I^1 \) supported on \( K \) there exists a continuous function \( g \in L^2(\mathbb{T}) \) satisfying \( \|g\|_{L^2} \leq C_K \|a\|_{L^2} \) with Fourier coefficients satisfying \( \hat{g}_{n_k} = c_{n_k} \). That is, a compensation signal exists.

In the case addressed in Theorem 7, a compensation signal can always be found. But, of course, the difference is that the density of \( K \) is zero: for every \( M \) elements of \( K \) we have roughly \( \lambda^M \) elements in the compensation set.

Now we give an example of when uncontrollable peaks can occur (see also Section 5.2 of [16]). Consider a sequence \( c \in I^1(\mathbb{Z}) \) supported on the positive integers, and such that \( \lim_{K \to \infty} \sum_{k=1}^K c_k = \infty \). Let \( M \) be a large positive integer and set \( I_{N,M} = N2^{-M} \) for any \( N \leq 2^{-M} \). Set \( K_{N,M} \) to be the smallest value for \( k \) such that \( 2^{k-M}N \) is an integer. We then have

\[
\sum_{k=1}^K c_k e^{2\pi i k l} = \sum_{k=1}^{K_{N,M}} c_k e^{2\pi i k l} + \sum_{k=K_{N,M}+1}^k c_k e^{2\pi i k l}. \tag{46}
\]

Clearly, the term on the left remains bounded, while the term on the right tends to infinity as \( K \to \infty \). However, since the indices \( \{2^k\}_{k=0}^\infty \) correspond to \( \lambda = 2 \) in Theorem 7, there exists a sequence \( a \) supported on \( \mathbb{Z} \setminus \{2^k\}_{k=0}^\infty \) such that \( \|a\|_{L^1} \leq C_{\text{Ex}} \|c\|_{L^1} \) for a constant independent of \( c \) and

\[
\left\| \sum_{k \in \mathbb{Z}} c_k e^{2\pi ikl} \right\|_{L^1(\{0,1\})} \leq C_{\text{Ex}}. \tag{47}
\]

One could ask if all subsets of \( Z \) of density zero might correspond to a subspace for which the PAPR reduction problem is solvable or, equivalently, for which a norm equivalence holds. Here, however, a famous result tells us that this is not the case. That is, Green and Tao have proved the following theorem.

**Theorem 8** (see [17]). The prime numbers contain arithmetic progressions of arbitrary length.

Thus, the primes are an example of a subset of the natural numbers with density zero, but for which the PAPR reduction problem is not solvable.

### 4. The Discrete Fourier Case

For completeness, we include a short section on the discrete Fourier case. The discrete setting is important, because it is often used to model or approximate the analog setting. Using sampling theorems, one can then relate results from the discrete setting to the continuous setting. This is done in the papers [6, 18, 19]. Also, if a signal has been digitized, then it falls into the discrete, finite-dimensional setting. The proofs and a more thorough discussion of the work presented here can be found in [20].

**Definition 6.** The \( N \times N \) inverse discrete Fourier transform (DFT) matrix is given by

\[
F_{jk} = \frac{1}{\sqrt{N}} e^{-2\pi i (j-1)(k-1)/N}. \tag{48}
\]

This matrix is denoted \( F \), and for \( x \in I^2_N \), \( Fx \) denotes this matrix applied to \( x \).

**Definition 7.** The unit ball in \( I^r_N \) is denoted \( B^r_N \); that is,

\[
B^r_N = \left\{ x \in I^r_N : \|x\|_{I^r_N} \leq 1 \right\}. \tag{49}
\]

Since for any finite \( N \) solvability in the finite-dimensional setting holds for some constant, we must consider a sequence
of subsets for an increasing sequence of dimensions and address whether the solvability constant remains bounded. Thus, we assume that \( \{N_k\}_{k=1}^\infty \) is an increasing sequence or natural numbers, and for each \( N_k \) we have a subset \( I_{N_k} \) of \( \{1,\ldots,N_k\} \). Here, \( I_{N_k} \) denotes \( \{1,\ldots,N_k\} \setminus I_{N_k} \). Similarly to Definition 2, the discrete PAPR problem is solvable for the sequences \( \{N_k\}_{k=1}^\infty \) and \( \{I_{N_k}\}_{k=1}^\infty \) if there exists a constant \( C_{E_{k}} \), such that for each \( k \), for all \( x \in \mathbb{R}_{N_k} \) with \( \text{supp}(x) \subset I_{N_k} \), there exists a compensation vector \( r \in \mathbb{R}_{N_k} \) supported in \( I_{N_k} \) such that

\[
\|F(x + r)\|_{\ell_{N_k}^2} \leq C_{E_{k}} \|x\|_{\ell_{N_k}^2}.
\]

**Theorem 9** (see [20]). Let \( \{N_k\}_{k=1}^\infty \) be a subsequence of \( \mathbb{N} \), and let \( I_{N_k} \) be a subset of \( \{1,\ldots,N_k\} \). Let \( Y_k = \{y \in \mathbb{R}_{N_k} : \text{supp}(F^* y) \subset I_{N_k}\} \). The discrete PAPR problem is solvable for the sequence of sets \( \{I_{N_k}\}_{k=1}^\infty \) with constant \( C_{E_{k}} \) if and only if

\[
\|y\|_{\ell_{N_k}^2} \leq C_{E_{k}} \|y\|_{\ell_{N_k}^2}
\]

for all \( y \in Y_k \).

Note that \( \|y\|_{\ell_{N_k}^2} \leq \|y\|_{\ell_{N_k}^2} \) holds for any vector \( y \). But as \( k \) increases, at some point \( C_{E_{k}}/\sqrt{N_k} < 1 \), and so the important point is that \( C_{E_{k}} \) remains fixed.

For a set \( A \), \( |A| \) denotes its cardinality.

**Theorem 10** (see [20]). Let \( \{N_k\}_{k=1}^\infty \) be a subsequence of \( \mathbb{N} \), and let \( I_{N_k} \) be the corresponding sets as defined earlier. If

\[
\lim_{n \to \infty} \frac{|I_{N_k}|}{N_k} > 0,
\]

then the discrete PAPR problem is not solvable.

The proof uses the finite-dimensional equivalence relation in Theorem 9 and Szemerédi’s Theorem, just as the proof of Theorem 6 used the \( L^2 - L^1 \) equivalence relation, Theorem 5, and Szemerédi’s Theorem.

### 5. Discussion and Conclusion

The approach to the PAPR reduction problem presented here has two parts: first, we showed an equivalence between solvability and a norm relation, and then we considered cases when we could violate or satisfy the norm relation. There is a very interesting difference between these two parts in that the equivalence statement holds for any orthonormal basis and proving the statement did not require working with the functions. The second step, however, required using the special structure of the Fourier system. In [20], we present the analogous result for the Walsh or CDMA system, and there we also use every special property of that system. We make the heuristic conjecture that if an orthonormal basis has structure, then one can construct a function in the span of a subset of elements with positive relative density that violates the norm relation.

Compare this conjecture with the following observation. Paley’s Theorem (Theorem I.B.24 in [15]) states that there exists a constant \( C \) such that for \( f(t) = \sum_{k=1}^n a_k e^{2\pi i a_k t} \) for any \( n \in \mathbb{N} \)

\[
\|f\|_{L^2((0,1))} \leq C\|f\|_{L^1((0,1))}.
\]

Thus, if we rearrange the Fourier basis by repeatedly taking one element from the set \( \{e^{2\pi i k}\}_{k=1}^\infty \) and one from its complement, we can achieve solvability for a subset with a relative density of one half, by the definitions used here. What is lost, however, is the structure. Here, the highest frequency grows exponentially with respect to the total number of signals.

The Khinchine inequality (I.B.8 in [15]) is the analogous statement to Paley’s Theorem for the Walsh basis. Here, again, one could gain solvability by taking the elements with indices \( \{2^k\}_{k=1}^\infty \), but one then loses the structure and the desireable properties. For example, with the natural ordering one can move from the infinite setting to the finite setting with the optimal projection constant 1. But if one deviates from this ordering, this very helpful property is lost. This relationship between structure and norm equivalences requires much further study.

This paper shows that the PAPR reduction problem is a challenging mathematical problem, both for general orthonormal systems and for specific systems, such as OFDM. Specific orthonormal systems demonstrate their own difficult properties: for example, with OFDM we have seen that the ordering of the elements plays an important role and allows one to introduce the density that was investigated here. Thus, the notion of density, as investigated here, is fundamentally tied to the OFDM system.

On the other hand, there are results that hold for all orthonormal systems. Here, for example, we see that the worst case PAPR behavior of an OFDM signal is just an instance of the \( \sqrt{N} \) behavior of any \( N \) orthonormal signals [16]. A theorem of Kashin and Tzafriri [21] states that under very reasonable conditions the expected maximum of a signal formed as a linear combination of \( N \) orthonormal functions on the unit interval behaves like \( \sqrt{\log(N)} \). This is discussed in [22], where it is also shown that a signal’s PAPR value is highly concentrated in probability around its expectation. This may be compared to probabilistic bounds for the OFDM system in [23]. Thus, even in expectation, the PAPR behavior of the OFDM system is, in fact, just an instance of a uniform behavior. Understanding these properties is particularly important for the emerging effort in communications research to address overall energy consumption, rather than considering only the energy of the transmitted signal. In particular, one must address how much energy is necessary to create a certain signal in terms of the signal’s properties. In this way, one can address the relationship between signal properties, such as PAPR, and circuit design with the goal of reducing energy consumption at the circuit level; see [24] for a recent discussion.

This paper is a contribution to the mathematical understanding of peak-to-average power ratio and, in particular, tone reservation. We have shown that tone reservation is inherently related to such mathematical areas as the Hahn-Banach Theorem, norm equivalences, and arithmetic
progressions, and presented results in each of these areas. The general result of the paper is that a fixed portion of subcarriers is always allocated for carrying information, then signal peaks cannot be held below a fixed threshold.

Acknowledgment

H. Boche was supported by startup funds of the Technische Universität München, and B. Farrell was supported by the German Research Foundation (DFG) under Project BO 1734/18-1.

References

[1] S. H. Han and J. H. Lee, “An overview of peak-to-average power ratio reduction techniques for multicarrier transmission,” IEEE Transactions on Wireless Communications, vol. 12, no. 2, pp. 56–65, 2005.

[2] J. Tellado and J. M. Cioffi, “Efficient algorithms for reducing par in multicarrier systems,” in Proceedings of the IEEE International Symposium on Information Theory (ISIT ’98), p. 191, Cambridge, Mass, USA, August 1998.

[3] B. S. Krongold and D. L. Jones, “PAR reduction in OFDM via active constellation extension,” IEEE Transactions on Broadcasting, vol. 49, no. 3, pp. 258–268, 2003.

[4] R. Bäuml, R. F. H. Fischer, and J. B. Huber, “Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping,” IEEE Transactions on Information Theory, vol. 32, pp. 2056–2057, 1996.

[5] M. Breiling, S. H. Müller-Weinfurtner, and J. B. Huber, “SLM peak-power reduction without explicit side information,” IEEE Communications Letters, vol. 5, no. 6, pp. 239–241, 2001.

[6] J. Ilic and T. Strohmer, “PAPR reduction in OFDM using kashin’s representation,” in Proceedings of the 10th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC ’09), pp. 444–448, June 2009.

[7] T. Tao, “A quantitative ergodic theory proof of Szemerédi’s theorem,” Electronic Journal of Combinatorics, vol. 13, no. 1, pp. 1–49, 2006.

[8] K. F. Roth, “Sur quelques ensembles d’entiers,” Comptes-Rendus de l’Académie des Sciences de Paris, vol. 234, pp. 388–390, 1952.

[9] E. Szemerédi, “On sets of integers containing no four elements in arithmetic progression,” Acta Mathematica Academiae Scientiarum Hungaricae, vol. 20, no. 1-2, pp. 89–104, 1969.

[10] E. Szemerédi, “On sets of integers containing no k elements in arithmetic progression,” Acta Arithmetica, vol. 27, pp. 199–245, 1975.

[11] T. Tao and V. Vu, Additive Combinatorics, Cambridge University Press, Cambridge, Mass, USA, 2006.

[12] W. Rudin, Real and Complex Analysis, McGraw-Hill, New York, NY, USA, 3rd edition, 1987.

[13] J. Bourgain, “Bounded orthogonal systems and the \( \Lambda(p) \)-set problem,” Acta Mathematica, vol. 162, no. 1, pp. 227–245, 1989.

[14] R. E. A. C. Paley, “On the lacunary coefficients of power series,” Annals of Mathematics, vol. 34, no. 3, pp. 615–616, 1933.

[15] P. Wojtaszczyk, Banach Spaces for Analysts, vol. 25 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, Mass, USA, 1991.

[16] H. Boche and V. Pohl, “Signal representation and approximation-fundamental limits,” European Transactions on Telecommunications, vol. 18, no. 5, pp. 445–456, 2007.

[17] B. Green and T. Tao, “The primes contain arbitrarily long arithmetic progressions,” Annals of Mathematics, vol. 167, no. 2, pp. 481–547, 2008.

[18] C. Tellambura, “Computation of the continuous-time PAR of an OFDM signal with BPSK subcarriers,” IEEE Communications Letters, vol. 5, no. 5, pp. 185–187, 2001.

[19] G. Wunder and H. Boche, “Peak value estimation of bandlimited signals from their samples, noise enhancement, and a local characterization in the neighborhood of an extremum,” IEEE Transactions on Signal Processing, vol. 51, no. 3, pp. 771–780, 2003.

[20] H. Boche and B. Farrell, “On the peak-to-average power ratio reduction problem for orthonormal transmission schemes,” Submitted.

[21] B. Kashin and L. Tzafriri, “Lower estimates for the supremum of some random processes,” East Journal on Approximations, vol. 1, no. 1, pp. 125–139, 1995.

[22] B. Farrell and H. Boche, “Worst case and expected peak-to-average power ratio for orthonormal systems,” in Proceedings of the IEEE International Conference on Communications (ICC ’11), 2011.

[23] G. Wunder and H. Boche, “Upper bounds on the statistical distribution of the crest-factor in OFDM transmission,” IEEE Transactions on Information Theory, vol. 49, no. 2, pp. 488–494, 2003.

[24] A. Mezghani and A. Nossek, “Power efficiency in wireless networks from a circuit perspective,” in Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS ’11), 2011.