

The PAPR Problem in OFDM Transmission

© ISTOCKPHOTO.COM/TRAFFIC_ANALYZER

[New directions for a long-lasting problem]

[Gerhard Wunder, Robert F.H. Fischer, Holger Boche, Simon Litsyn, and Jong-Seon No]

Digital Object Identifier 10.1109/MSP.2012.2218138
Date of publication: 15 October 2013

Energy efficiency matters in future mobile communications networks. The key driving factor is the growing energy cost of network operations that can make up as much as 50% of the total operational cost today [1]. In the context of green information and communication technology (ICT), this has led to many global initiatives such as the Green Touch consortium.

A major source for reducing energy costs is to increase the efficiency of the high-power amplifier (HPA) in the radio frequency (RF) front end of the base stations [1]. However, efficiency of the HPA is directly related to the peak-to-average power ratio (PAPR) of the input signal. The problem especially becomes serious in orthogonal frequency-division multiplexing (OFDM) multicarrier transmission, which is applied in many important wireless standards such as the Third Generation Partnership Project (3GPP) Long-Term Evolution Advanced (LTE-A) standard. In the remainder of this article, we refer to it simply as the PAPR problem. The PAPR problem still prevents OFDM from being adopted in the uplink of mobile communication standards, and, besides power efficiency, it can

also place severe constraints on output power and therefore coverage in the downlink.

In the past, there have been many efforts to deal with the PAPR problem resulting in numerous papers and several overview articles, e.g., [2]–[4]. However, with the upcoming novel systems, new challenges emerge that have been rarely addressed so far: 1) the envisioned boost in network energy efficiency (e.g., at least by a factor of 1,000 in the Green Touch consortium) will tighten the requirements on the component level so that the efficiency gap with respect to single-carrier transmission must considerably diminish; 2) multiple-input multiple-output (MIMO) multiplies the problem due to simultaneous control of parallel transmit signals particularly when considering a huge number of transmit antennas; and 3) multiuser (MU) (and multipoint) systems put additional side constraints on the parallel transmit signals that are difficult to implement on top of conventional approaches. Furthermore, many of the existing methods are not compatible with relevant standards and/or their prospective performance capabilities are not satisfactory. Yet, it is quite safe to say that no standard solution is available.

ENERGY EFFICIENCY IN MOBILE COMMUNICATION NETWORKS: A DRIVING SOURCE FOR INNOVATION

In this article, we will argue that, in light of these challenges, the PAPR metric itself has to be carefully reviewed within a much broader scope, overthrowing some of the common understanding and results. New metrics become more important since they enable the system designer to precisely adjust the algorithms to meet some given performance indicator. It is expected that such a design approach will no longer be treated like an isolated problem on the physical layer but will affect the design parameters on higher layers as well (e.g., resource allocation). For example, it has been discussed in [1] that from a ICT perspective, the system throughput should be related to input power rather than output power. To capture this paradigm on a HPA power efficiency level, different metrics are currently used, including total degradation and average distortion power. However, with respect to algorithm design, all these metrics are solely reflected by the standard PAPR figure of merit. This argument can also be extended to other situations: it has been recently shown in [5] that, if the only concern is average distortion power (instead of peak power), then a much less conservative design is possible compared to conventional design rules in OFDM transmission. Remarkably, such performance limits can be achieved efficiently using derandomization algorithms, establishing therefore a new powerful tool within the context of the PAPR problem. It is a major aim of this article to review and collect exactly those elements in the current literature that we believe represent the core of a more general theory.

THE PAPR METRIC ITSELF HAS TO BE CAREFULLY REVIEWED WITHIN A MUCH BROADER SCOPE OVERTHROWING SOME OF THE COMMON UNDERSTANDING AND RESULTS.

Besides this point of view, it is interesting to apply new signal processing and mathematical concepts to the PAPR problem. Compressed sensing [6], [7] is a new framework capturing sparsity in signals beyond Shannon's sampling theorem and has attracted a lot of attention in recent years. It is based on the observation that, for a sparse signal, a small number of linear projections (measurements) contain enough information for its recovery. Compressed sensing can be applied to the PAPR problem because sparsity frequently appears in the clipped portions of the distorted OFDM signals. There are currently many research efforts in this direction, but some challenges still remain, such as the degraded recovery performance

in noisy environment. The adoption of related mathematical concepts such as Banach space geometry [8], [9] complement this discussion. It is outlined that these theoretically deeply rooted concepts can help to understand some of the fundamental limits as well as to develop new algorithmic solutions for the PAPR problem.

To summarize, there is clearly a need for a fresh look at the PAPR problem under the general umbrella of the metrics theme discussed previously which will open up new research strands not yet explored. In this article, we are going to address and discuss some of the fundamentals, challenges, latest trends, and potential solutions that originate from this perspective and that we believe to be important to come to an innovative breakthrough for this long-lasting problem.

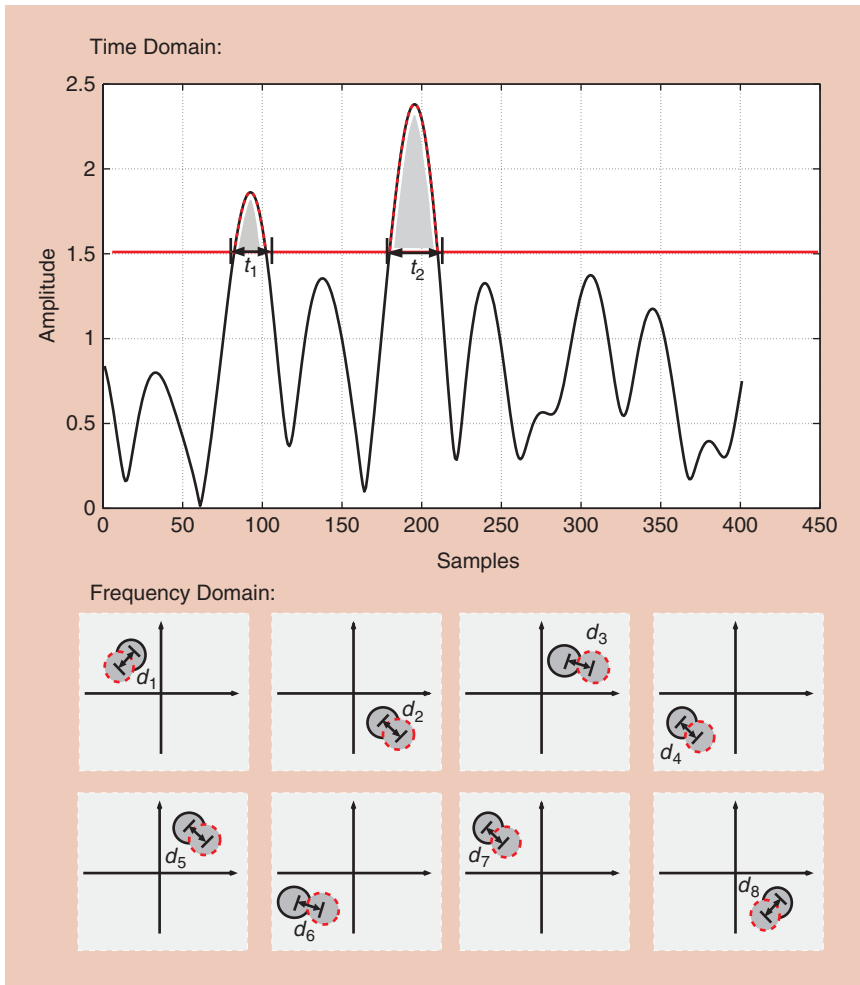
NOTATIONS

We recall the following standard notations: the frequency-domain OFDM symbol for each antenna (N_t in total) consists of N subcarriers. The multiplexed transmit symbols $C_{m,n}$ (carrying information and/or control data) are drawn from some common quadrature amplitude modulation/phase-shift keying (QAM/PSK) signal constellation and collected in the space-frequency codeword $C := [C_1, \dots, C_{N_t}]$, where $C_m := [C_{m,1}, \dots, C_{m,N^t}]$ is the transmit sequence of antenna m . In case of a single antenna we write $C_n := C_{1,n}$ and, correspondingly, $C := C_1 = [C_1, \dots, C_m]^T$. Given the inverse discrete Fourier transform (IDFT) matrix $F := [e^{j2\pi kn/(LN)}]_{0 \leq k < LN, 0 \leq n < N}$ the I -times oversampled discrete-time OFDM transmit symbols in the equivalent complex baseband at antenna m are given by $s_m = FC_m$. The average power of this signal might be normalized to one. We define the PAPR of the transmit signal at antenna m as

$$\text{PAPR}(s_m) := \|s_m\|_\infty^2. \quad (1)$$

COMMENT ON OVERSAMPLING

Please note that PAPR of the continuous-time passband signal differs roughly by 3 dB. Clearly, there is also still some overshooting between the samples, but, due to sufficiently high oversampling, the effect is negligible. The tradeoff between overshooting and oversampling is one of the few subproblems



[FIG1] An illustration of distorted OFDM signal (in time domain, clipping level = 3.5 dB) and corresponding impact on the subcarrier signaling points (in frequency domain): the distortion signal is typically a sequence of clips with clip duration t_1, t_2, \dots . The SDR metric then relates the mean useful signal energy to the mean distortion energy while EVM collects the mean of sum of squared errors d_1, d_2, \dots in the data sequence (due to the very same time domain distortion). In case of Nyquist sampling, both are actually equal (subject to a scaling factor) [11], [12]. (Figure courtesy of Gerhard Wunder.)

in OFDM transmission that is well understood. The best known results that hold even in the strict band-limited case are given in [10], where overshooting is proved to be below $1/\cos(\pi/2I)$.

THE DESIGN CHALLENGE

In OFDM transmission, many subcarriers (constructively or destructively) add up at a time that causes large fluctuations of the signal envelope; a transmission that is free from any distortion requires linear operation of HPA over a range N times the average power. As practical values of subcarriers are large, these high dynamics afford HPA operation well below saturation so that most of the supply power is wasted with deleterious effect on either battery life time in mobile applications (uplink) or energy cost of network operation (downlink). In practice, these values are not tolerable, and from a technology viewpoint it is also challenging to provide such a large linear range. Hence, the HPA output signal is inevitably cut off at some point relative to the average

power (clipping level) leading to in-band distortion in the form of intermodulation terms and spectral regrowth into adjacent channels. The effect is illustrated in Figure 1, where the distorted OFDM signal and corresponding impact on the signal points are depicted.

The PAPR problem brings up several challenges for the system designer: one challenge is to adjust HPA parameters (HPA backoff, digital predistortion) in some specific way so that power efficiency is traded against nonlinear distortion, which effects the data transmission on a global scale. How to capture this tradeoff by a suitable metric on a component level is not clear yet. Special HPA architectures such as Doherty [13] and others can help to improve on this tradeoff. We also mention that other design constraints such as costs might prevent specific architectures [14].

A second challenge is to process the baseband signal by peak power reduction algorithms in such a way that the key figures of merit in the aforementioned tradeoff are improved. This alternating procedure makes it apparent that the PAPR problem involves joint optimization of HPA, predistortion, and a signal processing unit. This interplay has only been marginally addressed so far, let alone in the context of multiuser systems equipped with multiple antennas such as LTE-A.

In the following, we discuss some potential metrics that can be used in the optimization.

THE RIGHT PARADIGM? ALTERNATIVE METRICS FOR PAPR

Classically, in OFDM transmission, the PAPR of the transmit signal is analyzed and minimized by applying transmitter-side algorithms. Meanwhile, it has been recognized that it may be reasonable to study other parameters as well. For example, when aiming at minimizing the energy consumption of the transmitter including the analog front end or when operating a low-cost, low-precision power amplifier—sometimes referred to as *dirty RF* [14]—potentially other signal properties need to be controlled.

Let us present some illustrative examples first. Suppose we are interested in the clipped energy instead of the PAPR (we give some justification for this in terms of capacity below). Naturally, since the total energy is approximately one, the clipped energy is finite as well, but when N increases, the required clipping level for asymptotically zero clipping energy might be of

interest for design purposes. Clearly, no clipping at all is trivially sufficient but, surprisingly, it is actually *not* necessary: it is proved in [5] that clipping level can be adjusted along the $\log\log(N)$ law so that it is practically almost constant. This stands in clear contrast to the $\log(N)$ PAPR scaling discussed in the section “Approaching the $\log(N)$ Barrier: Derandomization.”

Subsequently, some alternative metrics are briefly summarized.

- Of course, the PAPR itself has justification. As it relates peak and mean power, the PAPR is the adequate metric for quantifying the required power backoff of the power amplifier. When using higher-order modulation per carrier, the energy per OFDM frame is no longer constant and hence average power fluctuates. In this case, the peak power instead of PAPR is as suited metric. Both metrics are well-suited measures if purely limiting effects (modeled, e.g., as soft limiter) should be characterized. Since PAPR is random, we are also interested in the complementary cumulative distribution function (CCDF) $F(x)$ and other characteristic figures such as the mean $\mathbb{E}\{\cdot\}$, etc.

- Besides looking at the transmit signal itself, in many situations the impact of the nonlinear power amplifier to system performance is of interest. One possible approach is to quantify the nonlinear distortions caused by a particular power amplifier model. The signal-to-distortion ratio (SDR) [15], [16] and error vector magnitude (EVM) [11], [12] capture the in-band distortion of the OFDM signal and are immediately related to error rate of uncoded transmission. Both, SDR and EVM and their interdependence are illustrated in Figure 1.

- Within 3GPP, a power amplifier model that causes nonlinear distortions according to the third power of the RF transmit signal has become popular (so-called *cubic polynomial model*) [17]. In this case, the cubic metric (CM) measuring the mean of time domain sample energy to the third power is well adapted to the specific scenario [18], [19]. However, any additional clipping is not included here.

- CM metric is a special case of the amplifier-oriented metric (AOM) defined in [19] and [20]. AOM measures the mean squared absolute difference of desired and distorted HPA output. Here, the HPA output is calculated based upon some model such as the mentioned cubic polynomial model (or the well-known Rapp model).

- A very severe problem in communication systems is that nonlinear devices cause a spectral widening and, hence, generate out-of-band radiation. To avoid violating spectral masks imposed by the regulatory body, a metric quantifying the out-of-band power or the shoulder attenuation is desirable [20]. In this field, significant work has to be done.

- When applying strong channel coding schemes—which today is the state of the art in OFDM transmission—SDR and EVM are no longer suited performance characteristics. Instead, the end-to-end capacity of the entire OFDM scheme including the nonlinear devices matters. Unfortunately, neither the capacity of the continuous-time peak-power limited additive white Gaussian noise (AWGN) channel itself, nor the

capacity of OFDM over such channels are known. Moreover, when using (as often done) a statistical model for the behavior of the nonlinear device, only lower bounds on the capacity are obtained as the statistical dependencies within one OFDM frame are not exploited.

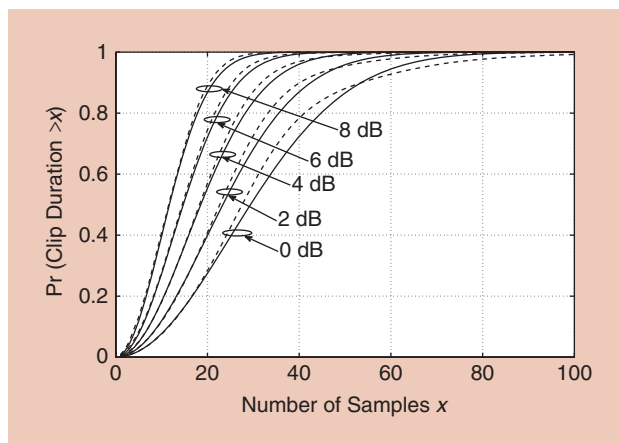
However, recent work [21], [22] indicates that clipped OFDM performs (almost) the same as unclipped OFDM with signal-to-noise ratio (SNR) reduced according to the clipping power loss. Hence, the main source of the loss are not the introduced distortions or errors but simply the reduced output power. This, in turn, leads to the conclusion that a suited metric for capacity maximization is simply the average power of the power amplifier output signal. In this case, unfortunately, the generation of out-of-band radiation is not penalized.

- The symbol error rate (SER) is a related measure that has been directly applied to peak power control algorithms in [23].

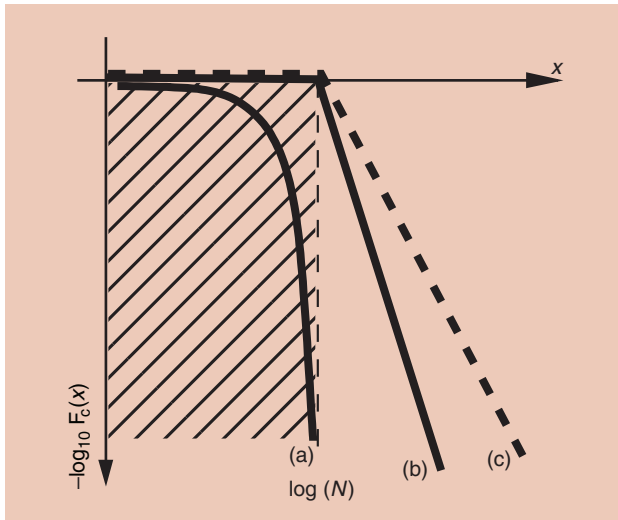
- In future applications, more than a single signal parameter will have to be controlled. For example, the capacity should be maximized but at the same time the out-of-band power should be minimized. Consequently, suited combinations of metrics capturing the desired trade off are requested.

COMMENT ON GAUSSIAN APPROXIMATION

It is worth mentioning that many metrics have been analyzed in the past with the help of the Gaussian approximation. This, however, is not in all cases a feasible path. It is true that as N gets large, the finite-dimensional distributions converge, and, since the signals are band limited, the process itself also converges. However, we mention that it is a local property, valid only for any finite interval. For example, in Figure 2, the empirical CCDF of the clip duration (see the illustration in Figure 1) is shown for OFDM signals and compared to a widely known result for envelopes of Gaussian processes (and their Hilbert transforms) for large clipping levels [24]. To be specific, [24] derives the so-called Palm distribution of the clip duration,



[FIG2] The palm distribution of clip duration of an OFDM signal with $N = 2,048$ and different clipping levels in decibels. Note that the theory (dotted) is valid only for large clipping levels but fits relatively well also for low clipping levels.



[FIG3] An illustration of the virtue of the LDP is given: (a) derandomization, (b) SLM, and (c) uncoded data.

which describes the statistical average after an upcrossing of the Gaussian process. It is seen that simulation and analysis match very well [23]. On the other hand, metrics such as EVM [12] and SER [25] have been shown not to match well.

APPROACHING THE $\log(N)$ BARRIER: DERANDOMIZATION

In this section, we establish several fundamental principles for the PAPR problem. We believe that all of them actually connect to a broader theory general enough to capture alternative metrics as well, and this will open the door for new, provably more efficient algorithms.

THE LDP

By analyzing PAPR of multicarrier signals, one faces a fundamental barrier that to overcome seems quite challenging: the $\log(N)$ barrier (recall that N is the number of subcarriers). In fact, it is an exercise in large deviations to show that multicarrier signals with statistically independent subcarriers have PAPR of $\log(N)$ in a probabilistic sense [26]–[28]. This means that with very high probability, the PAPR lies in an open, arbitrarily small interval containing $\log(N)$: this is what we call the *large deviation principle (LDP)*.

Implicitly, LDP affects the performance of many peak power control schemes. The LDP has been known since long in the context of random polynomials but in the OFDM context, the most general form is due to [29] where it is shown that as long as N is large enough and the subcarriers are independent that the following inequality is true:

$$\Pr\left\{\left|\sqrt{\text{PAPR}} - \sqrt{\log(N)}\right| > \gamma \frac{\log[\log(N)]}{\sqrt{\log(N)}}\right\} \leq \frac{1}{[\log(N)]^{2\gamma - \frac{1}{2}}}. \quad (2)$$

Here, $\gamma > 1/4$ is some design parameter that trades off probability decay over deviation from $\log(N)$. While the analysis is

tricky when it comes to show that PAPR is not below $\log(N)$ with high probability, it is a surprisingly easy task to show the converse: standard inequalities (such as Chernoff bounds) or any other Markov-style bound do the job. Some can be exploited for algorithm design as shown later in this section.

The inequality states that PAPR concentrates more and more around the value $\log(N)$, which therefore establishes an important theoretical scaling law. The proof is technical but the result might be surprising since 1) the factor before the logarithmic term is exactly unity and 2) the scaling law differs from the well-known law of iterated logarithm, which would suggest only doubly logarithmic scaling.

The LDP contains some valuable illustrative aspects that we are going to reveal now. The LDP in (2) is somewhat unaccessible and shall be rewritten in the more convenient form

$$\log(F_c(x)) = [\log(N) + O(\log[\log(N)]) - x]^- , \quad (3)$$

where we used the order notation $O(\cdot)$ and the definition $[x]^- := \min(0, x)$. Disregarding the order term $O(\log[\log(N)])$, we have the interpretation that the probability decreases linearly on a logarithmic scale from some cutoff point $\log(N)$, which is illustrated in Figure 3. The proximity to filter design terminology is intended and it makes obvious sense to speak of a passband and a stopband in the figure. Comparing this to the standard analysis where statistical independence and Nyquist sampling is assumed gives

$$\log(F_c(x)) = [\log(N) - x]^- , \quad (4)$$

where the order term is missing. Hence, we conclude that a careful non-Gaussian analysis for continuous-time OFDM signals entails an error of at most $O(\log[\log(N)])$.

The LDP is very useful for assessing the performance of peak-power control schemes. Before we show this we might ask why this concentration happens? Let C_1, C_2, \dots, C_N be a (data) sequence of independent random variables; when estimating PAPR without a priori information the expectation is the best possible choice. Using successive knowledge of already fixed data, we have the following estimations:

$$y_0 = \mathbb{E}\{\text{PAPR}(s)\} \quad (5)$$

$$y_1 = \mathbb{E}\{\text{PAPR}(s) | C_1\} \quad (6)$$

$$y_2 = \mathbb{E}\{\text{PAPR}(s) | C_1, C_2\} \quad (7)$$

...

$$y_N = \mathbb{E}\{\text{PAPR}(s) | C_1, C_2, \dots, C_N\}. \quad (8)$$

It can be shown that this process establishes a Martingale with bounded increments $|y_i - y_{i-1}|$ from which it follows (see [30]) measure concentration of the PAPR around its average via the Azuma–Hoeffding inequality or McDiarmid’s inequality. Another approach used in [30] for proving measure concentration of the PAPR around its median is based on the convex-hull distance inequality of Talagrand. The tails of the concentration inequalities are even exponential then. Let us now apply the LDP within the context of peak power control.

MULTIPLE SIGNAL REPRESENTATION AND PARTITIONING

The basic principle for most of the peak power control schemes is multiple signal representation, which roots in classical methods such as selected mapping (SLM) [31] and partial transmit sequences (PTS) [32], [33]. The idea is simple: instead of transmitting the original OFDM data frame, multiple redundant candidates are generated and the “best” candidate is singled out for transmission. By using suitable transforms or mappings, the main goal is to achieve statistical independence between the candidates’ metrics. Clearly, instead of PAPR, alternative metrics can also be used in the selection process [19], [20]. SLM and PTS are similar; the difference between SLM and PTS is that the mappings are applied to a subset of the data frame.

For SLM, many transforms have been proposed: in the original approach, the data frame is element-wise multiplied by random phases; other popular approaches include binary random scrambling and permutation of the data (see references in [19]; here is also where the transmission of side information is discussed). Similar for PTS, random phases have been used as well. While a full search is typically carried out, efficient algorithms to find the phases have been proposed. An exhaustive list can be found in [2].

SLM can be analyzed within the context of LDP. The transforms define U alternatives each assumed with independent PAPR. Clearly this independence assumption is crucial: it might be argued that it has to hold for the PAPR only but the model clearly fails when the number of alternatives is large. By exploiting the LDP we have simply then

$$\log(F_c(x)) = U [\log(N) - x] \quad (9)$$

so that the decay is U times faster as depicted in Figure 3. A similar analysis can be carried out when the selection is done, directed, or extended over the time domain [19]. Note that, in principle, PTS can be analyzed as well; however, since the transform is on subsets of the data, the independence assumption is far more critical. Another main problem so far is that side information is treated separately and not within the same communication model.

A better model is complete partitioning of the set of transmit sequences. The idea is illustrated in Figure 4. Suppose that the transmit sequence belongs to some set that is partitioned into many cells, all of them containing the same information. Note that if the actual cell selection is required at the receiver for decoding, side information is generated. This side information belongs in our general model to the transmit sequence itself and must be specially protected. This can be done via an embedded code that is decoded before or after the actual information decoding procedure [2]. Let us mark the subsets where PAPR is below some threshold: the reasoning is that by the mapping of codewords from one cell into another, sets with larger PAPR should be mapped to a marked subset by at least one mapping that will ensure peak power below the threshold. Obviously, the definition of such a mapping will determine the performance of the scheme.

One of the simplest examples is when the data is over some constellation and side information is encoded into a sequence of

binary phase-shift keying (BPSK) symbols: each sequence defines a specific BPSK vector determining the sign vector. Both modified information and side information sequence define the transmitted codeword. This method is called sequence balancing [34]. It is characteristic for this method that correlation is inserted in the stream by using suitable binary codes. We will call this the binary correlation model. Note, if the side information is purely redundant, the method reduces to tone reservation [4]. Moreover, if the selection defines phase relations between partial sequences then it is a version of PTS [32]. A related approach is Trellis shaping [2], [35].

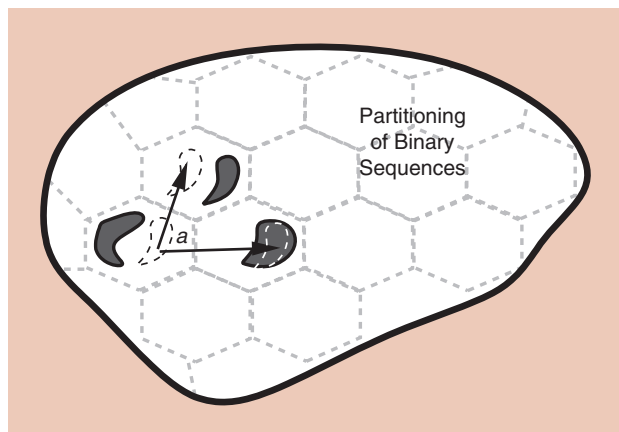
Sequence balancing using binary codes can achieve (even though easily generalized) already a sufficient fraction of the theoretically possible performance gain: the main required property of the set of binary vectors is their ability of as many sign changes as possible over any subvector, which is called the *strength of the code* [34]. Many binary codes have this property and are thus suited for this procedure. The strength is related to the dual distance. It can be shown that if the strength grows as $\log(N)$ then PAPR is below $\log(N)$ for large N . Unfortunately, similar to SLM and PTS, the number of candidates grows as well.

There are other methods that use partitioning as well, such as tone injection [4] where the constellation is artificially extended or translates of codes [36]. Schemes such as active constellation extension [37] introduce redundancy as well but can be continuously formulated so that other methods such as convex optimization can be applied.

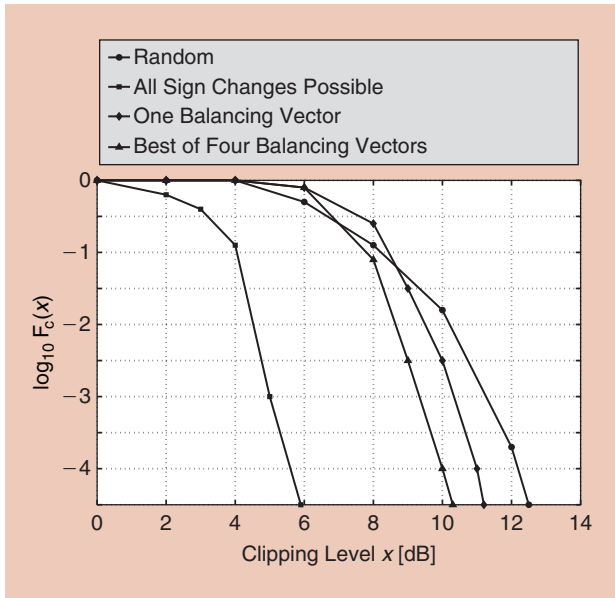
All discussed approaches assume to run a full cell selection search that is too complex in many situations. A better approach is discussed next.

DERANDOMIZATION OF CHOICES

The LDP provides a method to circumvent full search by assuming a suitable underlying probability model for the cell selection. By derandomizing the cell selection one can easily devise suitable algorithms guaranteeing a PAPR reduction very close



[FIG4] A general model for peak power control: each cell contains the same information; sets with large PAPR (dotted areas) are mapped to sets where PAPR is below some threshold (grey-shaded areas).



[FIG5] This figure compares the sequence balancing method of the section “Multiple Signal Representation and Partitioning” in terms of the CCDF of PAPR for a 128 subcarrier OFDM system using 1) derandomization algorithm where all possible subcarrier sign changes are allowed and 2) codes of given strength with randomly chosen balancing vectors. For the code a strength ten dual BCH code with only 18 redundant subcarrier is applied. Please note that the simulation matches very well the theory predicting cutoff point $\log(128) \approx 6.8$ dB.

to the $\log(N)$ barrier [29], [38], [39]. The basis algorithm goes back to Spencer [40], who called it the probabilistic method.

The derandomization method is best explained along an example: consider again the binary correlation model where any possible sign change for some information sequence C is allowed. Denote this sign vector by $A = [A_0, \dots, A_{N-1}]$ and the resulting transmit sequence by $A \odot C$ (respectively $s_{A \odot C}$). Suppose that all the sign changes happens at random with equal probability and each sign change is independent. As for the LDP, define the random variables $y_i(A_0^*, \dots, A_{i-1}^*) = \{\text{PAPR}(s_{A \odot C}) \mid A_0^*, \dots, A_{i-1}^*\}$. Then we can mimic the steps (5)–(8) and successively reduce randomness by applying

$$A_i^* := \arg \min_{a_i} y_i(A_0^*, \dots, A_{i-1}^*, a_i).$$

By the properties of (conditional) expectations

$$y_i(A_0^*, \dots, A_{i-1}^*, A_i^*) \leq y_i(A_0^*, \dots, A_{i-1}^*)$$

and finally $\text{PAPR}(s) \leq y_0$ since $y_N(A_0^*, \dots, A_{N-1}^*)$ is simply a nonrandom quantity. Finally, by the LDP $y_0 \leq \log(N)$ for N large enough. Since the expectations are somewhat difficult to handle instead of the $\text{PAPR}(s)$ typically the set function and corresponding bounds have been used. For example, Chernoff bounds have been used in [38], [29], and [41] showing good performance and low complexity. Moment bounds with better tail properties have been used but the complexity is higher [42]. Performance results of the derandomization method are reported in Figure 5 comparing sequence balancing (see the section “Multiple Signal Representation and Partitioning”) with and without

derandomization. The benefit of the derandomization method is clearly observed and corresponds to more than 4 dB gain in HPA backoff (at 10^{-3} outage probability), which mimics exactly the results of the LDP analysis (see the section “The LDP”). However, it comes at the cost of 1 bit/dimension rate loss. The relevant tradeoff between rate and PAPR have been rarely investigated so far.

Combining the derandomization method with partitioning yields several improved algorithms for standard problems. For example, the PTS method has been applied in [43]. It is proved that with derandomization method PTS can achieve $r \log(N)$ where r is the percentage of partial transmit sequences related to N . The tone reservation method has been treated using derandomization in [29] (see also the section “Tone Reservation and Szemerédi’s Theorem”). Implicitly, derandomization has been used in [36] to show that PAPR of some translate of a code C is below $|C| \log(N)$. Related derandomization algorithms have been used in [39] adopting the so-called *pusher-chooser game* from [40]. The idea is to choose l_p -norms and to prove a recursive formula similar to the Chernoff method. The approach can be generalized to alternative metrics if appropriate bounds are available: in [23] the SER has been reduced using derandomization showing that $\log(N)/2$ clipping level is sufficient asymptotically for zero error probability [instead of $\log(N)$]. Furthermore, in the recent paper [5], zero clipped energy is asymptotically achieved with clipping level $\log \log(N)$.

There is still plenty of room for improvements, e.g., by considering correlations between different samplings points and incorporating other metrics as well [5], [23]. It has not been noticed yet that this field is particularly underdeveloped and bears great potential for significant improvements of current systems. Another point to be improved is the rate loss imposed by the current methods.

ADDITIONAL RESOURCES: MIMO AND MULTIUSER SYSTEMS

While the most beneficial in terms of spectral efficiency, MIMO systems complicate the PAPR problem: in single-antenna systems, the PAPR (or other metrics) of only one transmit antenna has to be controlled. In the MIMO setting, a large number of OFDM signals are transmitted in parallel, and typically the worst-case candidate dictates the PAPR metric (e.g., due to out-of-band power) [19].

As a consequence, PAPR reduction methods tailored to this situation should be utilized instead of performing single-antenna PAPR reduction in parallel. Multiantenna transmitter provide additional degrees of freedom that can be utilized beneficially for PAPR reduction—the full potential has not yet been explored in the literature. Basically, the peak power can be redistributed over the antennas. By this, MIMO PAPR reduction may lead to an increased slope of the CCDF curves (cf. Figure 3), i.e., the probability of occurrence of large signal peaks can be significantly lowered compared to single-antenna schemes. This effect is similar to that of achieving some diversity gain.

When studying MIMO PAPR reduction schemes, two basic scenarios have to be distinguished: on the one hand,

in point-to-point MIMO transmission, joint processing of the signals at both ends (transmitter and receiver) is possible. On the other hand, in point-to-multipoint situations, i.e., multiuser downlink transmission joint signal processing is only possible at the transmitter side. In the multipoint-to-point scenario (multiple-access channel) no joint optimization of the transmit signals can be performed, hence this case is not amenable for MIMO approaches. This fact heavily restricts the applicability of PAPR reduction schemes.

For the point-to-point setting, a number of PAPR reduction schemes have been designed, particularly the extension of SLM. Besides ordinary SLM (conventional SLM is simply applied in parallel) simplified SLM (the selection is coupled over the antennas) has been proposed in [44]. Directed SLM [45] is tailored

to the MIMO situation and successively invests complexity (test of candidates) only where PAR reduction is really needed.

It might be sufficient that the PAPR stays below a tolerable limit, determined by the actual radio front end. Here, complexity can be saved if candidate generation and assessment is done successively and stopped if the tolerable value is reached. Interestingly, the average number of assessed candidates is simply given by the inverse of the cumulative distribution function (CDF) of PAPR of the underlying original OFDM scheme. It is noteworthy to mention that, for $\text{PAPR} = \log(N)$ and reasonably large number N of carriers, average complexity per antenna is in the order of $e = 2.718\dots$ (Euler's number) [46]. Alternative metrics have been used in [47].

Compared to point-to-point MIMO systems, developing PAPR reduction schemes applicable in point-to-multipoint scenarios (multiuser downlink) is a much more challenging task. Since no joint receiver-side signal processing is possible, for the candidate generation at the transmitter side, only those operations are allowed that can individually be reversed at each of the receivers. Among the SLM family, only simplified SLM can be used here. However, in this situation, the usually present transmitter side multiuser pre-equalization can be utilized for PAPR reduction. Applying Tomlinson–Harashima precoding, the sorting in each carrier can be optimized to lower PAPR at almost no cost in (uncoded) error rate [19]. The same is true when applying lattice-reduction-aided pre-equalization. Here the unimodular matrices (describing a change of basis) can be optimized to control the properties of the transmit signals [48]. There are also links from MIMO PAPR reduction and derandomization to code design (cf. the section “New Trends in Code Design”).

GOING BEYOND: OFDM CAPACITY FUNDAMENTALS

While the capacity of the discrete-time peak-power-constraint channel is known and computable, the capacity of the OFDM peak-power-constraint channel is still an open problem [49],

[50]. The problem is indeed intricate as it has been unknown until very recently that there are exponentially many OFDM signals with constant PAPR (cf. the section “New Trends in Code Design”). However, no practical encoding scheme is known that even comes close to this merely theoretical result. From this perspective the capacity problem awaits a more thorough theoretical solution.

Recent work [21], [22] on practical schemes indicates that clipped OFDM performs (almost) the same as unclipped OFDM with SNR reduced according to the clipping power loss. The main source of the loss is not the introduced distortions or errors but simply the reduced output power. Given the OFDM frame in frequency domain $C = [C_1, \dots, C_N]$ via IDFT the time-domain samples $s[k]$ are calculated. These samples

then undergo clipping in the amplifier front end. As usual, the clipping behavior can be described by a nonlinear, memoryless point symmetric function $g(x)$ (with $g(x) \leq x$, $x > 0$, applied element-wise to vectors). In frequency domain, the clipped signal is given by $Z = \text{DFT}\{g(\text{IDFT}\{C\})\}$. Note that clipping is a deterministic function and a one-to-one relation between the vector C of unclipped symbols and the vector $Z = [Z_1, \dots, Z_N]$ of clipped ones exist. Assuming an AWGN channel, at the receiver side, the vector Z , disturbed by additive white Gaussian noise, is present. In case of intersymbol-interference channels, the symbols Z_n are additionally individually scaled by the fading gain at the respective carrier.

This clipping behavior can be visualized for $N = 3$ and two PAM per carrier; see Figure 6(a). The initial hypercube with vertices given by all possible vectors C is distorted. However, the attenuation of the useful signal (the vector Z has lower energy) will be the dominating effect over deformation. This, in turn, leads to the conclusion that a suited metric for capacity maximization is simply the average power of the power amplifier output signal.

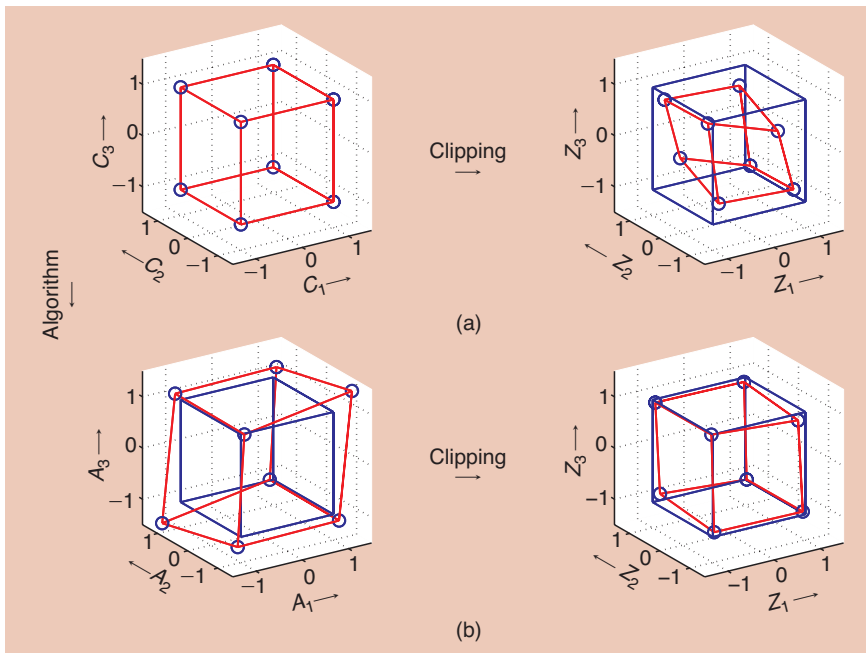
A possible strategy is shown in Figure 6(b). A signal shaping algorithm may adjust the signal points in $2N$ -dimensional real-valued space such that after clipping the set of all possible OFDM vectors in frequency domain forms (approximately) a hypercube with energy close to that of the initial constellation. The first work on using the strategy of active constellation extension for achieving its goal has been presented [51].

EMERGING SOLUTIONS: AN OPEN FIELD

NEW TRENDS IN CODE DESIGN

Jones et al. [52] were the first to describe block coding schemes in the present context. This framework has been put in systematic form by observing the connection of cosets of Reed–Muller codes and complementary sequences [53], [54]. Unfortunately, these approaches have limited potential for modern OFDM

COMPRESSED SENSING CAN BE REGARDED AS MINIMIZING THE NUMBER OF MEASUREMENTS WHILE STILL RETAINING THE INFORMATION NECESSARY TO RECOVER THE ORIGINAL SIGNAL WELL.



[FIG6] (a) A visualization of the effect of clipping on the set of possible OFDM frames (here: $N = 3$, two PAM per carrier). (b) A visualization of predistortion via an algorithm for maximizing the power of the signal after clipping.

systems due to their small coding rate. The fundamental trade-off between different code key properties such as rate, PAPR, etc., was explored and discussed in [55]. More recent ideas use the idea of sequence balancing and code extensions in form of erasure coding in other domains (e.g., MIMO [56]) to tackle the PAPR problem with an inner code, while error correction still is done via an outer code [34].

CODES AND SEQUENCES WITH LOW PAPR

Though most of multicarrier signals of length N have PAPR close to $\log(N)$, it turns out that signals with constant PAPR are not so rare. Using a remarkable result of Spencer [57] it is possible to show that the number of such BPSK modulated signals, is exponential in N . Specifically, there are at least $(2 - \delta_K)^N$ such signals with PAPR not exceeding K , where δ_K is a constant depending on K and tending to zero when K grows. It is an open question how to generate many signals for given K .

Much research was devoted to describing signals with low values of PAPR. For BPSK modulated signals an extreme example is provided by Rudin–Shapiro sequences defined recursively from $P_0 = Q_0 = 1$, and

$$P_{m+1} = (P_m, Q_m), \quad Q_{m+1} = (P_m, -Q_m).$$

These sequences of length being a power of two have PAPR at most two. More general examples of sequences with PAPR at most two arise from Golay complementary sequences. Two sequences constitute a complementary pair if the sum of the values of their aperiodic correlation functions sum up to zero. Many methods are known for constructing such sequences, see [2, Sec. 7.6]. Notice that it is not known if BPSK

modulated signals can have PAPR less than two. However, if one increases the size of multiphase constellations to infinity, there exist sequences with PAPR approaching one [2, Th. 7.37]. For constructions of multiphase complementary pairs from cosets of Reed–Muller codes, see [58] and references therein. PAPR of m -sequences and Legendre sequences is discussed in [2, Sec. 7.7] and [2, Sec. 7.8].

Often we need to know the biggest PAPR among sequences belonging to a code. Bounds on PAPR of codes on sphere as a function of their sizes and minimum Euclidean distances was studied in [55]. A relation between the distance distribution of codes and PAPR was derived in [29]. This yielded bounds on PAPR of long algebraic codes, such as BCH codes. Analysis of PAPR of codes with iterative decoding (for instance, LDPC codes) remains an open problem. PAPR of codes of small size was studied in [55]. In particular, it

was shown that PAPR of duals of length N BCH codes are at most $\log^2(N)$. Bounds on PAPR of Kerdock and Delsarte–Goethals codes were derived as well. In [34], it was shown that in every coset of a code dual to BCH code with the minimum distance of $\log(N)$ exists a codeword with PAPR at most $\log(N)$. At the same time, this leads to a very modest rate loss. Still, constructing codes having low PAPR and high minimum distance seems to be a challenge.

Computing PAPR of a given code is a computationally consuming problem. If a code has a reasonably simple maximum-likelihood decoding algorithm it is possible to determine efficiently its PAPR [59], [60].

In [56], off-the-shelf channel codes, in particular Reed–Solomon (RS) and Simplex codes are employed to create candidates, from which, as in SLM, the best are selected. The codes are thereby arranged over a number of OFDM frames rather than over the carriers. Such an approach is very flexible as due to the selection step any criterion of optimality can be taken into account. Moreover, instead of applying the approach to the MIMO setting, it can also be used if block of temporal consecutive OFDM frames are treated jointly. The method is illustrated in Figure 7.

CONSTELLATION SHAPING

In constellation shaping, we have to find a constellation in the N -dimensional frequency domain, such that the resulting shaping region in the time domain has low PAPR. At the same time, we would like to have a simple encoding method for the chosen constellation. Such shaping based on Hadamard transform was considered in [61]. The main challenge in constellation shaping is to find a unique way of mapping (encoding)

and its inverse (decoding) of reasonable complexity. The suggested approach in [61] and [62] is based on a matrix decomposition. Though the simulation results are quite promising, the implementation complexity still seems to be far from being affordable [62], [63].

BANACH SPACE GEOMETRY

An interesting new approach to the PAPR problem is that of using Banach space geometry. Banach space geometry relates norms and metrics of different Banach spaces to each other. For example, a question that often arises is: assume a Banach space with unit norm ball B_1 and another Banach space with unit ball B_2 ; both spaces are of finite, possibly different dimension. What is the relation between the norms if the projection of one ball covers the other ball? Furthermore, what is the dependence of this relation on the dimensions?

Interestingly, these relations turn out to be useful for the PAPR problem in several other ways depending on the underlying Banach spaces as the following examples show.

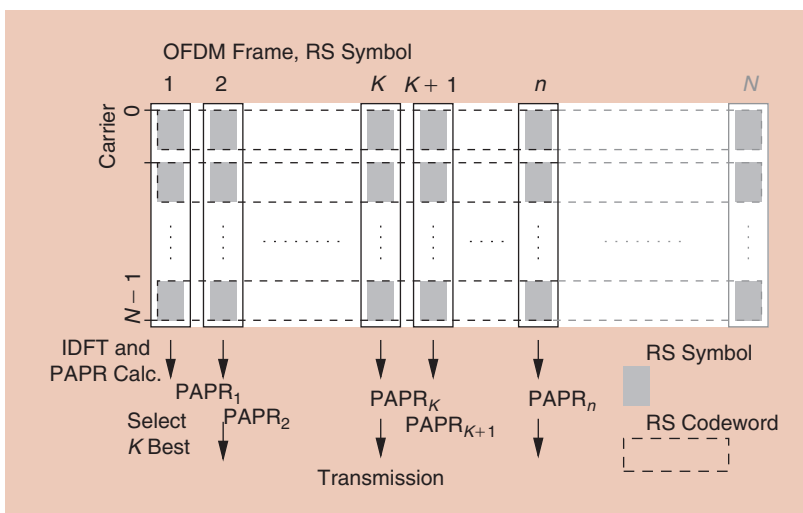
ALTERNATIVE ORTHONORMAL SYSTEMS: KASHIN AND TZAFRIRI'S THEOREM

In the section "Approaching the $\log(N)$ Barrier: Derandomization," it was shown that OFDM has unfavorable PAPR of order $\log(N)$ if N gets large. One might ask if this is an artifact of the underlying orthonormal signaling system. The answer is actually no, with the implication that OFDM plays no specific role among all orthonormal systems. In [64], it was already shown that the worst PAPR is of order N regardless of the signaling system, e.g., multicarrier code division multiplexing (CDM), etc. But even if we consider not the worst PAPR but look at the PAPR on average, the situation does not get better. In [65], Kashin and Tzafiriri proved that for any orthonormal system on a given finite time interval the expectation of PAPR is necessarily of order $\log(N)$. Again, changing the signaling is not beneficial in terms of PAPR. The underlying mathematical problem is that of estimating the supremum norm of a finite linear combination of functions weighted with random coefficients constrained in the energy norm.

IS PAPR OF SINGLE-CARRIER REALLY MUCH BETTER?

It is common engineering experience that single-carrier has better PAPR than multicarrier. But it might be worth raising this question again within the context of upcoming technological advances (such as LTE-A), which operate much closer to the Nyquist bandwidth and, moreover, use different modulation and coding schemes. Let us formalize this question.

Suppose, we send a transmit sequence C_1, \dots, C_N and use a band-limited filter to generate the continuous-time signal



[FIG7] An illustration of the coding method across symbols using off-the-shelf codes (such as RS codes).

(bandwidth is set to π for simplicity). The transmit signal can be described by

$$s(t) = \sum_{i=1}^N C_i \frac{\sin \pi(t - t_i)}{\pi(t - t_i)}$$

with sampling points $t_i \in \mathbb{Z}$. Naturally, band-limited signals of this form have very different PAPR behavior compared to OFDM since, obviously, if the coefficients are from some standard modulation alphabet, the signal is nailed down to some finite value at the sampling point independent of N . However, within the sampling intervals (on average) large PAPR could actually occur. Noteworthy, the worst case is growing without bounds linearly in N .

Surprisingly, the exact answer to this problem has not been explored until very recently [66], which is basically a result on large deviations in Banach spaces. It is proved in [66] that such bad PAPR cannot actually happen and that there is a constant $c_0 > 0$ such that

$$\mathbb{E}(\text{PAPR}) \leq c_0 \log \log(N).$$

But we also see the catch here. Modern communication systems use higher modulation sizes and, in that case, the influence of the data becomes dominant if the distribution becomes Gaussian-like. In that case we approach the $\log(N)$ again.

There is an interesting connection of the PAPR problem to the Hilbert transform context: since in many standard communication models, e.g., in Gabor's famous "Theory of Communication" [67], [68], the transmit signal is a linear combination of a signal and its Hilbert transform, properties such as PAPR in the transform domain become more and more important. Initiated by early works of Logan [69], who investigated the Hilbert transforms of certain bandpass signals, it was recognized not until very recently [70], [71] that the results are fragile for wideband signals containing spectral components in an interval around zero frequency. Then, in general, the domain of the Hilbert transform must be suitably extended; further, examples of band-limited wideband signals are provided where the PAPR

grows without bounds in the Hilbert transform domain [72]. Hence, for certain single-carrier analytic modulation schemes the transmit signal has to be shaped very carefully.

OVERCOMPLETE EXPANSIONS WITH UNIFORMLY BOUNDED PAPR

While the result for arbitrary orthonormal systems appears rather pessimistic, there is a possible solution in the form of frames. Frames are overcomplete systems of vectors in \mathbb{R}^n , $n < N$. Let us denote this description by $U = [u_1, \dots, u_N] \in \mathbb{R}^{n \times N}$, $N \geq n$. Then, if the rows are independent there is $x \in \mathbb{R}^N$ so that

$$y = U^T x \quad (10)$$

for any $y \in \mathbb{R}^n$ and the elements of U are a frame. If $U^T U = I_n$, where I_n is the identity matrix, then it is called a tight frame. In his seminal work, Kashin [8] interpreted the mapping (10) as an embedding of the Banach space with supremum norm l_∞ to the Banach space with standard Euclidean norm l_2 and asked for the growth factor $K(\lambda) > 0$, $\lambda := N/n$, between the two norms when the l_2 unit ball in \mathbb{R}^n should be covered by the unit ball l_∞ in \mathbb{R}^N . Such representations are called Kashin representations of level λ [73].

Clearly, if $N = n$ then $K(\lambda) = \sqrt{N}$. However, if $N > n$ (overcomplete expansion) then Kashin proved that there is a subspace in \mathbb{R}^N generated by a frame U such that $K(\lambda)$ is given by

$$K(\lambda) := c_1 \left(\frac{\lambda}{\lambda-1} \log \left(1 + \frac{\lambda}{\lambda-1} \right) \right)^{1/2}, \quad c_1 > 0.$$

Hence, the $K(\lambda)$ is uniformly bounded in n if $\lambda > 1$ is fixed. Good estimates of the constant $c_1 > 0$ are not known [73].

This intriguing result has been applied in the PAPR context in [73] and the implications for peak power reduction are immediate. The matrix U can be taken as a precoding matrix for classical OFDM transmission and achieve uniformly bounded PAPR. Unfortunately, the construction of the optimal subspace is not known [73]. Kashin representations exploiting the uncertainty principle of random partial Fourier matrices are presented in [74].

TONE RESERVATION AND SZEMERÉDI'S THEOREM

One of the oldest but still very popular schemes is tone reservation [75]. But, despite its simplicity, many questions remain unanswered, which does not come by coincidence: recent work in [76] has analyzed the performance of this method and uses an application of Szemerédi's famous theorem about arithmetic progressions.

Recalling the setting where a subset of subcarriers is solely reserved for peak power reduction, the challenge is to find for a given set of transmit sequence a subset and corresponding values such that the PAPR is reduced to the most possible gain. Until now, achievability and limits are not known (except for simple cases). Therefore, there is some incentive to look at this problem from a new perspective. Reference [76] has analyzed the case where the compensation set is arbitrary but fixed. In this typical case, it is proved that the efficiency of the system,

i.e., the ratio of cardinality of information and compensation sets, must decrease to zero with growing N if the peak power is constrained independent of the subcarriers. The proof shows that to maintain some strictly nonzero efficiency, a necessary condition is a relation between Banach spaces on multicarrier signals represented by the information set. This relation is shown not to hold asymptotically for sets with additive structure. However, Szemerédi's theorem states that such sets are included in every subset of cardinality δN where $\delta > 0$. In fact, such arithmetic progressions induce signals with bad PAPR behavior naturally to be excluded by the tone reservation method. The theorem shows that this is not possible.

In extended work [77], other families of orthogonal signaling such as Walsh sequences are also analyzed, where all of them showed basically the same discouraging result regarding the system's efficiency. This leads to the conjecture in [77] that all natural orthogonal signalling families have this behavior.

COMPRESSED SENSING

Compressed sensing [6], [7] is a new sampling method that compresses a signal simultaneously with data acquisition. Each element of the compressed signal or measurements consists of a linear combination of the elements in the original signal and this linear transformation is independent of instantaneous characteristics of each signal. In general, it is not possible to recover an unknown original signal from the measurements in the reduced dimension. Nevertheless, if the original signal has sparsity property, its recovery can be perfectly achieved at the receiver. Since sparsity frequently appears in the PAPR problem, compressed sensing can be a powerful tool to solve these problems.

Compressed sensing can be regarded as minimizing the number of measurements while still retaining the information necessary to recover the original signal well (i.e., beyond classical Nyquist sampling). The process can be briefly illustrated as follows. Let f denote a signal vector of dimension N and g be a measurement vector of dimension M with $M < N$ obtained by $g = \Phi f$, where Φ is called sensing matrix. At the transmitter, sampling and compression are performed altogether by simply multiplying Φ by f to obtain g . At the receiver, if f is an S -sparse signal, which means f has no more than S nonzero elements, it is shown in [7] that the exact f can be obtained from g by using l_1 minimization, that is,

$$\min_{\tilde{f}} \|\tilde{f}\|_1 \quad \text{subject to} \quad g = \Phi \tilde{f} \quad (11)$$

as long as Φ has some good property, which is called restricted isometry property (RIP). For some positive integer S , the isometry constant δ_S of a matrix Φ is defined as the smallest number such that

$$(1 - \delta_S) \|f\|_2^2 \leq \|\Phi f\|_2^2 \leq (1 + \delta_S) \|f\|_2^2$$

holds for all S -sparse vectors f . Under RIP with $\delta_{2S} < \sqrt{2} - 1$, (11) gives the exact solution for f [78]. This recovery method using l_1 -minimization is called basis pursuit (BP) [79], which

requires high computational complexity. Many greedy algorithms [80]–[83] have been developed to reduce the recovery complexity.

In many applications of compressed sensing such as the PAPR problem, it is required to recover f from the corrupted measurements $g' = g + z$, where z is a noise vector of dimension M . For this, recovery algorithms such as basis pursuit denoising [79], Lasso [84], and their variants have been developed while the existing recovery algorithms can also be used. However, these algorithms do not still show good performance enough to be adopted in wireless communication systems, which usually require very low error rate even in severely noisy environments.

Related to PAPR problems, the elements of compressed sensing such as sparsity, RIP, and recovery algorithms can be utilized in many PAPR reduction schemes. In [74] and [85], a new tone-reservation scheme is proposed, which is different from the existing tone reservation [4] in that it provides a guaranteed upper bound for PAPR reduction as well as guaranteed rates of convergence. This scheme exploits the RIP of the partial DFT matrix. In [86], a novel convex optimization approach is proposed to numerically determine the near-optimal tone-injection solution. Generally, tone injection [4] is an effective approach to mitigate PAPR problem without incurring bandwidth loss. However, due to its computational complexity, finding the optimal tone injection becomes intractable for OFDM systems with a large number of subcarriers. Therefore, a semidefinite relaxation needs to be adopted in the convex optimization [87]. Moreover, based on the observation that only a small number of subcarrier symbols are usually moved, l_0 minimization is required and naturally it can be relaxed to l_1 minimization similar to compressed sensing literature.

One of the popular solutions to PAPR reduction is clipping the amplitude of the OFDM signal although the clipping increases the noise level by inducing a clipping noise. Due to the sparsity of the clipping noise, compressed sensing can be used to recover and cancel the clipping noise. Before the clipping noise cancellation schemes using compressed sensing appear, some foundations of them have been presented. An impulse noise cancellation scheme using sparse recovery is firstly proposed in [88]. In practical systems, there exists a set of null tones not used for information transmission, which is exploited as measurements to estimate the impulse noise in time domain at the receiver. As an extension to [88], an alternative recovery algorithm with low complexity is proposed in [89], which exploits the structure of DFT matrix and available a priori information jointly for sparse signal recovery. In [90], the work in [88] is extended to the case of bursty impulse noise whose recovery is based on the application of block-based compressed sensing. Second, a clipping noise cancellation scheme using frame theory is proposed in [91]. Although this scheme uses not compressed sensing but frame expansion, the frame expansion can be viewed as a special case of compressed sensing problem with known positions of nonzero elements. Some additional reserved tones not including data are padded and they are used as the measurements to recover the clipping noise at the receiver.

Motivated by the above works, clipping noise cancellation schemes using compressed sensing have been proposed in [92] and its extended version in [93]. In [92] and [93], M reserved tones are allocated before clipping at the transmitter and they cause some data rate loss. These reserved tones can be exploited as measurements instead of null tones in [88]–[90] and [94]. Let us denote the transceiver model in frequency domain with clipping noise as

$$Y = H(C + D) + Z, \quad (12)$$

where C and Y are $N \times 1$ transmitted and received tone vectors, respectively, H is a diagonal matrix of the channel frequency response, D is $N \times 1$ clipping noise vector, and Z is AWGN vector. Starting from (12), we equalize the channel by multiplying with H^{-1} and select the rows whose indices correspond to locations of the reserved tones by multiplying with a $M \times N$ row selection matrix S_r . This results in

$$\underbrace{S_r H^{-1} Y}_g = \underbrace{S_r F D}_\Phi + \underbrace{S_r H^{-1} Z}_f, \quad (13)$$

where F is the DFT matrix and $D = Fd$. As seen in (13), the clipping noise on the reserved tones is used as measurements to recover the clipping noise d in time domain by sparse recovery algorithm. Additionally, in [93], a method exploiting a-priori information together with weighted l_1 minimization for enhanced recovery followed by Bayesian techniques is proposed. However, the performance of [92] and [93] is restricted due to weakness of the compressed sensing against noise.

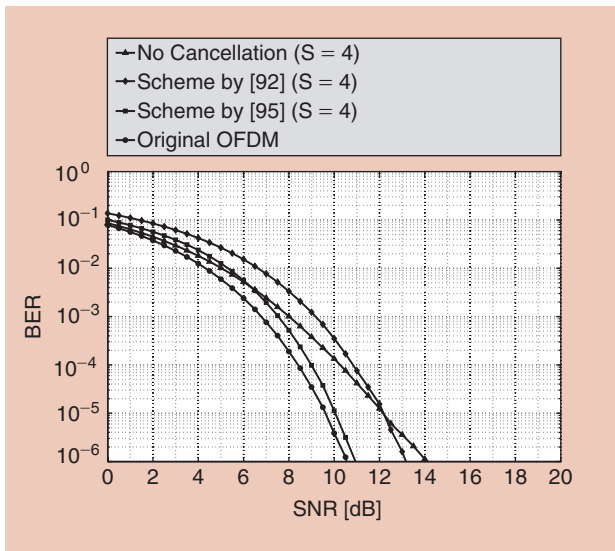
In [95], more enhanced clipping noise cancellation scheme using compressed sensing is proposed. Different from [92] and [93], the scheme in [95] does not cause data rate loss, because it exploits the clipping noise in frequency domain as measurements underlying in the data tones rather than the reserved tones. In this case, transmitted data and clipping noise are mixed in the data tones. To distinguish the clipping noise from the data tones well, this scheme exploits part of the received data tones with high reliability. To (12), we multiply H^{-1} and row selection matrix S_d , selecting the locations of reliable data tones, as

$$S_d H^{-1} Y = S_d C = S_d D + S_d H^{-1} Z. \quad (14)$$

Then, we estimate the $S_d \hat{C}$ and subtract them from (14) as

$$\underbrace{S_d H^{-1} Y - S_d \hat{C}}_g = \underbrace{S_d F d}_\Phi + \underbrace{S_d (C - \hat{C})}_f + S_d H^{-1} Z. \quad (15)$$

Then, from partially extracted clipping noise in frequency domain, we can recover the clipping noise d in time domain via sparse recovery algorithms. Furthermore, this scheme can adjust the number of the measurements M by changing the reliability of received data. Therefore, when there is AWGN noise, we can select the optimal number of measurements corresponding to the noise amount. Consequently, this scheme successfully realizes the clipping noise cancellation scheme by overcoming weakness of the compressed sensing against noise. Additionally, in [95], clipping noise cancellation for OFDMA access (OFDMA) systems is also proposed using compressed



[FIG8] The BER performance in OFDM AWGN channel using compressed sensing based clipping noise cancellation schemes as described in [92] and [95]. In [95], 23 out of 64 QPSK subcarriers are used for recovery of the $S = 4$ sparse clipping noise. In [92], 29 tones are selected by a (59, 29, 14) difference set and reserved. Baseline performance is the original OFDM signal with no clipping. It can be seen that provided the AWGN is not too strong, sparsity in the clipping noise can lead to significant performance gains of more than 3 dB given an error probability of 10^{-6} .

sensing. The fast Fourier transform (FFT) block of OFDM systems can be decomposed into the small FFT blocks. And, the subset of rows in the small sized DFT matrix can also be used as a sensing matrix, which can be used to recover the clipping noise for OFDMA systems via sparse recovery algorithm.

Figure 8 shows the bit error rate (BER) over SNR performance of the clipping noise cancellation schemes based on compressed sensing described in [92] and [95] for OFDM signals over the AWGN channel. The S -sparse clipping noise signal contaminates the original OFDM signal and the case of no clipping noise cancellation shows the worst BER performance among all schemes. In [92], the authors applied the compressed sensing technique to OFDM systems for the first time, but there is a benefit only for the high SNR region due to weakness of compressed sensing recovery against AWGN. The BER performance of the scheme in [95] is better because the number of the measurements can be adjusted corresponding to the AWGN level.

CONCLUSIONS

Despite two decades of intensive research, the PAPR problem remains one of the major problems in multicarrier theory with huge practical impact. This article provides a fresh look on this problem by outlining a new perspective using alternative metrics (including MIMO and multiuser systems as a special case), the corresponding theoretical foundations and related designs. This is followed by thorough discussion of current limits and new future directions.

ACKNOWLEDGMENTS

The work of Gerhard Wunder was supported by the Deutsche Forschungsgemeinschaft under grant WU 598/3-1. The work of Jong-Seon No was supported by the National Research Foundation of Korea grant 2012-0000186 funded by the Korean government.

AUTHORS

Gerhard Wunder (gerhard.wunder@hhi.fraunhofer.de) studied electrical engineering at the University of Hannover and the Technical University (TU) Berlin. He received the graduate degree in electrical engineering (Dipl.-Ing.) with highest honors in 1999 and the Ph.D. degree (Dr.-Ing.) in communication engineering on the PAPR problem in OFDM with distinction (summa cum laude) in 2003 from TU Berlin. In 2007, he also received the habilitation degree and became a Privatdozent at TU Berlin. Since 2003, he has been the head of a research group at the Fraunhofer Lab for Mobile Communications/Heinrich-Hertz-Institut and the TU Berlin. In 2000 and 2005, he was a visiting professor at the Georgia Institute of Technology and Stanford University. In 2009 he was a consultant at Alcatel-Lucent Bell Labs, New Jersey. He has been an editor of *IEEE Transactions on Wireless Communications* since 2011. In 2011, he received the award for outstanding scientific publication by the German Communication Engineering Society (ITG).

Robert F.H. Fischer (robert.fischer@uni-ulm.de) received the Dr.-Ing. degree (1996) and the habilitation degree (2001) from the University of Erlangen-Nürnberg, Germany. In 1997, he was with the IBM Research Laboratory, Zürich, Switzerland, and in 2005 he spent a sabbatical at ETH, Zürich, Switzerland. Since 2011, he has been full professor at the University of Ulm, Germany. His current interests are information theory, coded modulation, digital communications, signal processing, and especially precoding and shaping techniques for high-rate transmission schemes. He received the 1997 Dissertation Award from the Technische Fakultät, University of Erlangen-Nürnberg, the 2000 Publication Award of the German Society of Information Techniques (ITG), the 2002 Wolfgang Finkelburg Habilitation Award, and the Philipp-Reis-Preis in 2005. He authored *Precoding and Signal Shaping for Digital Transmission* (Wiley, 2002). He is a Senior Member of the IEEE.

Holger Boche (boche@tum.de) received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the Technische Universität Dresden, Germany, in 1990 and 1994, respectively. He graduated with a degree in mathematics from the Technische Universität Dresden in 1992. He received his Dr. rer. nat. degree in pure mathematics from the Technische Universität Berlin, Germany, in 1998. Since 2002, he has been a full professor of mobile communication networks with the Institute for Communications Systems, Technische Universität Berlin. Since 2004 he has also been director of the Fraunhofer Institute for Telecommunications (HHI), Berlin, Germany. He has been with the Institute of Theoretical Information Technology and full professor at the Technical University of Munich, Germany, since 2010. He was corecipient of the 2006 IEEE Signal

Processing Society (SPS) Best Paper Award and recipient of the 2007 IEEE SPS Best Paper Award. He is a Fellow of the IEEE.

Simon Litsyn (litsyn@eng.tau.ac.il) received the Ph.D. degree from the Leningrad Electrotechnical Institute in 1982. Since 1991, he has been with the School of Electrical Engineering of the Tel Aviv University, where he is currently a professor. He authored *Covering Codes* (Elsevier, 1997) and *Peak Power Control in Multicarrier Communications* (Cambridge University Press, 2007), more than 200 papers in profile journals, and 30 patents. He was an associate editor of *Coding Theory, IEEE Transactions on Information Theory* from 2000 to 2003. He is an associate editor of *Advances in Mathematics and Communications* and *Applicable Algebra in Engineering, Communications and Computing*. From 2000 to 2012, was a chief scientist and engineering fellow of SanDisk. His research interests include algorithmic methods in communications and storage and applications of discrete mathematics.

Jong-Seon No (jsno@snu.ac.kr) received the B.S. and M.S.E.E. degrees in electronics engineering from Seoul National University, Seoul, Korea, in 1981 and 1984, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1988. He was a senior member of technical staff at Hughes Network Systems from 1988 to 1990. He was also an associate professor in the Department of Electronic Engineering, Konkuk University, Seoul, from 1990 to 1999. In 1999, he joined the faculty of the Department of Electrical and Computer Engineering, Seoul National University, where he is currently a professor. He was the general cochair of the 2009 International Symposium on Information Theory, COEX, Seoul, Korea. His area of research interests includes error-correcting codes, sequences, cryptography, interference alignment, LDPC codes, compressed sensing, and wireless communication systems. He is a Fellow of the IEEE.

REFERENCES

- [1] L. M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gódor, G. Auer, and L. Van der Perre, "Challenges and enabling technologies for energy aware mobile radio networks," *IEEE Signal Process. Mag.*, vol. 48, no. 11, pp. 66–72, Nov. 2010.
- [2] S. Litsyn, *Peak Power Control in Multicarrier Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
- [3] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 56–65, Apr. 2005.
- [4] J. Tellado, "Peak to average power reduction for multicarrier modulation," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1999.
- [5] G. Wunder, "Analysis of alternative metrics for the PAPR problem in OFDM transmission," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Peterburg, Russia, Aug. 2011, pp. 479–483.
- [6] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [7] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [8] B. Kashin, "Diameters of some finite-dimensional sets and classes of smooth functions," *Izvestia*, vol. 41, no. 2, pp. 334–351, 1977.
- [9] A. M. Pinkus, *n-Widths in Approximation Theory*, vol. 7. Berlin, Germany: Springer-Verlag, 1985.
- [10] G. Wunder and H. Boche, "Peak value estimation of band-limited signals from its samples, noise enhancement and a local characterization in the neighborhood of an extremum," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 771–780, Mar. 2003.
- [11] I. Kotzer, S. Har-Nevo, S. Sodin, and S. Litsyn, "An analytical approach to the calculation of EVM in clipped multi-carrier signals," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1371–1380, May 2012.
- [12] I. Kotzer and S. Litsyn, "On the EVM computation of arbitrary clipped multi-carrier signals," *EURASIP J. Advances Signal Process.*, vol. 36, 2011.
- [13] F. H. Raab, P. Asbeck, S. Crippsa, P. B. Kenington, Z. B. Popovic, N. Pothecary, J. F. Sevic, and N. O. Sokal, "Power amplifiers and transmitters for RF and microwave," *IEEE Trans. Microw. Theory Tech.*, vol. 53, no. 3, pp. 814–826, 2002.
- [14] G. Fettweis, M. Loehning, D. Petrovic, M. Windisch, P. Zillmann, and W. Rave, "Dirty RF: A new paradigm," *Int. J. Wireless Inform. Netw.*, vol. 14, no. 2, pp. 133–148, June 2007.
- [15] H. Ochiai and H. Imai, "Performance analysis of deliberately clipped OFDM signals," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 89–101, Jan. 2002.
- [16] Q. Hua, R. Raich, and G. T. Zhou, "On the benefits of deliberately introduced baseband nonlinearities in communication systems," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Process.*, 2004, pp. 905–908.
- [17] 3GPP TSG RAN WG1 37, *Comparison of PAR and Cubic Metric for Power De-Rating*, Motorola Tdoc R1-040642, May 2004.
- [18] 3GPP, "User equipment radio transmission and reception (fdd)," in *3GPP TS 25.101*. Mar. 2009.
- [19] C. Siegl, "Peak-to-average power ratio reduction in multi-antenna OFDM via multiple signal representation," *Ph.D. dissertation*, Erlanger Berichte aus Informations- und Kommunikationstechnik, Band 29, Shaker Verlag, Aachen, Germany, 2010.
- [20] C. Siegl and R. F. H. Fischer, "Out-of-band power reduction using selected mapping with power-amplifier-oriented metrics," in *Proc. 14th Int. OFDM Workshop*, Hamburg, Germany, Sept. 2009.
- [21] F. Peng and W. E. Ryan, "On the capacity of clipped OFDM channels," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Seattle, WA, July 2006, pp. 1866–1870.
- [22] R. F. H. Fischer, "Capacity of clipped 4-QAM-OFDM," in *Proc. 8th Int. ITG Conf. Source Channel Coding*, Siegen, Germany, Jan. 2010.
- [23] G. Wunder and J. Schreck, "A PAPR algorithm that achieves asymptotically zero error probability with backoff below $\log(N)/2$," in *Proc. 15th Int. OFDM Workshop*, Hamburg, Germany, Sept. 2010.
- [24] Y. K. Belyaev and V. P. Nosko, "Characteristics of excursions above a high level for a Gaussian process and its envelope," *Theory Probability Its Applcat.*, vol. XIV, no. 2, pp. 296–309, 1969.
- [25] W. Henkel, V. Azis, and S. Trautmann, "The analytic treatment of the error probability due to clipping: A solved problem," in *Proc. Int. Symp. Inform. Theory*, 2004, pp. 1–5.
- [26] H. Ochiai and H. Imai, "On the distribution of the peak-to-average power ratio in OFDM signals," *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 282–289, Feb. 2001.
- [27] N. Dinur and D. Wulich, "Peak-to average power ratio in high-order OFDM," *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 1063–1072, June 2001.
- [28] G. Wunder and H. Boche, "Upper bounds on the statistical distribution of the crest-factor in OFDM transmission," *IEEE Trans. Inform. Theory*, vol. 49, no. 2, pp. 488–494, Feb. 2003.
- [29] S. Litsyn and G. Wunder, "Generalized bounds on the CF distribution of OFDM signals with application to code design," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 992–1006, Mar. 2006.
- [30] I. Sason, "On the concentration of the crest factor for OFDM signals," in *Proc. 8th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aachen, Germany, Nov. 2011, pp. 784–788.
- [31] R. Bäuml, R. F. H. Fischer, and J. Huber, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping," *Electron. Lett.*, vol. 32, no. 22, pp. 2056–2057, Oct. 1996.
- [32] S. Müller and J. Huber, "OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences," *Electron. Lett.*, vol. 33, no. 5, pp. 368–369, Feb. 1997.
- [33] A. Alavi, C. Tellambura, and I. Fair, "PAPR reduction of OFDM signals using partial transmit sequence: An optimal approach using sphere decoding," *IEEE Commun. Lett.*, vol. 9, no. 11, pp. 982–984, Nov. 2005.
- [34] S. Litsyn and A. Shpunt, "A balancing method for PMEPR reduction in OFDM signals," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 683–691, Apr. 2007.
- [35] T. T. Nguyen and L. Lampe, "On Trellis shaping for par reduction in OFDM systems," *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1678–1682, Sept. 2007.
- [36] K.-U. Schmidt, "Bounds on the PMEPR of translates of binary codes," *IEEE Commun. Lett.*, vol. 14, no. 11, pp. 1059–1061, Nov. 2010.
- [37] B. S. Krongold and D. L. Jones, "PAR reduction in OFDM via active constellation extension," *IEEE Trans. Broadcasting*, vol. 49, no. 3, pp. 525–528, 2003.
- [38] M. Sharif and B. Hassibi, "Existence of codes with constant PMEPR and related design," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2836–2846, 2004.
- [39] M. Sharif, V. Tarokh, and B. Hassibi, "Peak power reduction of OFDM signals with sign adjustment," *IEEE Trans. Commun.*, vol. 57, no. 2, pp. 2160–2166, July 2009.
- [40] J. Spencer, "Ten lectures on the probabilistic method," in *Proc. CBMS-NSF Regional Conf. Series*, 1994.
- [41] L. Wang and C. Tellambura, "Cross-entropy-based sign-selection algorithms for peak-to-average power ratio reduction of OFDM systems," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4990–4994, Oct. 2008.

- [42] M. G. Damavandi and A. Abbasfar, "Upper bounds on achievable PMEPR through subcarrier sign selection," in *Proc. 73rd IEEE Veh. Technol. Conf. (VTC Spring)*, May 2011, pp. 1–5.
- [43] M. Sharif and B. Hassibi, "High rate codes with bounded PMEPR for BPSK and other symmetric constellations," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 71–78, July 2006.
- [44] M.-S. Baek, M.-J. Kim, Y.-H. You, and H.-K. Song, "Semi-blind channel estimation and par reduction for MIMO-OFDM system with multiple antennas," *IEEE Trans. Broadcasting*, vol. 50, no. 4, pp. 414–424, Dec. 2004.
- [45] R. F. H. Fischer and C. Siegl, "PAPR reduction in single- and multi-antenna OFDM via directed selected mappings," *IEEE Trans. Commun.*, vol. 11, no. 11, pp. 3205–3208, Nov. 2009.
- [46] C. Siegl and R. F. H. Fischer, "Asymptotic performance analysis and successive selected mapping for PAR reduction in OFDM," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3228–3237, Apr. 2010.
- [47] R. F. H. Fischer and C. Siegl, "Probabilistic signal shaping in MIMO OFDM via successive candidate generation," *Eur. Trans. Telecommun.*, vol. 21, no. 7, pp. 648–654, Nov. 2010.
- [48] C. Siegl and R. F. H. Fischer, "Selected basis for par reduction in multi-user downlink scenarios using lattice-reduction-aided precoding," *EURASIP J. Advances Signal Process.*, July 2011.
- [49] J. G. Smith, "The information capacity of amplitude and variance constraint scalar Gaussian channels," *Inform. Control*, vol. 18, pp. 203–219, 1971.
- [50] S. Shamai and I. Bar-David, "The capacity of average and peak-powerlimited quadrature Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1060–1071, July 1995.
- [51] C. Siegl, R. F. H. Fischer, and E. Mustafin, "Active constellation extension for increasing the capacity of clipped OFDM," in *Proc. 15th Int. OFDM Workshop*, Hamburg, Germany, Sept. 2010.
- [52] A. E. Jones, T. A. Wilkinson, and S. K. Barton, "Block coding scheme for reduction of peak-to-mean envelope power ratio of multicarrier transmission schemes," *Electron. Lett.*, vol. 30, no. 25, pp. 2098–2099, 1994.
- [53] J. A. Davis and J. Jedwab, "Peak-to-mean power control in OFDM, Golay complementary sequences, and Reed-Muller codes," *IEEE Trans. Inform. Theory*, vol. 45, no. 7, pp. 2397–2417, Nov. 1999.
- [54] K.-U. Schmidt, "On cosets of the generalized first-order Reed-Muller code with low PMEPR," *IEEE Trans. Inform. Theory*, vol. 52, no. 7, pp. 3220–3232, July 2006.
- [55] K. G. Paterson and V. Tarokh, "On the existence and construction of good codes with low peak-to-average power ratios," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 1974–1986, Sept. 2000.
- [56] R. F. H. Fischer and C. Siegl, "Reed-Solomon and simplex codes for PAPR reduction in OFDM," *IEEE Trans. Inform. Theory*, vol. 55, no. 4, pp. 1519–1528, Apr. 2009.
- [57] J. Spencer, "Six standard deviations suffice," *Trans. AMS*, vol. 289, no. 2, pp. 679–706, 1985.
- [58] J. A. Davis, J. Jedwab, and K. G. Paterson, *Codes, Correlations and Power Control in OFDM*, vol. 542. Norwell, MA: Kluwer, 1999, pp. 85–112.
- [59] G. Wunder and H. Boche, "A baseband model for computing the PAPR in OFDM systems," in *Proc. 4th Int. ITG Conf. Source Channel Coding*, Berlin, Jan. 2002, pp. 273–280.
- [60] V. Tarokh and H. Jafarkhani, "On the computation and reduction of the peak-to-average power ratio in multicarrier communications," *IEEE Trans. Commun.*, vol. 48, no. 1, pp. 37–44, Jan. 2000.
- [61] H.-K.-C. Kwok and D. L. Jones, "Par reduction for Hadamard transform based OFDM," in *Proc. Conf. Inform. Sci. Syst.*, Princeton, NJ, Mar. 2000.
- [62] A. Mobasher and A. K. Khandani, "Integer-based constellation-shaping method for PAPR reduction in OFDM systems," *IEEE Trans. Commun.*, vol. 54, no. 1, pp. 119–127, Jan. 2006.
- [63] C. Eugen, D. Sterian, K. Wesolowski, Y. Wu, and M. Patzold, "Reducing the peak and average power for OFDM systems using QAM by constellation shaping," *Eur. Trans. Telecommun.*, vol. 21, no. 1, pp. 35–49, Jan. 2010.
- [64] H. Boche and V. Pohl, "Signal representation and approximation: Fundamental limits," *Eur. Trans. Telecommun.*, vol. 18, no. 5, pp. 445–456, Aug. 2007.
- [65] B. Kashin and L. Tzafiriri, "Lower estimates for the supremum of some random processes, 2," *East J. Approximations*, vol. 1, no. 3, pp. 373–377, 1995.
- [66] H. Boche, B. Farrell, M. Ledoux, and M. Wiese, "Expected supremum of a random linear combination of shifted kernels," *J. Fourier Anal. Appl.*, vol. 18, no. 4, pp. 790–802, 2012.
- [67] D. Gabor, "Theory of communication," *J. Institute Electr. Engineers*, vol. 93, no. 3, pp. 429–457, Nov. 1946.
- [68] H. Boche and U. J. Mönich, "Solution of the peak value problem for Gabor's theory of communication," in *Proc. IEEE Int. Symp. Signals Syst. Electron.*, Oct. 2012.
- [69] B. F. Logan, Jr., "Theory of analytic modulation systems," *Bell Syst. Tech. J.*, vol. 57, no. 3, pp. 491–576, Mar. 1978.
- [70] H. Boche and U. Mönich, "On the Hilbert transform of bounded band-limited signals," in *Problems of Information Transmission*, vol. 48, no. 3, Berlin, Germany: Springer-Verlag, 2012, pp. 217–238.
- [71] H. Boche and U. Mönich, "Extension of the Hilbert transform," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 3697–3700.
- [72] H. Boche and U. Mönich, "Complete characterization of band limited signals with bounded Hilbert transform," in *Proc. 20th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2012, pp. 1895–1899.
- [73] B. Farrell and P. Jung, "A Kashin approach to the capacity of the discrete amplitude constrained Gaussian channel," in *Proc. Int. Conf. Sampling Theory Applicat. (SampTA)*, 2009, vol. 3.
- [74] J. Ilic and T. Strohmer, "PAPR reduction in OFDM using Kashin's representation," in *Proc. IEEE 10th Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, Perugia, Italy, 2009, pp. 444–448.
- [75] Y. Wang, W. Chen, and C. Tellambura, "Genetic algorithm based nearly optimal peak reduction tone set selection for adaptive amplitude clipping PAPR reduction," *IEEE Trans. Broadcasting*, vol. 58, no. 3, pp. 462–471, Sept. 2012.
- [76] H. Boche and B. Farrell, "PAPR and the density of information bearing signals in OFDM," *EURASIP J. Adv. Signal Process.*, vol. 2011, 2011.
- [77] H. Boche and B. Farrell, "On the peak-to-average power ratio reduction problem for orthogonal transmission schemes," *Internet Math.*, vol. 9, no. 2–3, pp. 265–296, 2013.
- [78] E. J. Candes and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [79] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.
- [80] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [81] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *IEEE J. Sel. Signal Process.*, vol. 4, no. 2, pp. 310–316, Apr. 2010.
- [82] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Computat. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.
- [83] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inform. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.
- [84] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. Royal Statist. Soc. Series B*, vol. 58, no. 1, pp. 267–288, 1996.
- [85] J. Ilic and T. Strohmer, "OFDM, PAPR and RIP," in *Proc. Inform. Theory Applicat. Workshop (ITA)*, San Diego, CA, Feb. 2010.
- [86] Z. Ding and N. Jacklin, "Sparse signal reconstruction approach to tone injection for OFDM PAPR reduction," in *Proc. Inform. Theory Applicat. Workshop (ITA)*, La Jolla, CA, Feb. 2011.
- [87] S. Boyd and S. Vandenberghe, *Convex Optimization*. Cambridge, MA: Cambridge Univ. Press, 2004.
- [88] G. Caire, T. Y. Al-Naffouri, and A. K. Narayanan, "Impulse noise cancellation in OFDM: An application of compressed sensing," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, July 2008, pp. 1293–1297.
- [89] T. Y. Al-Naffouri, A. A. Quadeer, and G. Caire, "Impulsive noise estimation and cancellation in DSL using orthogonal clustering," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Aug. 2011, pp. 2841–2845.
- [90] L. Lampe, "Bursty impulse noise detection by compressed sensing," in *Proc. IEEE Int. Symp. Power Line Commun. Its Applicat. (ISPLC)*, Apr. 2011, pp. 29–34.
- [91] L. Valbonesi and R. Ansari, "Frame-based approach for peak-to-average power ratio reduction in OFDM," *IEEE Trans. Commun.*, vol. 54, no. 9, pp. 1604–1613, Sept. 2006.
- [92] E. B. Al-Safadi and T. Y. Al-Naffouri, "On reducing the complexity of tone-representation PAPR reduction techniques by compressive sensing," in *Proc. IEEE Globecom*, Honolulu, HI, Dec. 2009, pp. 1–6.
- [93] E. B. Al-Safadi and T. Y. Al-Naffouri, (2011, Jan.). Peak reduction and clipping mitigation by compressive sensing. *IEEE Trans. Signal Process.* [Online]. Available: <http://arxiv.org/pdf/1101.4335v1.pdf>
- [94] M. Mohammadnia-Avval, A. Ghassemi, and L. Lampe, "Compressive sensing recovery of nonlinearly distorted OFDM signals," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kyoto, Japan, June 2011, pp. 1–5.
- [95] K.-H. Kim, H. Park, J.-S. No, and H. Chung, (2011, Oct.). Clipping Noise Cancellation for OFDM and OFDMA Systems Using Compressed Sensing [Online]. Available: <http://arxiv.org/pdf/1110.4174.pdf>