

On the Peak-to-Average Power Ratio Reduction Problem for Orthogonal Transmission Schemes

Holger Boche and Brendan Farrell

Abstract. High peak values of transmission signals in wireless communication systems lead to wasteful energy consumption and out-of-band radiation. However, reducing peak values generally comes at the cost of some other resource. We provide a theoretical contribution toward understanding the relationship between peak value reduction and the resulting cost in information rates. In particular, we address the relationship between peak values and the proportion of transmission signals allocated for information transmission when one is using a strategy known as tone reservation. We show that when tone reservation is used in both OFDM and DS-CDMA systems, if a peak-to-average power ratio criterion is always satisfied, then the proportion of transmission signals that may be allocated for information transmission must tend to zero. We investigate properties of these two systems for sets of both finite and infinite cardinalities. We present properties that OFDM and DS-CDMA share in common as well as ways in which they fundamentally differ.

1. Introduction

Recent studies by two prominent consulting firms have estimated that 2% of global CO₂ emissions are attributable to the use of information and

communication technology, a contribution comparable to aviation [Boccaletti et al. 08, Gartner 07, POST 08]. While this impact is already significant, the amount of information communicated electronically is growing exponentially, and the emissions percentage is expected to increase to 3% by 2020. A large portion of the energy consumption that causes emissions is due to wireless communications, and within wireless systems, a significant portion of this energy consumption occurs at the amplifiers. Communications companies must deal with a tradeoff between expensive amplifiers that are efficient, a capital expenditure, and inexpensive amplifiers and high energy costs, an operating cost. The increase in the volume of wireless communication requires the implementation of systems that place more and more individual signals in a frequency band, and this, inherently, leads to larger signal amplitudes. Thus, from both an environmental and financial perspective, the interplay between information capacity, signal peak values, amplifier performance, and energy consumption is an essential area for research. The current understanding is that amplifiers are more efficient when transmitting signals that have smaller peak values. See [Raab et al. 02] for an overview.

While keeping the broader considerations of power consumption, amplifier efficiency, and capacity in mind, we focus in this paper on the relationship between signal peak values and the proportion of signal resources that can be allocated for information transmission. We address the balance between allocating resources toward reducing signal peak values and allocating them for transmitting information. To the best of the authors' knowledge, there has been very little theoretical work done in this area, and little is known about the fundamental relationships between these various aspects of wireless communication.

We focus our attention on two of the most important contemporary communications systems, namely orthogonal frequency division multiplexing (OFDM) and direct sequence-code division multiple access (DS-CDMA). Both of these use a classical basis for transmission signals: OFDM uses the Fourier basis, and DS-CDMA uses the Walsh basis. In either case, coefficients are chosen to represent a message, and the linear combination corresponding to these coefficients is transmitted. As part of transmission, the signal passes through an amplifier. Every amplifier has a threshold beyond which it cannot linearly amplify the signal, but distorts or "clips" it. We say that a signal is clipped at magnitude M if the signal is left undisturbed when its magnitude is less than M , and its magnitude is reduced to M when it is greater than M while leaving the phase unchanged. Both distortion and clipping take place only above some threshold, so that one may say that only the peaks are affected.

Transmission signals generally also satisfy a frequency-band requirement, that is, their Fourier transforms are supported in a specified region or the signal is band-limited. This results in the signals being analytic, i.e., infinitely

differentiable with convergent power series. If the distorted signal differs from the original only when the original magnitude exceeds a threshold, then the two signals agree whenever the original is below that threshold. If the distorted function is also band-limited, then it is analytic, and the difference between original and distorted signals is zero on an open interval. A basic theorem of complex analysis then implies that their difference is identically zero. Thus, the distorted function cannot be band-limited. That is, if clipping or distortion occurs, then the amplified signal is not band-limited, and out-of-band radiation occurs. This gives the motivation for the approach taken in this paper.

If the transmitted signal is not band-limited, it interferes with other frequency bands. If one is interested in capacity or error rates without a requirement that out-of-band radiation not occur, then a probabilistic approach to the peak behavior is appropriate. However, in many instances, out-of-band radiation may be strictly prohibited.

With the transition from analog to digital television transmission, the improved efficiencies allowed new frequency bands to be redistributed, in particular for wireless communications. This is commonly called the “digital dividend.” Very strict quality of service requirements have been imposed on operators for some of these bands, and a percentage-based compliance is insufficient. An example is wireless microphones, whereby users are of course very sensitive to a disruption of service. In these cases, statistical models are inadequate.

The strategy we consider here is known in the OFDM setting as tone reservation. This method was introduced in [Tellado and Cioffi 98, Tellado 99]; an overview is given in [Han and Lee 05, Litsyn 07], and a survey of recent advances is given in [Wunder et al. 13]. We will apply this strategy as well to the CDMA setting. Here one separates the available transmission signals into two subsets. Coefficients that carry the message are then applied to signals in one subset, and then coefficients are determined for the signals in the second subset such that the peak of the entire composite signal is ideally below a threshold. There are other methods to reduce the peak value, such as selected mapping and clipping and filtering. However, tone reservation is canonical in that the coefficients to be transmitted are not altered in any way, the auxiliary coefficients may simply be ignored by the receiver, and there is no additional overhead to transmission. We note that the literature on these topics is enormous, with some papers cited several hundred times. Extensive references of the most important works are available in [Litsyn 07, Wunder et al. 13].

Here we address how the ratio of information-bearing signals to compensation signals behaves with respect to a peak threshold as the total number of available signals increases. The two main results presented here, Theorem 3.4 in the Fourier-OFDM case and Theorem 4.2 in the Walsh-CDMA case, show that if a

peak threshold must always be satisfied, then the proportion of signals that may be used to carry information converges to zero. While the OFDM and CDMA share this property in common, they behave quite differently in other significant ways, which we discuss in later sections. These main results are coupled with two other main points. This first is a relationship between what we will call *solvability* and a norm equivalence, and the second is a fundamentally different behavior when these questions are addressed for sets of finite or infinite cardinality. The relationship between solvability and norm equivalence is presented in Section 2. The Fourier–OFDM case is addressed in Section 3, and the density result for the Walsh–CDMA case is presented in Section 4.1. Section 4.2 gives further properties of the Walsh system and in particular, emphasizes their localized nature. Some of these results were presented at a special session of the IEEE Conference on Vehicular Technology in Dresden, Germany, in June 1013.

While this paper addresses Fourier–OFDM and Walsh–CDMA systems, we note that recent results for the peak-value behavior of single-carrier (sinc) systems were obtained in [Boche et al. 12]. There it is shown that the expected peak of a random linear combination of shifted sinc functions grows with the number of such functions. This underscores the prevalence of high peak amplitudes in communications systems.

2. Solvability and a Norm Equivalence for Orthonormal Systems

2.1. Introductory Facts

We first formalize our problem and then introduce an important aspect of our approach. We begin with the following definition, where without loss of generality, we take $[0, 1]$ as the symbol interval.

Definition 2.1. Given a set of orthonormal functions $\{\phi_n\}_{n=1}^N \subset L^2([0, 1])$, we define the *peak-to-average power ratio (PAPR)* of a set of coefficients $a \in \mathbb{C}^N$ by

$$\text{PAPR}(\{\phi_n\}_{n=1}^N, a) = \text{ess sup}_{t \in [0, 1]} \frac{|\sum_{n=1}^N a_n \phi_n(t)|}{\|a\|_{\ell^2_N}}.$$

The following simple proposition shows that PAPR values of order \sqrt{N} can occur for any orthonormal system. We include the proof just for the sake of completeness.

Proposition 2.2. [Boche and Pohl 07, Theorem 6] *Let $\{\phi_k\}_{k=1}^N$ be N orthonormal functions in $L^2([0, 1])$. Then there exists a sequence $a \in \ell_N^2$ with norm $\|a\|_{\ell_N^2} = 1$ such that*

$$\text{ess sup}_{t \in [0,1]} \left| \sum_{n=1}^N a_n \phi_n(t) \right| \geq \sqrt{N}.$$

Proof. First we observe that

$$N = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{n=1}^N |\phi_n(t)|^2 dt \leq \text{ess sup}_{t \in [0,1]} \sum_{n=1}^N |\phi_n(t)|^2.$$

So, for every $\epsilon > 0$, there exists $t_0 \in [0, 1]$ such that all $\{\phi_k\}_{k=1}^N$ are defined at t_0 and

$$N - \epsilon \leq \sum_{n=1}^N |\phi_n(t_0)|^2. \tag{2.1}$$

Now set

$$a_n = \frac{\overline{\phi_n(t_0)}}{\sqrt{\sum_{n=1}^N |\phi_n(t_0)|^2}}.$$

Using inequality (2.1), we have

$$\begin{aligned} \sup_{t \in [0,1]} \sum_{n=1}^N a_n \phi_n(t) &\geq \sum_{n=1}^N a_n \phi_n(t_0) = \frac{\sum_{n=1}^N \overline{\phi_n(t_0)} \phi_n(t_0)}{\sqrt{\sum_{n=1}^N |\phi_n(t_0)|^2}} = \left(\sum_{n=1}^N |\phi_n(t_0)|^2 \right)^{1/2} \\ &\geq \sqrt{N - \epsilon}. \end{aligned}$$

Since ϵ is arbitrary, we have proved the proposition. □

Thus, for every orthonormal basis $\{\phi_n\}_{n=1}^\infty$, we have

$$\sup_{\|a\|_2=1} \text{PAPR}(\{\phi_n\}_{n=1}^N, a) \geq \sqrt{N}.$$

In fact, \sqrt{N} is also a bound on the PAPR for both the OFDM and the DS-CDMA systems. Since the transmission signals in each of these cases are uniformly bounded by 1, this follows from applying the Cauchy–Schwarz inequality pointwise to the linear combination. Therefore, OFDM does not offer any advantages with regard to worst-case performance for PAPR. Proposition 2.2 shows that the upper bound on PAPR for these two systems is also a lower bound on PAPR for all orthonormal systems.

2.2. Solvability and a Norm Equivalence

We recall from the introduction that the strategy addressed in this paper is to reserve one subset of orthonormal functions for carrying the information-bearing coefficients and to determine coefficients for the remaining orthonormal functions, so that the combined sum of functions has a small peak value. We formalize this in the following definition.

Definition 2.3. The PAPR reduction problem is *solvable* for the orthonormal system $\{\phi_n\}_{n=1}^\infty$ and the subset $K \subset \mathbb{N}$ with constant C_{ex} if for every $a \in l^2(K)$, there exists $b \in l^2(K^c)$ satisfying $\|b\|_{l^2(K^c)} \leq C_{\text{ex}} \|a\|_{l^2(K)}$ such that

$$\text{ess sup}_{t \in [0,1]} \left| \sum_{n \in K} a_n \phi_n(t) + \sum_{n \in K^c} b_n \phi_n(t) \right| \leq C_{\text{ex}} \|a\|_{l^2(K)}.$$

We may view the map from the coefficient vector a to a function with a small peak as an *extension operator*. This operator is a map from $l^2(K)$ to $L^2([0,1])$ given by

$$E_K a = \sum_{n \in K} a_n \phi_n(t) + \sum_{n \in K^c} b_n \phi_n(t).$$

Note that this map is not necessarily unique and is generally not linear; we shall not discuss the construction of such a map. Nonetheless, since the map gives a correspondence between $l^2(K)$ and $L^2([0,1])$, we refer to it as the extension operator. Thus, we equivalently say that the PAPR reduction problem is solvable for $\{\phi_n\}_{n=1}^\infty$ and K with extension norm C_{ex} if there exists an extension operator E_K such that

$$\|E_K\|_{l^2(K) \rightarrow L^2([0,1])} \leq C_{\text{ex}}.$$

Note that we are interested only in the existence of an extension operator, and that uniqueness is not part of the discussion. Clearly, the operator is generally not linear.

The main results of this paper concern the proportion of signals that may be used for information transmission under a peak-value constraint. Our approach, however, builds on a further point, namely a relationship between PAPR reduction and an L^1 - L^2 -norm equivalence. Given an orthonormal system $\{\phi_n\}_{n=1}^\infty$ for $L^2([0,1])$ and a subset $K \subset \mathbb{N}$, we define

$$X := \left\{ f : f \in L^1([0,1]), f = \sum_{n \in K} a_n \phi_n \right\}.$$

The torus is defined by

$$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}.$$

Theorem 2.4. [Boche and Farrell 10] *Let $\{\phi_k\}_{k \in \mathbb{N}}$ be an orthonormal basis for $L^2([0, 1])$, let K be a subset of \mathbb{N} , and let X be as just defined. The PAPR problem is solvable for the pair K and $\{\phi_k\}_{k \in \mathbb{N}}$ with extension norm C_{ex} if and only if*

$$\|f\|_{L^2(\mathbb{T})} \leq C_{\text{ex}} \|f\|_{L^1(\mathbb{T})} \tag{2.2}$$

for all $f \in X$.

While Theorem 2.4 is proved in [Boche and Farrell 10], we include the proof here so that the role of the Hahn–Banach theorem is apparent. In particular, the proof relies on the existence of a function, denoted by r below, for whose construction there generally exists no method.

Proof. Assume first that the PAPR problem is solvable. Then for all $s(t) = \sum_{k \in K} a_k \phi_k(t)$, $\|a\|_{l^2(\mathbb{Z})} \leq 1$, we have

$$\|E_K a\|_{L^\infty(\mathbb{T})} \leq C_{\text{ex}} \|s\|_{L^2(\mathbb{T})} \leq C_{\text{ex}}.$$

Since $L^\infty(\mathbb{T}) \subset L^2(\mathbb{T})$, it follows that

$$E_K s = \sum_{k \in K} a_k \phi_k + \sum_{k \in \mathbb{N} \setminus K} b_k \phi_k.$$

Let $f \in X$, $f(t) = \sum_{k \in K} c_k \phi_k(t)$, be arbitrary. Then

$$\begin{aligned} \left| \sum_{k \in K} a_k \bar{c}_k \right| &= \left| \sum_{k \in K} a_k \bar{c}_k + \sum_{k \in \mathbb{Z} \setminus K} b_k \bar{c}_k \right| = \left| \frac{1}{2\pi} \int_{\mathbb{T}} f(t) \overline{E_K s(t)} dt \right| \\ &\leq \|f\|_{L^1(\mathbb{T})} \|E_K s\|_{L^\infty(\mathbb{T})} \leq C_{\text{ex}} \|f\|_{L^1(\mathbb{T})}. \end{aligned}$$

Set

$$a_k = \begin{cases} c_k / \|c\|_{l^2}, & c_k \neq 0, \\ 0, & c_k = 0. \end{cases}$$

Then

$$\|f\|_{L^2(\mathbb{T})} = \|c\|_{l^2} = \left| \sum_{k \in K} a_k \bar{c}_k \right| \leq C_{\text{ex}} \|f\|_{L^1(\mathbb{T})}.$$

Assume now that $\|f\|_{L^2(\mathbb{T})} \leq C_{\text{ex}} \|f\|_{L^1(\mathbb{T})}$ for all $f \in X$. Let $a \in l^2(\mathbb{Z})$ be a sequence, supported in K , with finitely many nonzero terms and satisfying

$\|a\|_{l^2(\mathbb{Z})} \leq 1$. Set $s(t) = \sum_{k \in K} a_k \phi_k(t)$. For $f \in X$, $f(t) = \sum_{k \in K} c_k \phi_k(t)$, define the functional Ψ_a by

$$\Psi_a f = \sum_{k \in K} a_k \bar{c}_k.$$

Since

$$|\Psi_a f| \leq \|a\|_{l^2(\mathbb{Z})} \|c\|_{l^2(\mathbb{Z})} \leq \|f\|_{L^2(\mathbb{T})} \leq C_{\text{ex}} \|f\|_{L^1(\mathbb{T})},$$

Ψ_a is continuous on X . Since X is a closed subspace of $L^1(\mathbb{T})$, it follows by the Hahn–Banach theorem that the functional Ψ_a has an extension Ψ_E to all of $L^1(\mathbb{T})$, where $\|\Psi_a\| = \|\Psi_E\|$. The dual of $L^1(\mathbb{T})$ is $L^\infty(\mathbb{T})$. Thus, for some $r \in L^\infty(\mathbb{T})$,

$$\Psi_E f = \langle f, r \rangle,$$

for all $f \in L^1(\mathbb{T})$, so that $\|\Psi_E\| = \|r\|_{L^\infty(\mathbb{T})}$. Since $L^\infty(\mathbb{T}) \subset L^2(\mathbb{T})$, r possesses the unique expansion

$$r(t) = \sum_{k \in \mathbb{N}} d_k \phi_k(t)$$

for some $d \in l^2(\mathbb{Z})$. The sequences d and a agree on K , and we define $E_K s := r$. □

We will also address the case in which we have a finite set of basis functions intended for information and a finite set reserved for peak reduction. We then have a finite set $\{\phi_k\}_{k=1}^N$, which, of course, is then not an orthonormal basis for $L^2([0, 1])$. Consequently, we in general cannot represent the function r in the proof above in terms of $\{\phi_k\}_{k=1}^N$. Nonetheless, we have one direction of Theorem 2.4, which we state as a corollary.

Corollary 2.5. *Let $\{\phi_k\}_{k \in \mathbb{N}}$ be a set of orthonormal functions in $L^2([0, 1])$, let K be a subset of \mathbb{N} , and let X be as previously defined. If the PAPR problem is solvable for the pair K and $\{\phi_k\}_{k \in \mathbb{N}}$ with extension norm C_{ex} , then*

$$\|f\|_{L^2(\mathbb{T})} \leq C_{\text{ex}} \|f\|_{L^1(\mathbb{T})} \tag{2.3}$$

for all $f \in X$.

In the finite setting, if one can show that a constant C_{ex} does not exist such that the norm property in line (2.3) holds for all finite cardinalities, then it follows that solvability cannot hold. Thus, to prove that solvability does not hold, we do not use the two-way statement of Theorem 2.4, but just this one-way statement. This will be our approach later.

3. The Discrete Fourier Case

3.1. Density of Information Sets for OFDM

The discrete Fourier case is interesting for several reasons. First, the discrete case implies the continuous case, and therefore delivers the result on the density of tone reservation sets for OFDM. Additionally, the problem considered is applicable to a large number of areas and is valuable in its own right. The discrete case is important for the PAPR problem because much of the work done with signals is, of course, done with discretized versions of the signals. For example, oversampling and zero-padding are used in the papers [Tellambura 01, Ilic and Strohmer 09]. In some settings it is possible, using sampling results, to relate discrete properties to analog properties, and therefore it is valuable to understand the behavior in the discrete setting. In [Wunder and Boche 03], for example, bounds on the PAPR of an OFDM signal are given in terms of samples of the signal and the oversampling rate.

Definition 3.1. The $N \times N$ inverse discrete Fourier transform (DFT) matrix is given by

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

This matrix is denoted by F , and for $x \in l_N^2$, Fx denotes this matrix applied to x . We denote by F^* the Hermitian transpose or adjoint of F .

Definition 3.2. We denote by l_N^p the set \mathbb{C}^N viewed as a linear space with norm $\|x\|_{l_N^p} = (\sum_{i=1}^N |x_i|^p)^{1/p}$. The unit ball in l_N^p is denoted by B_N^p , i.e.,

$$B_N^p = \left\{ x \in l_N^p : \|x\|_{l_N^p} \leq 1 \right\}.$$

Let $\{N_k\}_{k=1}^\infty$ be a subsequence of \mathbb{N} , and for each N_k , let I_{N_k} be a subset of $\{1, \dots, N_k\}$; $I_{N_k}^c$ denotes $\{1, \dots, N_k\} \setminus I_{N_k}$. In analogy to Definition 2.3, we say that 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_{ex} such that for each k , for all $x \in l_{N_k}^2$ with $\text{supp}(x) \subset I_{N_k}$ there exists a compensation vector $r \in l_{N_k}^2$ supported in $I_{N_k}^c$ such that

$$\|F(x + r)\|_{l_{N_k}^\infty} \leq \frac{C_{\text{ex}}}{\sqrt{N_k}} \|x\|_{l_{N_k}^2}.$$

Theorem 3.3. *Let $\{N_k\}_{k=1}^\infty$ be a subsequence of \mathbb{N} , and let I_{N_k} be a subset of $\{1, \dots, N_k\}$. Let $Y_k = \{y \in l_{N_k}^2 : \text{supp}(F^*y) \subset I_{N_k}\}$. The discrete PAPR problem is solvable for the sequence of sets $\{I_{N_k}\}_{k=1}^\infty$ with constant C_{ex} if and only if*

$$\|y\|_{l_{N_k}^2} \leq \frac{C_{\text{ex}}}{\sqrt{N_k}} \|y\|_{l_{N_k}^1}$$

for all $y \in Y_k$.

In general, one has $\|y\|_{l_{N_k}^2} \leq \|y\|_{l_{N_k}^1}$ for every vector y . Here, though, as k increases, we eventually have $C_{\text{ex}}/\sqrt{N_k} < 1$, and so the important point is that C_{ex} remains fixed.

Proof. Assume first that PAPR is solvable with constant C_{ex} . Let N be an element of $\{N_k\}_{k=1}^\infty$. For $I_N \subset \{1, \dots, N\}$, let $I_N^c = \{1, \dots, N\} \setminus I_N$. Then for every $x \in \mathbb{C}^N$ with $\text{supp}(x) \subset I_N$, we can find an extension $r \in \mathbb{C}^N$ with $\text{supp}(r) \subset I_N^c$ such that

$$F(x+r) \in \frac{C_{\text{ex}}}{\sqrt{N}} B_N^\infty.$$

We denote by $l_N^p(I_N)$ elements of l_N^p with support contained in I_N . Denote by E_{I_N} the operator that maps x to the compensated vector $x+r$. Then

$$\|FE_{I_N} x\|_{l_N^\infty} \leq \frac{C_{\text{ex}}}{\sqrt{N}} \|x\|_{l_N^2}, \quad (3.1)$$

and so $\|FE_{I_N}\|_{l_N^2(I_N) \rightarrow l_N^\infty} \leq C_{\text{ex}}/\sqrt{N}$. As in the analog case, we take a vector b with $\text{supp}(b) \subset I_N$, and observe that

$$|\langle b, E_{I_N} x \rangle| = |\langle Fb, FE_{I_N} x \rangle| \leq \|Fb\|_{l_N^1} \|FE_{I_N} x\|_{l_N^\infty} \leq \|Fb\|_{l_N^1} \frac{C_{\text{ex}}}{\sqrt{N}} \|x\|_{l_N^2}.$$

By setting

$$x_k = \begin{cases} \frac{b_k}{\|b\|_{l_N^2}}, & b_k \neq 0, \\ 0, & b_k = 0, \end{cases}$$

we obtain

$$\|b\|_{l_N^2} = |\langle b, b \rangle| = |\langle b, E_{I_N} x \rangle| \leq \frac{C_{\text{ex}}}{\sqrt{N}} \|Fb\|_{l_N^1}.$$

Now let N be an element of $\{N_k\}_{k=1}^\infty$. We take an element $c \in l_N^2$ with $\text{supp}(c) \subset I_N$. Let Ψ_c be the functional acting on Y by

$$\Psi_c y = \langle c, F^*y \rangle. \quad (3.2)$$

We then have

$$|\Psi_c y| \leq \|c\|_{l_N^2} \|y\|_{l_N^2} \leq \frac{C_{\text{ex}}}{\sqrt{N}} \|y\|_{l_N^1} \|c\|_{l_N^2},$$

so that

$$\|\Psi_c\| \leq \frac{C_{\text{ex}}}{\sqrt{N}} \|c\|_{l_N^2}. \tag{3.3}$$

Since Y is a closed subspace of l_N^1 , by the Hahn–Banach theorem there exists an extension Ψ_E of Ψ_c to all of l_N^1 such that $\|\Psi_c\| = \|\Psi_E\|$, and then Ψ_E can be represented by a vector r such that

$$\Psi_E y = \langle r, y \rangle$$

for all $y \in l_N^1$. Let $\bar{c} = Fr$. If $y \in Y$ and $y = Fx$, then $\langle r, y \rangle = \langle F^* \bar{c}, F^* x \rangle = \langle \bar{c}, x \rangle$. Comparing this with (3.2), we see that c and \bar{c} must agree on I_N . That is, \bar{c} is an extension of c . Lastly, using (3.3), we obtain

$$\|\Psi_E\| = \|r\|_\infty = \|F\bar{c}\|_\infty \leq \frac{C_{\text{ex}}}{\sqrt{N}} \|c\|_{l_N^2}.$$

□

For a set A , $|A|$ denotes its cardinality.

Theorem 3.4. *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*

$$\limsup_{n \rightarrow \infty} \frac{|I_{N_k}|}{N_k} > 0,$$

then the discrete PAPR problem is not solvable.

The proof will use arithmetic progressions and Szemerédi’s theorem, Theorem 3.6.

Definition 3.5. An *arithmetic progression of length m* is a subset of \mathbb{Z} that has the form $\{a, a + d, a + 2d, \dots, a + (m - 1)d\}$ for some integer a and some positive integer d .

Theorem 3.6. [Tao 06, Theorem 1.2] *For every integer $m \geq 1$ and $0 < \delta \leq 1$, there exists an integer $N_{\text{Sz}}(m, \delta) \geq 1$ such that for every $N \geq N_{\text{Sz}}(m, \delta)$, every set $A \subset \{1, \dots, N\}$ of cardinality $|A| \geq \delta N$ contains at least one arithmetic progression of length m .*

Proof of Theorem 3.4. By Theorem 3.6, there exists an integer N in the subsequence $\{N_k\}_{k=1}^\infty$ such that I_N contains an arithmetic progression of length m . Assume again that this progression is $\{a + dl\}_{l=0}^{m-1}$. Let D denote the vector of length N with the value $e^{2\pi i(a+dl)t/N} / \sqrt{m}$ at the entries of the arithmetic progression, where t will be addressed shortly. Then

$$\begin{aligned} \|FD\|_{l_N^1} &= \sum_{j=1}^N |(FD)_j| = \sum_{j=1}^N \left| \sum_{l=1}^N F_{jl} D_l \right| \\ &= \sum_{j=1}^N \left| \frac{1}{\sqrt{N}} \frac{1}{\sqrt{m}} \sum_{l=0}^{m-1} e^{-2\pi i(a+dl)t/N} e^{2\pi i dlj/N} \right| \\ &= \frac{1}{\sqrt{m}} \frac{1}{\sqrt{N}} \sum_{j=1}^N \left| \sum_{l=0}^{m-1} e^{2\pi i dl(t-j)/N} \right|. \end{aligned} \tag{3.4}$$

This calculation holds for every t , so we may take the t that minimizes the absolute value:

$$\begin{aligned} \min_{t \in [0,1]} \sum_{j=1}^N \left| \sum_{l=0}^{m-1} e^{\frac{2\pi i dl(t-j)}{N}} \right| &= \min_{t \in [0,1]} \sum_{j=1}^N \left| \frac{\sin \frac{\pi dm(t-j)}{N}}{\sin \frac{\pi d(t-j)}{N}} \right| \\ &\leq \int_0^1 \sum_{j=1}^N \left| \frac{\sin \frac{\pi dm(t-j)}{N}}{\sin \frac{\pi d(t-j)}{N}} \right| dt = N \int_0^1 \left| \frac{\sin \frac{\pi dmt}{N}}{\sin \frac{\pi dt}{N}} \right| dt \leq N \log m, \end{aligned}$$

where the last step is the bound on the Dirichlet kernel. Now, returning to (3.4) and defining D using the t that results in the minimum in the calculation above, we have $\|FD\|_{l_N^1} \leq \sqrt{N} \log m / \sqrt{m}$. If the discrete PAPR problem is solvable, then by Theorem 3.3, we have a norm equivalence with a factor C_{ex} / \sqrt{N} . However, we have just shown that C_{ex} must be arbitrarily small. This contradiction proves Theorem 3.4. □

Our next result shows that one can have a norm equivalence on subspaces given by subsets of the columns of the DFT matrix when the density converges to zero fast enough.

Corollary 3.7. *Let $\{N_k\}_{k=1}^\infty$ be a subsequence of \mathbb{N} and let I_{N_k} be the corresponding sets as defined earlier. Assume that the compensation set is finite with indices $\{-N, \dots, N\} \setminus I_{N_k}$. If*

$$\limsup_{k \rightarrow \infty} \frac{|I_{N_k}|}{N_k} > 0,$$

then the PAPR problem is not solvable for the sequence of information sets $\{e^{2\pi i l \cdot}\}_{l \in I_{N_k}}$.

Proof. The proof of the previous theorem gives that inequality (3.1) cannot hold for a common constant. Since the discrete case gives the values of the continuous case on the regular N -point grid, it follows that there does not exist a universal constant such that

$$\left\| \sum_{l=1}^N e^{-2\pi il \cdot} x_l + \sum_{l=1}^N e^{-2\pi il \cdot} y_l \right\|_{L^\infty([0,1])} \leq C_{\text{ex}} \|x\|_{l^2(I_N)}, \tag{3.5}$$

where $\text{supp}(x) \subset I_N$ and $\text{supp}(y) \subset R_N$. □

Note that the results of this section also give a bound on the size of subsets for which one does have solvability. Namely, from Theorems 3.3 and 3.4, we see that solvability with constant C_{ex} implies that if m is the length of the longest arithmetic progression in an information set I_N , then $\sqrt{m}/\log m \leq C_{\text{ex}}$. Thus, for a given index set I_N , one can determine its long arithmetic progression and obtain a lower bound on the extension norm.

3.2. Examples of Solvability and Projection Properties

The next theorem gives a condition on a subset of \mathbb{N} such that one has a bounded extension constant. The sacrifice, however, is that the information set has a density converging exponentially fast to 0.

Theorem 3.8. *Let $R_N = \{r_l\}_{l=1}^L$ be a subset of $\{1, \dots, N\}$ satisfying $r_k \geq \lambda r_{k-1}$ for some $\lambda > 1$, and where N is chosen such that $N = \lambda r_L$. There exists a constant $C(\lambda)$ depending only on λ such that for all $a \in l^2_N$ supported on R_N ,*

$$\|a\|_{l^2_N} \leq \frac{C(\lambda)}{\sqrt{N}} \|Fa\|_{l^1_N}.$$

The theorem should be compared with Theorem 4.14, its counterpart for the Walsh setting. In the theorem just stated, λ provides for both a proportion of functions with frequencies higher than r_L and the solvability in the first place. This latter aspect is due to the following result of Paley.

Theorem 3.9. [Wojtaszczyk 91, Section I.B.8] *Let $0 < p < \infty$ and let $\{n_k\}_{k=1}^\infty$ be a subsequence of \mathbb{N} such that $\inf_{k \in \mathbb{N}} n_{k+1}/n_k = \lambda > 1$. Then there exist constants $A(\lambda, p)$ and $B(\lambda, p)$ such that for all sequences with only finitely many nonzero terms,*

$$A(\lambda, p) \left\| \sum_{i=k}^\infty a_i e^{ik \cdot} \right\|_{L^p([0,1])} \leq \left\| \sum_{k=1}^\infty a_k e^{ik \cdot} \right\|_{L^2([0,1])} \leq B(\lambda, p) \left\| \sum_{k=1}^\infty a_k e^{ik \cdot} \right\|_{L^p([0,1])}.$$

Proof of Theorem 3.8. Assume that $a \in l_N^2$ is supported on R_N , and define

$$f(t) = \sum_{l=1}^L a_{r_l} e^{ir_l t}.$$

We define a set of kernels that allow us to represent f in terms of its samples:

$$\mathcal{K}_{N,\lambda} = \left\{ K(t) = \sum_{k=-r_L}^{r_L} e^{ikt} + \sum_{k=-N}^{-r_L-1} d_k e^{ikt} + \sum_{k=r_L+1}^N d_k e^{ikt}, \right. \\ \left. \text{where } d_k = d_{-k} \text{ for } k = r_L + 1, \dots, N \right\}.$$

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

$$f(t) = \frac{1}{N} \sum_{l=0}^N f\left(\frac{2\pi l}{N}\right) K\left(t - \frac{2\pi l}{N}\right).$$

Then, using Theorem 3.9, we obtain

$$\|a\|_{l_N^2} = \|f\|_{L^2([0,1])} \tag{3.6}$$

$$\begin{aligned} &\leq B(\lambda, 1) \|f\|_{L^1([0,1])} = B(\lambda, 1) \left\| \frac{1}{N} \sum_{l=0}^N f\left(\frac{2\pi l}{N}\right) K\left(\cdot - \frac{2\pi l}{N}\right) \right\|_{L^1([0,1])} \\ &\leq \frac{B(\lambda, 1)}{N} \sum_{l=0}^N \left| f\left(\frac{2\pi l}{N}\right) \right| \left\| K\left(\cdot - \frac{2\pi l}{N}\right) \right\|_{L^1([0,1])} \\ &= \frac{B(\lambda, 1)}{N} \sum_{l=0}^N \left| f\left(\frac{2\pi l}{N}\right) \right| \|K\|_{L^1([0,1])} \\ &= B(\lambda, 1) \|K\|_{L^1([0,1])} \frac{1}{N} \sum_{l=0}^N \left| f\left(\frac{2\pi l}{N}\right) \right|, \end{aligned} \tag{3.7}$$

where in (3.7) we use the shift-invariance of the $\|\cdot\|_{L^1(\mathbb{T})}$ -norm for any K .

Setting $d_k = (N - r_L - k/2)$ for $k = r_L, \dots, N$ and

$$K(t) = \frac{1}{N - r_L} \left(\sum_{k=0}^N D_k(t) - \sum_{k=0}^{r_L} D_k(t) \right),$$

we may define $f + g_{\lambda N}$ to be the convolution of $f + g$ with K . The Fourier expansion of $f + g_N$ is supported on $\{-N, \dots, N\}$ and agrees with a on $\{r_k\}_{k=1}^\infty$.

Using P_K to denote the projection given by convolution with K , we have

$$\|f + g_N\|_{L^\infty(\mathbb{T})} \leq \|P_K\| \|f + g\|_{L^\infty(\mathbb{T})} \leq C_{\text{ex}} \|a\|_{l^2(\mathbb{Z})} \|P_K\|. \tag{3.8}$$

The norm $\|P_K\|$ is the $\|\cdot\|_{L^1(\mathbb{T})}$ -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

$$F_n(t) = \frac{1}{n} \sum_{k=0}^{n-1} D_k = \left(\frac{\sin \frac{nt}{2}}{\sin \frac{t}{2}} \right)^2. \tag{3.9}$$

Thus, for every $m > l$,

$$\begin{aligned} \sum_{k=0}^m D_k - \sum_{k=0}^l D_k &= (m-l) \sum_{k=0}^{2l-m} (e^{ikt} + e^{-ikt}) \\ &\quad + \sum_{k=1}^{2(m-l)} \left(m-l - \frac{k}{2} \right) \left(e^{i(2l-m+k)t} + e^{-i(2l-m+k)t} \right). \end{aligned}$$

Using the positivity given in (3.9), we obtain

$$\begin{aligned} \|K\|_{L^1([0,1])} &= \int_0^1 \left| \frac{1}{N-r_L} \left(\sum_{k=0}^N D_k(t) - \sum_{k=0}^{r_L} D_k(t) \right) \right| dt \\ &\leq \frac{1}{N-r_L} \int_0^1 \sum_{k=0}^{\lambda r_L} D_k(t) + \sum_{k=0}^{r_L} D_k(t) dt \\ &\leq \frac{2N}{N-r_L} \leq \frac{2\lambda}{\lambda-1}. \end{aligned}$$

Returning to (3.8), we have

$$\|f + g_N\|_{L^\infty(\mathbb{T})} \leq \frac{2\lambda}{\lambda-1} C_{\text{ex}} \|a\|_{l^2(\mathbb{Z})},$$

where the Fourier expansion of g_N is supported on $\{N, \dots, N\} \setminus I_N$.

Returning to (3.6), we now have

$$\begin{aligned} \|a\|_{l_N^2} &\leq \frac{2\lambda B(\lambda, 1)}{\lambda-1} \frac{1}{N} \sum_{l=0}^N \left| f\left(\frac{2\pi l}{N}\right) \right| \\ &= \frac{2\lambda B(\lambda, 1)}{\lambda-1} \frac{1}{\sqrt{N}} \sum_{l=0}^N \left| \frac{1}{\sqrt{N}} \sum_{k=1}^L a_{r_k} e^{2\pi i r_k l/N} \right| \\ &= \frac{2\lambda B(\lambda, 1)}{\lambda-1} \frac{1}{\sqrt{N}} \|Fa\|_{l_N^1}. \end{aligned}$$

□

4. The Walsh or CDMA Case

4.1. PAPR and Density for Walsh or CDMA Systems

PAPR has been studied extensively in OFDM systems. In this section, we show that the same type of behavior occurs as well in the downlink of direct sequence–code division multiple access (DS-CDMA) systems. We assume again, without loss of generality, that the symbol period is normalized to length 1. Assume that a base station communicates with $N = 2^n$ users. DS-CDMA then uses 2^n orthogonal functions that take the value 1 or -1 on intervals of length 2^{-n} . These functions are the spreading sequences. We denote these sequences by $\{w_k\}_{k=1}^{2^n}$, and we will give their specific values shortly. The base station then transmits

$$s(t) = \sum_{k=1}^{2^n} d_k w_k(t) \quad \text{for } t \in [0, 1] \quad (4.1)$$

to transmit the coefficient d_k to user k .

However, Proposition 2.2 applies here, and so the function in (4.1) can achieve peaks of size $\sqrt{2^n}$. We note that in the uplink, each user transmits only one signal w_k , and so there is not the accumulation of signals that leads to the high peaks that can occur in the downlink. Thus, in order to reduce the PAPR of DS-CDMA downlink signals, one could reserve certain spreading sequences to be used for compensation, in analogy to tone reservation for OFDM.

There are several ways to define the Walsh system, though the definitions involve only a different ordering. We present one definition now, and will comment on another in Section 4.3. The various definitions may be found in the first several pages of [Schipp et al. 90]. The following is the original ordering given in [Walsh 23], and is the system used in the downlink for the IS-95 standard and UMTS-IMT-2000.

Definition 4.1. The *Rademacher functions*, denoted by r_0, r_1, \dots , are defined on $[0, 1]$ by

$$r_k(t) = \operatorname{sgn} \sin(2\pi 2^k t),$$

where we define $\operatorname{sgn} 0 = -1$. The *Walsh functions*, denoted by w_1, w_2, \dots , are defined using the Rademacher functions by

$$w_1(t) = 1 \quad \text{and} \quad w_{2^k + m}(t) = r_k(t) \cdot w_m(t)$$

for $k = 0, 1, 2, \dots$ and $m = 1, \dots, 2^k$.

See [Fine 49] for the fundamental properties of the Walsh functions.

We are able to obtain specific bounds on C_{ex} in the Walsh case and therefore can state more precise results than in the discrete Fourier case. Our main result concerning the Walsh functions is the following.

Theorem 4.2. *Let $\delta \in (0, 1)$ be a density and assume that $N = 2^n$ ($n \in \mathbb{N}$) satisfies $N \geq (\frac{2}{\delta})^{m+1}$ for some $m \in \mathbb{N}$. If the PAPR problem is solvable with constant C_{ex} for a subset of indices $Y \subset \{1, \dots, N\}$ for $|Y|/N \geq \delta$, then*

$$C_{\text{ex}} \geq \frac{2^m - m^2}{1 + m}.$$

We need several lemmas before we can prove the theorem. We begin with a definition.

Definition 4.3. The correlation function between w_r and I is

$$C(w_r, I) = \int_0^1 w_r(x) \left| \sum_{k \in I} w_k(x) \right|^2 dx.$$

Further, for w_r and I , we define the set

$$\mathcal{M}(w_r, I) = \{k \in I : \text{there exists } \tilde{k} \in I \text{ s.t. } w_k(x)w_{\tilde{k}}(x) = w_r(x) \forall x \in [0, 1]\}.$$

We could equivalently look at the set of all pairs (k, \tilde{k}) such that $w_k w_{\tilde{k}} = w_r$, which would always include the permutation (\tilde{k}, k) . For every $k \in \mathcal{M}(w_r, I)$, there is exactly one \tilde{k} in $\mathcal{M}(w_r, I)$ such that the pair k and \tilde{k} satisfy the requirement given for $\mathcal{M}(w_r, I)$. To see this, suppose that \tilde{k}_1 and \tilde{k}_2 both satisfy

$$w_k w_{\tilde{k}_1} = w_r = w_k w_{\tilde{k}_2}. \tag{4.2}$$

Then since $w_k w_k = 1$ for every k , multiplying (4.2) by w_k , we have

$$w_{\tilde{k}_1} = w_{\tilde{k}_2}. \tag{4.3}$$

Remark 4.4. The reader unfamiliar with the Walsh functions may find it helpful to read the proof of Theorem 4.11. The properties used in the proof below are contained in the latter proof. In particular, the central idea of the proof of the main lemma, Lemma 4.9, relies on the identity in (4.8). This equality gives a representation of the sum of Walsh functions as a product of factors $(1 + r_k)$ for the appropriate Rademacher functions. This property and the fact that multiplying the set of Walsh functions (other than w_1) by a Walsh function gives a permutation of the Walsh functions leads to the idea of expressing a

linear combination of Walsh functions as a product. If this is done correctly, one can obtain the L^1 and L^2 properties necessary for the theorem.

Lemma 4.5. *The following equality holds:*

$$C(w_r, I) = |\mathcal{M}(w_r, I)|.$$

Proof. We have

$$\begin{aligned} C(w_r, I) &= \int_0^1 w_r(x) \left(\sum_{k_1 \in I} w_{k_1}(x) \right) \left(\sum_{k_2 \in I} w_{k_2}(x) \right) dx \\ &= \sum_{k_1 \in I} \sum_{k_2 \in I} \int_0^1 w_r(x) w_{k_1}(x) w_{k_2}(x) dx. \end{aligned} \quad (4.4)$$

For $1 \leq r, k_1, k_2 \leq N$, we have $w_r w_{k_1} w_{k_2} = w_{\tilde{r}}$ for some $1 \leq \tilde{r} \leq N$. Therefore, the integral in (4.4) is nonzero only when $w_r w_{k_1} w_{k_2} = w_1 = 1$, that is, when $w_r = w_{k_1} w_{k_2}$. This is the set of all pairs of k_1 and k_2 such that this holds. In light of the comment following the definition of $\mathcal{M}(w_r, I)$, the lemma is proved. \square

An important monotonicity property follows from this lemma.

Corollary 4.6. *Assume that $C(r, I) = 0$. Then $C(r, \tilde{I}) = 0$ for all $\tilde{I} \subset I$.*

Proof. This follows from

$$0 = C(w_r, I) = |\mathcal{M}(w_r, I)| \geq |\mathcal{M}(w_r, \tilde{I})| = C(w_r, \tilde{I}).$$

\square

Lemma 4.7. *If $N = 2^n$ for a positive integer n , then*

$$\sum_{r=1}^N C(w_r, I) = |I|^2.$$

Proof. Recall that for $x \in [0, \frac{1}{N})$, $w_r(x) = 1$ for all r . And since $N = 2^n$, it follows that $\sum_{r=1}^N w_r(x) = 0$ for all $x \in [\frac{1}{N}, 1]$. See, for example, [Fine 49]. Therefore,

$$\begin{aligned} \sum_{r=1}^N C(w_r, I) &= \int_0^1 \sum_{r=1}^N w_r(x) \left| \sum_{k \in I} w_k(x) \right|^2 dx = N \int_0^{1/N} \left| \sum_{k \in I} w_k(x) \right|^2 dx \\ &= N \int_0^{1/N} |I|^2 dx = |I|^2, \end{aligned}$$

which completes the proof. \square

Definition 4.8. If f is a linear combination of Walsh functions, we define $\text{supp}(f)$ to be the set of indices corresponding to the functions in the sum.

Lemma 4.9. Let δ be a density, that is, $\delta \in (0, 1)$. For $m > 1$, if $N = 2^n$, $n \in \mathbb{N}$, satisfies $N \geq (2/\delta)^{m+1}$, then for every subset $I \subset \{1, \dots, N\}$ satisfying $|I|/N \geq \delta$, there exists a function f supported on I such that

$$\|f\|_{L^2([0,1])} \geq 2^m - m^2 \quad \text{and} \quad \|f\|_{L^1([0,1])} \leq 1 + m.$$

Proof. First we just take N to be large and I to be a subset of $\{1, \dots, N\}$. We will construct a function with the norm properties given, and then we will show that m can be made large in dependence on δ and N .

Suppose that $\mathcal{M}(w_r, I) = \{l_1, \dots, l_{2k}\}$. Then

$$w_r \cdot \{w_{l_1}, \dots, w_{l_{2k}}\} = \{w_{l_1}, \dots, w_{l_{2k}}\}.$$

Also, suppose that $w_j w_{l_{j+1}} = w_r$ for $j = 1, \dots, k$. Then $w_r w_j = w_{l_{j+1}}$ for each j . Therefore, we may select the subset $\{w_{l_1}, w_{l_3}, \dots, w_{l_{2k-1}}\}$ and obtain the properties

$$w_r \cdot \{w_{l_1}, w_{l_3}, \dots, w_{l_{2k-1}}\} = \{w_{l_2}, w_{l_4}, \dots, w_{l_{2k}}\}$$

and

$$w_r \cdot \{w_{l_1}, w_{l_3}, \dots, w_{l_{2k-1}}\} \cap \{w_{l_2}, w_{l_4}, \dots, w_{l_{2k}}\} = \emptyset.$$

We use this splitting to select a set that does not allow for any factoring of $w^{(1)}$. We use the notation $I_A^{(1)}$ and $I_B^{(1)}$ to denote a splitting of $I^{(1)}$ in the way just described. Since $C(w^{(1)}, I^{(1)}) = |\mathcal{M}(w^{(1)}, I^{(1)})|$ and $w^{(1)}$ cannot be realized as the product of any two Walsh functions with indices in $I^{(1)}$, we have $C(w^{(1)}, I^{(1)}) = 0$. Lastly, we also have $|I^{(1)}| = \frac{1}{2} |\mathcal{M}(w_r, I)|$.

Let $w^{(1)}$ be the Walsh function with the highest correlation with I , that is,

$$C(w^{(1)}, I) = \max_{w=w_2, \dots, w_N} C(w, I).$$

Since $C(w, I) = |\mathcal{M}(w, I)|$, the maximizer $w^{(1)}$ is the Walsh function whose index corresponds to the subset of I with the greatest number of splittings contained in I .

Now set $I^{(2)} = \mathcal{M}(w^{(1)}, I^{(1)})$, and define $I_A^{(2)}$ and $I_B^{(2)}$ accordingly. We repeat this until ultimately, $I^{(m)} = \{i_{m,1}, i_{m,2}\}$ for two indices such that $w_{i_{m,1}} w_{i_{m,2}} = w^{(m)}$. We have then selected Walsh functions $w^{(1)}, w^{(2)}, \dots, w^{(m)}$ and determined

sets $I_A^{(1)}, I_B^{(1)}, \dots, I_A^{(m)}, I_B^{(m)}$ such that for $j = 1, \dots, m$,

$$\begin{aligned} w^{(j)} I_A^{(j)} &= I_B^{(j)}, & I_A^{(j)} &= w^{(j)} I_B^{(j)}, \\ w^{(j)} I_A^{(j)} \cap I_B^{(j)} &= \emptyset, & I_A^{(j)} \cup I_B^{(j)} &= I^{(j)}, \end{aligned} \quad (4.5)$$

Define

$$\mathcal{F}_0 = \left\{ f : f(x) = \sum_{k \in I} \alpha_k w_k(x), \alpha_k \in \{0, 1\}, \text{ at least one } \alpha_k \neq 0 \right\},$$

and for $l = 1, \dots, m$, define

$$\mathcal{F}_l = \left\{ f : f(x) = \sum_{k \in I^{(l)}} \alpha_k w_k(x), \alpha_k \in \{0, 1\}, \text{ at least one } \alpha_k \neq 0 \right\}.$$

Now we construct our desired function. We have $w^{(m)} = w_{i_{m,1}} w_{i_{m,2}}$, so that $w^{(m)} w_{i_{m,2}} = w_{i_{m,1}}$. Then

$$(1 + w_{i_{m,2}}) w^{(m)} = w_{i_{m,1}} + w_{i_{m,2}} \in I^{(m)} \subset I_A^{(m-1)}. \quad (4.6)$$

The two Walsh functions in the sum (4.6) are unique. Thus

$$(1 + w_{i_{m,2}}) w^{(m)} \in \mathcal{F}^{(m)},$$

and we set

$$f^{(m)} = (1 + w_{i_{m,2}}) w^{(m)}.$$

Now we repeat this by looking at

$$\left(1 + w^{(m-1)}\right) f^{(m)}.$$

By the properties given in (4.5) and (4.6), we have $\text{supp}(w^{(m-1)} f^{(m)}) \subset I_B^{(m-1)}$, while $\text{supp}(f^{(m)}) \subset I_A^{(m-1)}$. Thus $(1 + w^{(m-1)}) f^{(m)}$ has four unique terms and is contained in $\mathcal{F}^{(m-1)}$. We denote it by $f^{(m-1)}$. Similarly, $(1 + w^{(m-2)}) f^{(m-1)}$ has eight unique terms and is contained in $\mathcal{F}^{(m-2)}$. We continue this and ultimately arrive at $f^{(0)} \in \mathcal{F}^{(0)}$; $f^{(0)}$ is the sum of 2^m unique Walsh functions, and in particular,

$$1 + f^{(0)} = \prod_{l=1}^m (1 + w^{(l)}) - \sum_{l=1}^m w^{(l)}.$$

This yields the bounds

$$\begin{aligned} \int_0^1 |1 + f^{(0)}(x)|^2 dx &= \int_0^1 \left| \prod_{l=1}^m (1 + w^{(l)}(x)) - \sum_{l=1}^m w^{(l)} \right|^2 dx \\ &\geq \int_0^1 \left| \prod_{l=1}^m (1 + w^{(l)}(x)) \right|^2 dx - m^2 = \int_0^1 \prod_{l=1}^m |1 + w^{(l)}(x)|^2 dx - m^2 \\ &= \int_0^1 \prod_{l=1}^m 2(1 + w^{(l)}(x)) dx - m^2 = 2^m - m^2 \end{aligned}$$

and

$$\begin{aligned} \int_0^1 |1 + f^{(0)}(x)| dx &= \int_0^1 \left| \prod_{l=1}^m (1 + w^{(l)}(x)) - \sum_{l=1}^m w^{(l)} \right| dx \\ &\leq \int_0^1 \left| \prod_{l=1}^m (1 + w^{(l)}(x)) \right| dx + m = \int_0^1 \prod_{l=1}^m (1 + w^{(l)}(x)) dx + m = 1 + m. \end{aligned}$$

Now it remains to obtain a lower bound on m . We have

$$|I^{(1)}| = \frac{1}{2} \left| \mathcal{M}(w^{(1)}, I) \right|.$$

Using Lemma 4.7, we obtain

$$|I^{(1)}| = \frac{1}{2} \left| \mathcal{M}(w^{(1)}, I) \right| = \frac{1}{2} \max_{w_r \neq w_1} C(w_r, I) \geq \frac{1}{2} \frac{1}{N} \sum_{r=1}^N C(w_r, I) = \frac{1}{2} \frac{|I|^2}{N} \geq \frac{\delta}{2} |I|.$$

Similarly,

$$|I^{(k)}| \geq \frac{\delta}{2} |I^{(k-1)}| \quad \text{when } |I^{(k-1)}| \geq 4.$$

If the process goes until m , then $|I^{(m)}| \geq (\frac{\delta}{2})^m \delta N$. If $N \geq (\frac{2}{\delta})^{m+1}$, then $(\frac{\delta}{2})^m \delta N \geq 2$, and thus the set splitting can be performed m times. \square

Proof of Theorem 4.2. By Corollary 2.5, solvability implies that for all f with support in Y , we have

$$\|f\|_{L^2(\mathbb{T})} \leq C_{\text{ex}} \|f\|_{L^1(\mathbb{T})}.$$

By Lemma 4.9, we have

$$C_{\text{ex}} \geq \frac{2^m - m^2}{1 + m}$$

and the proof is complete. \square

We return to the statement of Theorem 4.2. Suppose that a certain δ is fixed. Then as N increases, the largest value for m that still satisfies $N \geq (\frac{2}{\delta})^{m+1}$

increases. Thus the lower bound on C_{ex} increases as N increases. That is, given a density δ and a number of Walsh functions N , one knows a lower bound for the best possible extension norm. In particular, for a fixed δ , there does not exist a uniform extension constant.

As stated in Remark 4.4, the two fundamental properties here are the permutation of the Walsh functions when multiplied by another Walsh function and the representation of the sum as a product. Indeed, we only needed to find m Walsh functions such that their products are all unique and supported in I . The m functions also need not have indices in I . The product terms then disappear for the L^1 -norm but not for the L^2 -norm. The first steps of this approach also work for the Fourier case, since the properties just listed hold there as well. One can also define a correlation function and look for splittings of maximum cardinality. The approach encounters a difficulty, though, when we try to obtain an upper L^1 bound on the product that was easily bounded in the Walsh case.

4.2. Solvability of PAPR for Direct Sequence CDMA and Localized Behavior of the Walsh System

We have several positive results for cases in which the density of the information set converges to zero. Here, the Fourier and Walsh systems have a property in common. In particular, similar to Theorem 3.8 in the Fourier case, we show that if the information set is the dyadic integers, then any information-bearing signal can be compensated by a signal supported on the remaining indices. This is Theorem 4.13. Therefore, the Fourier and Walsh systems behave similarly as far as density and solvability are concerned, and they differ in terms of their projection properties. En route to the last results, we require several definitions. We finish the section with Theorem 4.19, which gives a matrix embedding for Hadamard matrices.

Theorem 4.10. (Khintchine's inequality.) [Wojtaszczyk 91, I.B.8] *There exist constants A_p and B_p , $0 < p < \infty$, such that for all finite sequences of scalars $(a_i)_{i=1}^n$,*

$$A_p \left\| \sum_{i=1}^n a_i r_i \right\|_{L^p([0,1])} \leq \|a\|_{l^2(\mathbb{N})} = \left\| \sum_{i=1}^n a_i r_i \right\|_{L^2([0,1])} \leq B_p \left\| \sum_{i=1}^n a_i r_i \right\|_{L^p([0,1])} .$$

The Walsh system has a very special property: projection operators mapping bounded functions onto the span of the first 2^k Walsh functions are uniformly bounded with norm 1. Thus, when we combine this projection property with Khintchine's inequality, we obtain a statement for finite sets. While the following material may be found in [Fine 49], we include the proof of Theorem 4.11 for the reader's convenience and to emphasize the unique properties of the Walsh

system. Theorems 4.13 to 4.17 are easy to understand once one has read through the proof below. Equation (4.10) of the proof shows that the dyadic Lebesgue constants are 1 for the Walsh system, which is a fundamentally different behavior from the $\log N$ behavior that occurs in the Fourier case.

Theorem 4.11. *Let P_{2^n} denote the projection onto $\{w_1, \dots, w_{2^n}\}$. Then*

$$\|P_{2^n}\|_{C([0,1]) \rightarrow L^\infty([0,1])} = 1$$

for all $n \in \mathbb{N}$, and if $f \in C([0, 1])$, then for all $x \in [0, 1]$,

$$\lim_{n \rightarrow \infty} (P_{2^n} f)(x) = f(x).$$

Proof. The projection of $f \in L^2([0, 1])$ onto w_n is

$$c_n = \int_0^1 w_n(x) f(x) dx.$$

We consider the projection onto $\{w_k\}_{k=1}^n$ at the point x and denote this by $s_n(x, f)$:

$$s_n(x, f) = \sum_{k=1}^n c_k w_k(x).$$

Equivalently,

$$s_n(x, f) = \int_0^1 \sum_{k=1}^n f(t) w_k(t) w_k(x) dt.$$

Then

$$s_{2^n}(x, f) = \int_0^1 \sum_{k=1}^{2^n} f(t) w_k(t) w_k(x) dt. \tag{4.7}$$

We are interested in the sum

$$\sum_{k=1}^{2^n} w_k(t) w_k(x);$$

however, this is just the sum of all possible products of $\{r_0(x)r_0(t), \dots, r_n(x)r_n(t)\}$, and so

$$\sum_{k=1}^{2^n} w_k(t) w_k(x) = \prod_{k=1}^n (1 + r_k(x)r_k(t)). \tag{4.8}$$

If x and t are in the same dyadic interval of length 2^{-n} , then $r_k(x)r_k(t) = 1$ for $k = 1, \dots, n$. But if there exists k less than or equal n such that x and t are not

in the same dyadic interval of length 2^{-k} , then $r_k(x)r_k(t) = -1$. This is due to the fact that one term must equal 1 and the other must equal -1 . In this case, the product (4.8) must equal zero. Defining

$$D_{2^n}(x, t) = \sum_{k=1}^{2^n} w_k(t)w_k(x),$$

we then have

$$D_{2^n}(x, t) = \begin{cases} 2^n, & x, t \text{ in the same dyadic interval of length } 2^{-n}, \\ 0, & x, t \text{ not in the same dyadic interval of length } 2^{-n}. \end{cases} \quad (4.9)$$

Then

$$(P_{2^n} f)(x) = \int_0^1 f(t)D_{2^n}(x, t)dt.$$

If $x \in I_m := [\frac{m-1}{2^n}, \frac{m}{2^n}]$, then using (4.9), we obtain

$$\|P_{2^n}\|_{\mathcal{C}([0,1]) \rightarrow \mathcal{C}([0,1])} = \sup_{\substack{f \in \mathcal{C}([0,1]), \\ \|f\|_{L^\infty([0,1])} = 1}} \int_0^1 f(t)D_{2^n}(x, t)dt = \int_{I_m} D_{2^n}(x, t) dt = 1. \quad (4.10)$$

This proves the first claim.

For $x \in [0, 1]$ define $\alpha_n(x)$ and $\beta_n(x)$ by taking them to satisfy the following statement for an appropriate integer m :

$$\alpha_n(x) = m2^{-n} \leq x < (m + 1)2^{-n} = \beta_n(x).$$

Now returning to (4.7), we have

$$s_{2^n}(x, f) = \int_0^1 f(t)D_{2^n}(x, t) dt = 2^n \int_{\alpha_n(x)}^{\beta_n(x)} f(t) dt = \frac{F(\beta_n(x)) - F(\alpha_n(x))}{\beta_n(x) - \alpha_n(x)},$$

where $F(x)$ is an integral of $f(x)$. Since f is assumed to be continuous, we have

$$\lim_{n \rightarrow \infty} (P_{2^n} f)(x) = f(x),$$

and we are done. □

We include here the following theorem, which is proved in [Boche and Farrell 10], in order to contrast the projection behavior of the Fourier and Walsh bases. The additional redundancy of a factor of λ frequencies in the compensation set is necessary to obtain the theorem below. For a given extension norm, as the size of the information-bearing set increases, not only does the compensation set have to increase proportionally, but the set of extra frequencies included beyond the highest frequency must also grow proportionally. We will return to

this when we discuss the behavior of the Walsh system, where one can project sharply.

Theorem 4.12. [Boche and Farrell 10] *Suppose that I_N is a subset of $\{-N, \dots, N\}$ and that for every $a \in l^2(\mathbb{Z})$ supported on I_N , the PAPR reduction problem is solvable with an extension sequence supported on $\mathbb{Z} \setminus I_N$ and with extension bound C_{ex} . Assume $\lambda > 1$ and that λN is an integer. Then the PAPR reduction problem is also solvable with an extension sequence supported on $\{-\lambda N, \dots, \lambda N\} \setminus I_N$ with extension constant $\frac{2\lambda}{\lambda-1}C_{\text{ex}}$.*

The next theorem addresses the same question for the Walsh system. The significant point here is that with Walsh functions, one may work with only the dyadic set in which the information-bearing coefficients are contained. There is nothing gained or lost by using or not using any Walsh functions beyond this dyadic set.

Theorem 4.13. *Suppose that I_N is a subset of $\{1, 2, 3, \dots, N\}$ and that for every function f of the form*

$$f(t) = \sum_{k \in I_{2^n}} a_k w_k(t),$$

there exists a compensation function with coefficient vector a supported on $\mathbb{N} \setminus I_N$ such that the combined signal has $\|\cdot\|_{L^\infty(\mathbb{T})}$ -norm at most $C_{\text{ex}}\|a\|_{l^2(\mathbb{N})}$. Let n be the smallest integer such that $N \leq 2^n$. Then there exists a compensation function with coefficients supported on $\{k\}_{k=1}^{2^n} \setminus I_N$ such that the $\|\cdot\|_{L^\infty(\mathbb{T})}$ -norm of the combined signal is still at most $C_{\text{ex}}\|a\|_{l^2(\mathbb{N})}$.

Proof. By Theorem 4.11, we may simply project the original combined function onto the span of $\{w_k\}_{k=1}^{2^n}$ and maintain the same norm. □

Now we may consider a special case of Theorem 4.13, for which we know that the PAPR reduction problem is solvable, namely when I_N is the set of dyadic integers.

Theorem 4.14. *Let B_1 be the constant given in Khintchine’s inequality (Theorem 4.10). Then for every function of the form*

$$\sum_{k=1}^n a_{2^k} w_{2^k}(t),$$

there exists a vector $b \in l_{2^n}^2$ supported on $\{k\}_{k=1}^{2^n} \setminus \{2^k\}_{k=1}^n$ and with norm $\|b\|_{l_{2^n}^2} \leq B_1 \|a\|_{l_{2^n}^2}$ such that

$$\left\| \sum_{k=1}^n a_{2^k} w_{2^k} + \sum_{\{k\}_{k=1}^{2^n} \setminus \{2^k\}_{k=1}^n} b_k w_k \right\|_{L^\infty([0,1])} \leq B_1 \|a\|_{l_{2^n}^2}.$$

Proof. We may take the subset K in Theorem 2.4 to be $\{2^k\}_{k=1}^n$. (Theorem 2.4 of course holds when $L^p(\mathbb{T})$ is replaced by $L^p([0, 1])$.) Khintchine’s inequality (4.10) gives the norm equivalence, and thus there exists a sequence $b \in \mathbb{N} \setminus \{2^k\}_{k=1}^n$ such that

$$\left\| \sum_{k=1}^n a_{2^k} w_{2^k} + \sum_{\mathbb{N} \setminus \{2^k\}_{k=1}^n} b_k w_k \right\|_{L^\infty([0,1])} \leq B_1 \|a\|_{l_{2^n}^2}. \tag{4.11}$$

Applying Theorem 4.13 to the function in (4.11) proves the theorem. □

We define an optimal subset size for the Walsh system.

Definition 4.15. (Optimal subset size–Walsh.) We define

$$\mathcal{E}_N(C_{\text{ex}}, W) = \max\{|I_N|, I_N \subset \{1, \dots, N\}, \text{ such that } \text{PAPR is solvable for } I_N \text{ with constant } C_{\text{ex}}\},$$

where W refers to the Walsh system.

The following result shows that for a given extension constant, the efficiency of the optimal subset does not increase as the dimension increases. This means that we have strict monotone convergence on dyadic subsets of \mathbb{N} . The result is stronger than an asymptotic statement: it holds for all dyadic Walsh subsets.

Theorem 4.16. *For the Walsh system, the optimal subsets satisfy the following inequality:*

$$2\mathcal{E}_{2^m}(C_{\text{ex}}, W) \geq \mathcal{E}_{2^{m+1}}(C_{\text{ex}}, W)$$

for all constants C_{ex} .

Proof. We will call a subset of $\{0, \dots, 2^m\}$ an optimal subset for the constant C_{ex} if the PAPR reduction problem is solvable for the subset with constant C_{ex} and there is no other subset of greater cardinality for which this holds. Let $I_{2^{m+1}}^*$ denote an optimal subset of $\{0, \dots, 2^{m+1}\}$ and $I_{2^m}^*$ an optimal subset of

$\{0, \dots, 2^m\}$. First define

$$I^{(1)} = \{0, 1, \dots, 2^m\} \cap I_{2^{m+1}}^*$$

For $c \in l^2(I^{(1)})$, let $f_{c,2^{m+1}}$ denote the extension function in $\text{span}\{w_0, \dots, w_{2^{m+1}}\}$ that satisfies

$$\|f_{c,2^{m+1}}\|_{L^\infty([0,1])} \leq C_{\text{ex}} \|c\|_{l^2(I^{(1)})}.$$

By Theorem 4.11, we also have

$$\|P_{2^m} f_{c,2^{m+1}}\|_{L^\infty([0,1])} \leq \|f_{c,2^{m+1}}\|_{L^\infty([0,1])} \leq C_{\text{ex}} \|c\|_{l^2(I^{(1)})}.$$

So, if we define $f_{c,2^m} = P_{2^m} f_{c,2^{m+1}}$, then $f_{c,2^m}$ is also a solution for $c \in l^2(I^{(1)})$ with constant C_{ex} . Now, since we assumed that $I_{2^m}^*$ is an optimal subset, we must have $|I^{(1)}| \leq |I_{2^m}^*|$.

Now we define $I^{(2)}$ by

$$I^{(2)} = \{2^m, \dots, 2^{m+1}\} \cap I_{2^{m+1}}^*$$

We define $P^{(2)}$ and $Q^{(2)}$ by

$$\begin{aligned} (P^{(2)} f)(t) &= (P_{2^{m+1}} f - P_{2^m})(t) = \sum_{k=2^m}^{2^{m+1}} c_k(f) w_k(t) = w_{2^m} \sum_{k=2^m}^{2^{m+1}} c_k(f) w_{k-2^m}(t) \\ &= w_{2^m}(t) (Q^{(2)} f)(t). \end{aligned}$$

Note that $Q^{(2)}$ maps into $\text{span}\{w_0, \dots, w_{2^m}\}$. Since $|w_{2^m}(t)| = 1$, we have

$$\|P^{(2)} f\|_{L^\infty([0,1])} = \|Q^{(2)} f\|_{L^\infty([0,1])}.$$

We can now do the same calculation for the Lebesgue constants for $Q^{(2)}$ that was done in Theorem 4.11 and determine that $\|Q^{(2)}\| = 1$. Thus,

$$\|P^{(2)} f\|_{L^\infty([0,1])} = \|Q^{(2)} f\|_{L^\infty([0,1])} \leq \|f\|_{L^\infty([0,1])}.$$

Let $c \in l^2(I^{(2)})$, and suppose that $f_{c,2^{m+1}}^{(2)}$ is the extension function in

$$\text{span}\{w_0, \dots, w_{2^{m+1}}\}$$

that satisfies

$$\|f_{c,2^{m+1}}^{(2)}\|_{L^\infty([0,1])} \leq C_{\text{ex}} \|c\|_{l^2(I^{(2)})}.$$

Then

$$\|Q^{(2)} f_{c,2^{m+1}}^{(2)}\|_{L^\infty([0,1])} \leq \|f_{c,2^{m+1}}^{(2)}\|_{L^\infty([0,1])} \leq C_{\text{ex}} \|c\|_{l^2(I^{(2)})}.$$

Setting

$$\underline{I}^{(2)} = \left\{ j \in \{0, \dots, 2^m\} : j + 2^m \in I^{(2)} \right\},$$

since $Q^{(2)}$ maps onto $\text{span}\{0, \dots, 2^m\}$, we have established solvability with the same constant on $\underline{I}^{(2)}$, so that $|\underline{I}^{(2)}| \leq |I_{2^m}^*|$. Finally,

$$|I_{2^{m+1}}^*| = |I^{(1)} \cup I^{(2)}| = |I^{(1)}| + |I^{(2)}| = |I^{(1)}| + |\underline{I}^{(2)}| \leq 2|I_{2^m}^*|,$$

which completes the proof. \square

The following proposition is a simple consequence of the projection property for the Walsh system. It shows that if the information coefficients are all supported on a subset of $\{1, \dots, 2^m\}$, then there is nothing to be gained by including Walsh functions with index higher than 2^m for the compensation. Thus, compensation in the Walsh system is a local problem for dyadic sets of integers. This is in strong contrast to the Fourier system, where, as can be seen in Theorem 4.12, it is necessary to include tones at a certain redundancy factor beyond the largest element of the subset.

We introduce two new terms. For $c \in l^2(I_{2^m}^*)$ and $r \geq m$, we define

$$\text{Interp}(c, r, I_{2^m}^*) = \{f \in \text{span}\{w_0, \dots, w_{2^r}\} : c_k(f) = c_k \forall k \in I_{2^m}^*\}$$

and

$$\underline{C}_{\text{ex}}(r, I_{2^m}^*) = \sup_{\|c\|_{l^2(I_{2^m}^*)}} \left(\inf_{f \in \text{Interp}(c, r, I_{2^m}^*)} \|f\|_{L^\infty([0,1])} \right).$$

Proposition 4.17. *For all $r \geq m$,*

$$\underline{C}_{\text{ex}}(r, I_{2^m}^*) = \underline{C}_{\text{ex}}(m, I_{2^m}^*).$$

Proof. Suppose that

$$\underline{C}_{\text{ex}}(r, I_{2^m}^*) < \underline{C}_{\text{ex}}(m, I_{2^m}^*). \quad (4.12)$$

Then for every $\epsilon > 0$,

$$\inf_{f \in \text{Interp}(c, r, I_{2^m}^*)} \|f\|_{L^\infty([0,1])} < (\underline{C}_{\text{ex}}(r, I_{2^m}^*) + \epsilon) \|c\|_{l^2(I_{2^m}^*)}$$

and

$$\begin{aligned} \inf_{f \in \text{Interp}(c, r, I_{2^m}^*)} \|P_{2^m} f\|_{L^\infty([0,1])} &\leq \inf_{f \in \text{Interp}(c, r, I_{2^m}^*)} \|f\|_{L^\infty([0,1])} \\ &< (\underline{C}_{\text{ex}}(r, I_{2^m}^*) + \epsilon) \|c\|_{l^2(I_{2^m}^*)}, \end{aligned}$$

which contradicts the assumption (4.12). \square

4.3. Matrix Results Related to the Walsh System

We give a result for matrices that is analogous to the DFT matrix for the Walsh system. As we have seen throughout this section and the last, the Walsh system differs from the Fourier system in that results for the Walsh system hold for the dyadic sets of integers and do not require the redundancy needed in the Fourier case. This is seen again by comparing Proposition 4.19 here with Proposition 3.8 in the Fourier case.

Definition 4.18. The Rademacher matrix $R_k : \mathbb{C}^k \rightarrow \mathbb{C}^{2^k}$ is defined by

$$(R_k)_{i,j} = \frac{1}{\sqrt{2^k}} r_j \left(\frac{i}{2^k} \right).$$

The *Hadamard matrices* are defined inductively as follows:

$$H_1 = [1], \quad H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

and

$$H_{2^{k+1}} = \frac{1}{\sqrt{2^{k+1}}} \begin{bmatrix} H_{2^k} & H_{2^k} \\ H_{2^k} & -H_{2^k} \end{bmatrix}$$

for $k = 2, 3, \dots$

The Hadamard matrices are orthogonal and correspond to the Walsh system in that

$$(H_{2^k})_{i,j} = \frac{1}{\sqrt{2^k}} w_j \left(\frac{i}{2^k} \right).$$

Moreover, the Walsh system can be obtained from the Haar system by multiplying the Haar basis elements, represented as finite vectors, by the appropriate-size Hadamard matrix. See [Schippe et al. 90, Section 1.4].

Proposition 4.19. *There exists a constant B_1 such that for all k ,*

$$\|a\|_{l_k^2} \leq \frac{B_1}{\sqrt{2^k}} \|R_k a\|_{l_{2^k}^1}.$$

And in dimension 2^k , for all vectors a supported on the set $D = 1, 2, 4, 8, \dots, 2^k$,

$$\|a\|_{l_{2^k}^2} \leq \frac{B_1}{\sqrt{2^k}} \|H_{2^k} a\|_{l_{2^k}^1}.$$

Proof. We use that r_1, \dots, r_k are constant on the intervals $(\frac{j-1}{2^k}, \frac{j}{2^k}]$ for $j = 1, \dots, 2^k$. Using the Khintchine inequality, we have

$$\begin{aligned} \|a\|_{l^2_k} &= \left\| \sum_{i=1}^n a_i r_i \right\|_{L^2([0,1])} \leq B_1 \left\| \sum_{i=1}^n a_i r_i \right\|_{L^1([0,1])} = B_1 \int_0^1 \left| \sum_{i=1}^k a_i r_i(t) \right| dt \\ &= B_1 \frac{1}{2^k} \sum_{l=1}^{2^k} \left| \sum_{i=1}^k a_i r_i \left(\frac{l}{2^k} \right) \right| = B_1 \frac{1}{\sqrt{2^k}} \sum_{l=1}^{2^k} \left| \frac{1}{\sqrt{2^k}} \sum_{i=1}^k a_i r_i \left(\frac{l}{2^k} \right) \right| \\ &= \frac{B_1}{\sqrt{2^k}} \|R_k a\|_{l^1_{2^k}}. \end{aligned}$$

For the second statement, we simply note that for a supported on $1, 2, 4, \dots, 2^k$,

$$\left\| \sum_{i=1}^{2^k} a_i w_i \right\|_{L^2([0,1])} = \left\| \sum_{i=1}^k a_{2^i} r_i \right\|_{L^2([0,1])},$$

so that the same calculation proves the claim. \square

5. Conclusion and Discussion

We have provided a contribution toward an understanding of the relationship between the peak values of a signal and the proportion of orthonormal signals that can be used for information transmission when one is using tone reservation for PAPR reduction. Our results show that for the two most common wireless systems, OFDM and DS-CDMA, a strict amplitude constraint requires that the proportion of signals used to carry information must decrease as the total number of signals used increases when one is using tone reservation.

One could naively ask whether this is the case for all orthonormal systems. However, we gave examples for both the Fourier and the Walsh cases of subsequences such that the corresponding subspaces have norm equivalence, and thus the PAPR reduction problem is solvable for the infinite subsequence. By simply rearranging the original basis by alternately taking one function from the special subsequence and one from its complement, one has solvability on a subset with density $1/2$. We have seen, though, that when one restricts to finite sets, one no longer has solvability in this setting. Thus, the behavior depends on the properties of the finite set. This can be seen for Walsh functions in the matrix setting as well. Suppose that one alternately selected a Rademacher function and a Walsh function that is not a Rademacher function and represented them as columns in a matrix. The number of rows would then grow exponentially with respect to

the number of columns. Thus, the norm equivalence would not occur on spaces of the same dimension or even proportional dimension.

In the Walsh case, we have seen that the three orderings for the Walsh system all yield the same result. This is because the necessary properties, namely that products only permute the functions within the appropriate dyadic block, are common to all the orderings. We state the informal conjecture that as far as the topics addressed here are concerned, any basis with a useful structure (and uniformly bounded) will behave similarly to the Fourier and Walsh bases.

Acknowledgments. We note that the PAPR problem for CDMA systems was posed by Bernd Haberland and Andreas Pascht, of Bell Labs, in 2000, and the first author thanks them for discussions since that time. We thank Andreas Kortke, of the Technische Universität Berlin, and Wilhelm Keusgen, of the Heinrich Hertz Institute, for discussions concerning power amplifiers, in particular as part of the Smart Radio Frontend Project. We thank Gerhard Wunder, of the Heinrich Hertz Institute, for help with the PAPR literature. The approach taken to solve the PAPR reduction problem for CDMA was motivated by work of Timothy Gowers on quasirandom groups [Gowers 08], which he discussed at his Institute for Advanced Studies lecture in summer 2010. The first author thanks the IAS for its hospitality during the Workshop on Pseudorandomness in Mathematics 2010. Lastly, we mention that parts of this paper were presented in the first author's Bell Labs lecture in late 2010. This work was completed while both authors were with the Technische Universität Berlin and was supported by the German Science Foundation (DFG) under Project BO 1734/18-1.

References

- [Boccaletti et al. 08] B. Boccaletti, M. Löffler, and J. Oppenheim. "How IT Can Cut Carbon Emissions." *McKinsey Quarterly* (October 2008), 0–5.
- [Boche and Farrell 10] H. Boche and B. Farrell. "PAPR for OFDM and the Proportion of Information Bearing Signals for Tone Reservation." In *International Symposium on Information Theory and Its Applications (ISITA), 2010*, pp. 1058–1063, October 2010.
- [Boche and Pohl 07] H. Boche and V. Pohl. "Signal Representation and Approximation: Fundamental Limits." *European Trans. Telecomm. (ETT)* 18:5 (2007), 445–456.
- [Boche et al. 12] H. Boche, B. Farrell, M. Ledoux, and M. Wiese. "Expected Supremum of a Random Linear Combination of Shifted Kernels." *J. Fourier Anal. Appl.* 18:4 (2012), 790–802.
- [Fine 49] N. J. Fine. "On the Walsh Functions." *Trans. Amer. Math. Soc.* 65 (1949), 372–414.
- [Gartner 07] Gartner. "Gartner Estimates ICT Industry Accounts for 2 Percent of Global CO₂ Emissions." Available at <http://www.gartner.com/it/page.jsp?id=503867>, 2007.
- [Gowers 08] W. T. Gowers. "Quasirandom Groups." *Combin. Probab. Comput.* 17:3 (2008), 363–387.

- [Han and Lee 05] S. Han and J. Lee. “An Overview of Peak-to-Average Power Ratio Reduction Techniques for Multicarrier Transmission.” *IEEE Transactions on Wireless Communications* 12:2 (2005), 56–65.
- [Ilic and Strohmer 09] J. Ilic and T. Strohmer. “PAPR reduction in OFDM Using Kashin’s Representation.” In *IEEE SPAWC 2009.*, pp. 444–448, June 2009.
- [Litsyn 07] S. Litsyn. *Peak Power Control in Multicarrier Communications*. Cambridge Univ. Press, 2007.
- [POST 08] “ICT and CO₂ emissions.” *Postnote*. Parliamentary Office of Science and Technology (UK), December 2008.
- [Raab et al. 02] F. Raab, P. Asbeck, S. Cripps, P. Kenington, Z. Popovic, N. Potheary, J. Sevic, and N. Sokal. “Power Amplifiers and Transmitters for RF and Microwave.” *IEEE Transactions on Microwave Theory and Techniques* 50:3 (2002), 814–826.
- [Schipp et al. 90] F. Schipp, W. R. Wade, and P. Simon. *Walsh Series: An Introduction to Dyadic Harmonic Analysis*. Adam Hilger Ltd., 1990.
- [Tao 06] T. Tao. “A Quantitative Ergodic Theory Proof of Szemerédi’s Theorem.” *Electron. J. Combin.* 13:1 (2006), Research Paper 99.
- [Tellado 99] J. Tellado. “Peak to Average Power Reduction for Multicarrier Modulation.” PhD thesis, Stanford University, 1999.
- [Tellado and Cioffi 98] J. Tellado and J. Cioffi. “Efficient Algorithms for Reducing PAR in Multicarrier Systems.” In *Proc. IEEE ISIT, 1998.*, p. 191, August 1998.
- [Tellambura 01] C. Tellambura. “Computation of the Continuous-Time PAR of an OFDM Signal with BPSK Subcarriers.” *IEEE Communications Letters* 5:5 (2001), 185–187.
- [Walsh 23] J. L. Walsh. “A Closed Set of Normal Orthogonal Functions.” *Amer. J. Math.* 45:1 (1923), 5–24.
- [Wojtaszczyk 91] P. Wojtaszczyk. *Banach Spaces for Analysts*, Cambridge Studies in Advanced Mathematics 25. Cambridge University Press, 1991.
- [Wunder and Boche 03] 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* 51:3 (2003), 771–780.
- [Wunder et al. 13] G. Wunder, R. Fischer, H. Boche, S. Litsyn, and J.-S. No. “The PAPR Problem in OFDM Transmission: New Directions for a Long-Lasting Problem.” To appear in *IEEE Signal Processing Magazine*, 2013, arXiv:1212.2865.

Holger Boche, Institute of Theoretical Information Technology, Technical University Munich, Theresienstr. 90, 80333 München, Germany (boche@tum.de)

Brendan Farrell, Computing and Mathematical Sciences, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA (farrell@cms.caltech.edu)