On Reducing the Complexity of Tone-Reservation Based PAPR Reduction Schemes by Compressive Sensing

Eprahim B. Al-Safadi
Department of Electrical Engineering
King Fahd University of Petroleum and Minerals
P.O. Box 7443 Dhahran 31261, Saudi Arabia

Tareq Y. Al-Naffouri
Department of Electrical Engineering
King Fahd University of Petroleum and Minerals
Dhahran, Saudi Arabia

Abstract—In this paper, we describe a novel design of a Peak-to-Average-Power-Ratio (PAPR) reducing system, which exploits the relative temporal sparsity of Orthogonal Frequency Division Multiplexed (OFDM) signals to detect the positions and amplitudes of clipped peaks, by partial observation of their frequency content at the receiver. This approach uses recent advances in reconstruction of sparse signals from rank-deficient projections using convex programming collectively known as compressive sensing.

Since previous work in the literature has focused on using the reserved tones as spectral support for optimum peak-reducing signals in the time-domain [5], the complexity at the transmitter was always a problem. In this work, we alternatively use extremely simple peak-reducing signals at the transmitter, then use the reserved tones to detect the peak-reducing signal at the receiver by a convex relaxation of an other-wise combinatorially prohibitive optimization problem.

This in effect completely shifts the complexity to the receiver and drastically reduces it from a function of N (the number of subcarriers in the OFDM signal), to a function of m (the number of reserved tones) which is a small subset of N.

I. INTRODUCTION

In the last decade, the problem of high peak-to-average-power ratio (PAPR) in OFDM systems has been tackled by a variety of approaches, including coding techniques, constellation reshaping, tone-reservation, and selective mapping, to name a few. Although many of these reduction techniques are brilliant and extremely effective, the main obstacle limiting their actual implementation is commonly related to high complexity.

Our main contribution in this paper is to demonstrate how the complexity of PAPR reduction techniques based on tone-reservation in particular can be drastically reduced using compressive sensing, at the cost of suboptimal performance of the PAPR reduction algorithm.

After establishing the main technique, we describe and analyze the performance of a couple of peak reducing schemes that are suitable for this design, then compare the results with both suggested and previous attempts of PAPR reduction techniques based on tone-reservation. The complexity of these variants is stressed throughout, and brief indications of possible further improvements being investigated by the authors are also given.

Since an OFDM signal is typically constructed by the superposition of a large number of modulated waveforms, its envelope fluctuates with significant variance, causing the high PAPR. The highest of these temporal peaks can be modeled as extreme events such that, given an appropriate limiting threshold γ, or a fixed number of ordered statistics s, a vector registering the positions and values of such peaks would be controllably sparse.

When these rare events occur and clipping is used, such a nonlinear transformation of a random process results in frequency distortion that spreads un-evenly over the subcarriers. Now if a sufficient subset of these tones were reserved at the transmitter, then using appropriate techniques of compressive sensing, the receiver can estimate the entire temporal sparse vector containing the locations and amplitudes of the clipped parts of the OFDM signal by observing the frequency distortion on this subset contaminated by noise.

It is worth mentioning that the objective of this paper is to establish the serious potential for compressive sensing techniques in PAPR reduction in particular. Interesting approaches to fine-tune the basic algorithm such as optimizing the reserved tone-set and clipping criteria, exploiting auto-covariance for oversampled signals, and angle steering techniques all require rigorous analysis for improving the estimation process at the receiver and reducing the error rate with minimum complexity. These are being pursued by the authors and the details are avoided here for lack of space.

II. TRANSMISSION MODEL

We define the time-domain complex base-band channel model as

$$y_k = \sum_{i=0}^{L-1} h_l x_{k-l} + z_k \tag{1}$$

where $\{x_k\}$ and $\{y_k\}$ denote the channel input and output, $h = (h_0, h_1, \ldots, h_{L-1})$ is the impulse response of the channel, and the noise model $\{z_k\}$ is assumed a circularly symmetric Gaussian random variable with variance N_0. To compactly notate the OFDM system and the operations to come, we

express the channel model in matrix form $y = Hx + z$ where y and x are the time-domain OFDM receive and transmit signal blocks (after cyclic prefix removal) and $z \sim \mathcal{CN}(0, N_0 I)$.

Due to the presence of the cyclic prefix, H is a circulant matrix describing the cyclic convolution of the channel impulse response with the block x and can be decomposed as $H = F^H DF$ where the superscript H denotes Hermitian conjugation, F denotes a unitary Discrete Fourier Transform (DFT) matrix with (k, l) element

$$[F]_{k,l} = \frac{1}{\sqrt{N}} e^{\frac{-j2\pi kl}{N}}, \quad k, l \in 0, 1, \ldots, N-1$$

$D = diag(\check{h})$ and $\check{h} = \sqrt{N}Fh$ is the DFT of the channel impulse response (whose coefficients are found, by construction, on the first column of H).

III. PAPR REDUCTION DESIGN

The time-domain OFDM signal x is typically constructed by taking the IDFT of the data vector $\check{d}$ whose entries are drawn from a QAM constellation. Since this signal is typically of very high PAPR, we add a peak-reducing signal c of arbitrary spectral support at the transmitter and then estimate it and subtract it from the demodulated signal at the receiver.

In what follows, the main condition we impose on c is that it be sparse in time. This is basically the case if we set a clipping threshold on the envelope of the OFDM symbols or, better yet for the performance of this algorithm, if the transmitter were to clip the highest s peaks.

Denoting Ω as the set of frequencies in an OFDM signal of cardinality $|N|$, let $\Omega_d \subset \Omega$ be the set of frequencies that are used for data transmission and $\Omega_c = \Omega \backslash \Omega_d$ the complementary set reserved for measurement tones of cardinality $|\Omega_c| = m$. Note that for compressive sensing purposes, a near optimal strategy is to use a random assignment of tones for estimating c [3]. However, based on results in [11] it was found in [6] that using difference sets, one is able to boost the performance of the recovery algorithm and reduce the symbol error rate.

The data symbols $\check{d}_i$ are drawn from a QAM constellation of size M and are supported by Ω_d of cardinality $|\Omega_d| = N - m = k$. Consequently, the transmitted peak-reduced time-domain signal is

$$\bar{x} = x + c = F^H S_x \check{d} + c \tag{2}$$

where S_x is an $N \times k$ selection matrix containing only one element equal to 1 per column, and with $m = N - k$ zero rows. The columns of S_x index the subcarriers that are used for data transmission in the OFDM system. Similarly, we denote by S the $N \times m$ matrix with a single element equal to 1 per column, that span the orthogonal complement of the columns of S_x.

Demodulation amounts to computing the DFT

$$\begin{aligned}
\check{y} &= Fy = F(H\bar{x} + z) \\
&= F(F^H DF(F^H S_x \check{d} + c) + z) \\
&= DS_x \check{d} + DFc + \check{z}
\end{aligned} \tag{3}$$

where $\check{z} = Fz$ has the same distribution of z since F is unitary. Assuming the channel is known at the receiver, we can now estimate c by projecting $\check{y}$ onto the orthogonal complement of the signal subspace leaving us with

$$\acute{y} = S^T \check{y} = S^T DFc + \acute{z}. \tag{4}$$

Note that $\acute{z} = S^T Fz$ is an $m \times 1$ i.i.d Gaussian vector with variance N_0. We denote the $m \times N$ projection matrix obtained by a row selection of F (according to S) by $\Psi = S^T DF$.

The observation vector $\acute{y}$ is a projection of the sparse N-dimensional peak-reducing signal c onto a basis of dimension $m << N$ corrupted by $\acute{z}$. To demonstrate how such an N-dimensional vector can be estimated from m linear measurements, a few relevant notes about compressive sensing are given next. The interested reader is referred to [1] and [3] for details.

An N-dimensional vector v of s nonzero entries such that $s << N$ is said to be s-sparse. We denote by $|s|$ or simply s the cardinality of v, and refer to the positions of the nonzero entries of v as the sparsity pattern, which we formalize by an $s \times 1$ indexing vector, $\jmath_v$. The upshot of compressive sensing techniques is that the sparsity pattern of such a vector can be recovered non-combinatorially (i.e. without the need to minimize $\|v\|_0$) with a penalty on the additional number of measurements m, so that $s < m << N$.

Finding the exact number of measurements m required for exact recovery of v has proven to be difficult in the literature, especially when noise is considered. This critical number not only depends on the dimension of the sparse vector and its cardinality, but also varies according to the measurement matrix Ψ, the recovery algorithm used for estimating v, the estimation Signal-to-Noise-Ratio (SNR), and the Minimum-to-Average-Ratio (MAR) of the nonzero entries of v. Unfortunately, whereas our design uses random row selections of a unitary Fourier matrix for measurement (scaled by the channel coefficients), most results in the literature consider Gaussian ensembles exclusively, and we have yet to find sharp bounds for our case. (refer to [6] for sufficient conditions in the special case of a Fourier matrix indexed by a difference set and to [7] for a thorough treatment of random submatrices).

Assuming the cardinality is known, and for a given SNR and MAR, a necessary and sufficient scaling with respect to N and s for asymptotically reliable sparsity pattern recovery is given in [2] as $m = O(s \log(N - s))$. More specifically, for Lasso, a necessary condition is that

$$m > \frac{2}{MAR \cdot SNR} s \log(N - s) + s - 1 \tag{5}$$

where the SNR is defined as

$$SNR = \frac{E[\|\Psi c\|^2]}{E[\|\acute{z}\|^2]} = \frac{E[\|S^T DF\|^2]}{E[\|S^T Fz\|^2]} \tag{6}$$

and

$$MAR = \frac{min_{j \in u_c} |c_j|}{\|c\|^2 / s} \tag{7}$$

A comparable bound to [2] can be found in [4]. Note that in our case, the number of measurements m is equivalent to the number of reserved tones, while the number of clips is

equivalent to $|s|$, and hence for a fixed number of tones m the number of clips should be below this bound for reliable recovery. By definition, $MAR \in (0, 1]$ with the upper limit occuring when all the nonzero entries of c have the same magnitude. The effect of the SNR is also obvious in controlling s for a fixed m.

Now coming back to our problem, assume the peak reducing signal c is s-sparse in time, given $\acute{y}$ in (4), we use compressive sensing at the receiver to estimate c using a Second Order Cone Program (SOCP) adapted to complex vectors:

$$\begin{aligned} min \quad & \| \tilde{c} \|_1 \\ s.t. \quad & \|\acute{y} - \Psi\tilde{c}\|_2 \leq \varepsilon \end{aligned} \quad (8)$$

The resulting solution by compressive sensing alone is an estimate $\hat{c}_{cs}$ of the peak reducing signal which not only reliably detects the positions of its nonzero entries, but also gives a good approximation to the corresponding amplitudes.

IV. DESIGN AND PERFORMANCE ANALYSIS OF THE PEAK REDUCING SIGNALS

So far we have shown that a sparse, additive peak-reducing signal c can be estimated at the receiver by partially observing its spectral support. Since m is normally fixed, the ability to reduce the PAPR is directly related to increasing s. The bottleneck here is the Symbol Error Rate (SER); the more we increase s, the more prone the recovery is to erroneous sparsity pattern detection and amplitude estimation.

Consequently, an initial intelligent choice of how clipping should be performed has proven to be critical in the performance of the recovery algorithm. What's more, in keeping with our desire to reduce the complexity of the PAPR reduction scheme at the transmitter, unlike [5] we do not need to perform an optimization search for the time-domain clipper that minimizes the PAPR supported by a specified set of reserved tones.

Instead, we use simpler peak-reducing signals that directly operate on the highest temporal peaks. The reserved tones in our design are for estimating c by the frequency distortion naturally induced by clipping, which is completely handled at the receiver.

A. Annihilators based on Level Crossing

The simplest design, requiring the least amount of temporal and spatial complexity at the transmitter, would depend on setting an optimal clipping threshold γ^* such that, for a given SNR and recovery algorithm, the probability of exceeding a number of clips s^* tolerable by (5) is small. The peak reducing signal will hence take the form $c = -\sum_{i:|x_i|\geq\gamma} x_i\delta(n - n_i)$ where n_i corresponds to the location at which $|x_i|$ exceeds γ. Such an approach can guarantee a maximum peak $P_{max} \leq \gamma^2$ provided the recovery of c is reliable.

Note that in this rudimentary work, once $|x_i|$ is above the threshold, we don't clip the portion of which $|x_i|$ exceeds γ, but rather annihilate x_i altogether. This avoids the issue of exploiting the direction of a complex entry of c or $\bar{x}$ to enhance the recovery at the receiver should differential clipping be

used, which we postpone to another more advanced exposition. This also simplifies the comparison with an analogous system in which only least squares is used without compressive sensing. We will henceforth refer to c as an *annihilator*, not *clipper*.

In what follows, for simplicity we assume annihilation is done on the OFDM signal sampled at Nyquist rate, causing the entries of x to be uncorrelated. What's more, following [8] and [9] we assume that the real and imaginary parts of x are asymptotically Gaussian processes for large N. This directly implies that the entries of x are independent and that the envelope of x can be modeled as a sequence of i.i.d Rayleigh random variables with a common cumulative distribution function (CDF) $F_{|X|}(|x|)$. It can be shown that this is a worst case scenario for recovery of c, and that the estimation can actually be enhanced by knowledge of the autocorrelation of an oversampled signal.

By a convenient abuse of notation, we let s be the number of crossings of x to a fixed threshold γ, i.e. the number of entries in x who's magnitude exceeds γ. Consequently, the sparsity s becomes a random variable with a CDF

$$\begin{aligned} F_S(s;\gamma) &= \sum_{i=0}^{s} \binom{N}{i} (1 - F_{|X|}(\gamma))^i F_{|X|}(\gamma)^{N-i} \\ &\approx e^{-N(1-F_{|X|}(\gamma))} \sum_{i=0}^{s} \frac{[N(1-F_{|X|}(\gamma))]^i}{i!}. \quad (9) \end{aligned}$$

For a given number of reserved tones m, we need to ensure that the probability of s violating (5) is low. This entails finding a corresponding annihilating threshold $\gamma^* = \sup\{\gamma : \bar{F}_S(s;\gamma) \leq \epsilon\}$, where $\bar{F}_S(s;\gamma)$ is the complementary CDF of s. This can be found by first determining $\lambda^* = \sup\{\lambda : e^{-\lambda} \sum_{i=0}^{s} \frac{\lambda^i}{i!} \geq 1-\epsilon\}$ where $\lambda = N(1-F_{|X|}(\gamma))$, then simply substituting in $\gamma^* = F_{|X|}^{-1}(1 - \frac{\lambda^*}{N}) = \left(-2\sigma^2 \ln(\frac{\lambda^*}{N})\right)^{1/2}$ where σ is the parameter of the Rayleigh distribution function $F_{|X|}(|x|)$.

Although such an approach will be the simplest to implement at the transmitter, the recovery at the receiver will be poor in general since the variance of the required number of measurement tones m (which is a function of s) will be large. For instance, if we let $m(s) = s\log(N - s) + s$ and use a second order Taylor approximation for evaluating the variance of m in terms of the moments of s we get

$$\begin{aligned} \sigma^2_{m(s)} &\approx \{\dot{m}(E[s])\}^2\sigma_s^2 \\ &= \left(\frac{1}{\ln 2}[1 + \frac{N}{E[s] - N} + \ln(N - E[s])] + 1\right)\sigma_s^2 \\ &= N(1 - F_{|X|}(\gamma))\{1 + \frac{1}{\ln 2}[1 - \frac{1}{F_{|X|}(\gamma)} \\ &+ \ln(NF_{|X|}(\gamma))]\}^2 \quad (10) \end{aligned}$$

resulting in a significantly fluctuating demand for reserved tones. Hence for a fixed m in practice, it's either that γ must be set too high for $P_{max} \leq \gamma^2$ to be significantly lower then the non-limited envelope, or that the symbol error rate will

be high. Consequently, either the recovery performance or the PAPR reduction capability of the system will be sacrificed. The following slightly more complex design can greatly enhance the recovery capability.

B. Annihilators based on Order Statistics

By far, the most valuable function of compressive sensing algorithms is providing a good estimate of the sparsity *pattern* of a sparse vector with a number of measurements much smaller then the vector dimension. However, if these positions were known beforehand, say by a genie-augmented receiver, then estimating the amplitudes at these positions by least squares is optimum (assuming no side-information is used). In-between these two extremes lies an intermediate solution which only requires supplying the receiver with the cardinality s of c at initialization. This means the transmitter can only annihilate s peaks for each OFDM symbol, and these should obviously be the maximum s peaks of the time domain signal for maximum PAPR reduction.

The performance of this scheme is directly related to the tolerable sparsity s of c. Since the cost paid for PAPR reduction is m, this parameter is the main bottleneck in our design, so that the larger m, the more clipping we can perform and recover at the receiver and hence the more PAPR we can reduce. This approach offers significant improvement in terms of SER at the minor cost of increasing the number of comparisons from N (with γ) in the previous thresholding case to a number of ordered selections upper bounded by $N - s + (s-1)\lceil \log_2(N-s+2) \rceil$ [13], which is actually very close to N for $s \ll N$.

Let $x_{i:N}$ be the ith order statistic of x, then the peak-reducing signal can be expressed as $c = -\sum_{j=N-s+1}^{N} x_{j:N} \delta(n - n_j)$. In this case the PAPR is a random variable. The power P of a coefficient in x can be approximated by a Chi-square random variable of two degrees of freedom, and denoting by F_P the CDF of P, the CCDF of the highest power P_{max} is

$$
\begin{aligned}
\bar{F}_{P_{max}}(P) &= Pr\{x^2_{N-s:N} > P\} \\
&= 1 - \sum_{k=N-s}^{N} \binom{N}{k} (F_P)^k (1 - F_P)^{N-k} \\
&= 1 - I_{F_P}(N-s,\ s+1)
\end{aligned}
\tag{11}
$$

where the last expression is Pearson's incomplete beta function [10]. Analysis of the PAPR reduction capability of this scheme amounts to computing the distribution or moments of the difference between the maximum and $(N-s)^{th}$ order statistics which we do not pursue here.

The advantages of this approach include more efficient utilization of the measurement tones and increased reliability of the recovery algorithm. This includes improving both the sparsity pattern and amplitude recovery, resulting in a significantly lower SER.

The reason is this. In the presence of noise, the estimate $\hat{c}_{cs}$ of c will not be a strictly sparse vector, but rather one that has relatively smaller values at the locations of which the original

vector is null (assuming the estimation is good). This requires setting a threshold on the entries of the estimated vector to differentiate between actual nonzero entries of c and noise, and finding such an optimal threshold that would minimize the probability of error is not straightforward.

However, if the the receiver were to know the cardinality of c beforehand, then the compressive sensing estimate of the sparsity pattern can be enhanced by selecting the positions corresponding to the s maximum values in $\hat{c}_{cs}$ as the true support of c. Once these positions have been registered, the amplitude estimate can now be refined by least squares.

Let $\hat{\jmath}^{cs}$ denote the indexing set that registers the positions of the highest s nonzero entries in $\hat{c}_{cs}$ and $\hat{S}_c$ the $N \times s$ selection matrix corresponding to $\hat{\jmath}^{cs}$. Also denote by u_c the $s \times 1$ vector that contains the non-zero entries of c. Then we write $c = \hat{S}_c u_c$ and update (4) to

$$
\begin{aligned}
\acute{y} &= S^T D F \hat{S}_c u_c + \acute{z} \\
&\equiv \Phi_c u_c + \acute{z}
\end{aligned}
\tag{12}
$$

then using the least-squares estimate $\hat{u}_c^{ls} = (\Phi_c^H \Phi_c)^{-1} \Phi_c^H \acute{y}$ we get the corresponding LS estimate of c

$$
c^{ls} = \hat{S}_c (\Phi_c^H \Phi_c)^{-1} \Phi_c^H \acute{y}
\tag{13}
$$

To see the value of this approach, we compare its performance (in the section after the following) to a system which relies on Least Squares alone with an estimate $\hat{\jmath}^{\bar{x}}$ of the sparsity pattern of c based on selecting the smallest s values of $\bar{x}$ instead of compressive sensing such that

$$
\hat{\jmath}^{\bar{x}} = \arg_\jmath \{ |\bar{x}|_{1:N}, |\bar{x}|_{2:N}, \dots, |\bar{x}|_{s:N} \}
\tag{14}
$$

The reason is that, since we annihilate the highest peaks of x, we expect the sparsity pattern of c after noise addition at the receiver to correspond to the positions of where $\hat{\bar{x}}$ is minimum.

V. Comparison with Tellado: Peak reduction vs. Complexity

In the familiar work of Tellado (Ch 4 of [5] in particular), a peak reducing scheme using tone reservation is also implemented. However, the design offered in that work is fundamentally different; The basic idea in [5] is to use the reserved tones as a spectral support for the time-domain signal that minimizes the peak of the OFDM signal, and hence requires an optimization search at the transmitter for each OFDM signal expressed in our notation in the form

$$
\begin{aligned}
&\min_{\check{c}} \quad t \\
&s.t.\ \|x + F^H S \check{c}\|^2 \leq t
\end{aligned}
\tag{15}
$$

where $\check{c} = Fc$ is nonzero only on Ω_c from the definition of S. Clearly, this optimization approach will result in significantly more PAPR reduction compared to our design, since for the same number of reserved tones m, we can only clip $s < m$ maximum peaks, whereas by Tellado's method no such restriction exists. However, the complexity factor is critical in any PAPR reduction scheme, and this is where our method justifies its utility.

Firstly, the main complexity (i.e. the stage at which the optimization search is performed) in Tellado's technique is at the transmitter. As for our technique, the peak reducing signal is automatically constructed from annihilating the maximum peaks, leaving the problem of estimating the positions and amplitudes of the clips by a SOCP to the receiver. Secondly, although Tellado's general Quadratically Constrained Quadratic Program can be replaced with a simpler suboptimal iterative algorithm, the complexity is nevertheless always a function of N, whereas ours is a function of $m << N$, which can also be replaced with simpler recovery techniques such as basis pursuit. [1]

Thirdly, a very important concept which must be stressed is that compressive sensing by definition is inherently a complexity-reducing scheme. Although one could theoretically search for the sparsity pattern of a s-sparse vector corrupted by AWGN using Maximum Likelihood (ML) with $s+1$ measurements instead of m [2], the associated complexity is enormous, and this is why one pays an extra $m-s$ measurements to use convex programs instead. Hence in all disregard to complexity, one could theoretically beat the PAPR reduction capability of Tellado's approach with the same number of peak reducing tones by amplifying and shifting the complexity to the receiver.

It is worth mentioning that Tellado also suggests a suboptimal system that avoids repetitive optimization by designing *peak reducing kernels*. These are used for directly clipping the peaks which exceed a threshold (similar to our basic threshold-based annihilator), but are, however, designed at initialization to have a confined spectral support based on the provided peak reducing tones. The problem is that due to this confined spectral support, these kernels have temporal sidelobes that cause variable peak-regrowth among the OFDM signals, and Tellado attempts to optimize the location of the peak reducing tones to minimize these sidelobes. Clearly, our approach sidesteps this annoying problem since we need not confine the spectral support of c.

Finally, it appears to the authors that the two techniques (our's and Tellado's) can be combined in a mobile-user setting, such that the mobile station uses our technique to transmit, leaving the burden of estimating c to the base station, and allowing the base station to use Tellado's technique when transmitting, by that simplifying the complexity at the mobile station in both stages of transmission and reception.

VI. SIMULATIONS AND RESULTS

In support of a comprehensive analysis of our design, we tested the PAPR reduction capability, the SER, and the temporal complexity for different scenarios and compared against possible alternatives. For $N = 256$, $M = 32$ and $m = 20\%$ of N, we compared the PAPR resulting from annihilating the highest 8 peaks of 1000 OFDM blocks to the PAPR of the original signal and to Tellado's primary algorithm under the same conditions. Figure 1 shows the CCDF of

the PAPR in the three cases. Results show that at a PAPR excursion probability of 10^{-3}, the PAPR of the original OFDM signal was approximately 11 dB, the PAPR after using (15) was near 4.3 dB, while using the basic annihilator with $s = 8$ the PAPR was about 6.5 dB. Obviously once the detection process is improved we can safely clip more then 8 peaks and reduce the PAPR even further.

In terms of SER, we compared the performance of the compressive sensing algorithm in (8) with the case where we augment the algorithm by least squares given the knowledge of s as expressed in (13), and with the case where no compressive sensing is used and least squares is done on the pattern estimated by (14). In figure 2, we show the SER as we varied s for the three aforementioned cases given the same settings in the previous simulation. We note that the knowledge of s in (13) offers significant advantage over (8), especially once we plotted the SER against SNR which we do not show here for lack of space. Most importantly, these two cases drastically outperform the case where compressive sensing is not used. The naive estimation of the sparsity pattern conveyed by (14) only begins to outperform the two once the sparsity condition in (5) is grossly violated, and in that case the estimates given by all three are useless.

To test the effect of measurement-tone positioning on the SER of (13), we ran the algorithm for three different cases: A) m tones drawn randomly from N, B) m tones positioned uniformly over the spectrum of x, and C) m and N selected to compose a difference set. In this case we were limited in our selection of N and m, and the closest setting we found to our original one was $N = 197$ and $m = 49$ with Ω_c matching a (197, 49, 12) difference set found in [14].

Figure 3 confirms that uniform tone positioning results in extremely poor performance of compressive sensing algorithms. It also confirms that random tone assignment is nearly optimal as stated in [3]. Although the use of difference sets appears to outperform random positioning, the optimality of such an approach and the extent to which they can improve recovery is still unknown. A last point conveyed by figure 3 is that the SER does not increase monotonically with s, indicating that choosing m too high or s too low for a fixed m is not always advantageous due to the possibility of detecting false positives.

Lastly, to compare the relative temporal complexity of each technique, in the last figure we plot the empirical CCDF of the normalized relative time required to execute 1000 runs of (8) and (15). Using [12] with $\varepsilon = 10^{-3}$ and the same settings as in the first simulation, the average time required to perform the optimization in (8) was less then (15) by 89.34%.

---

[1] Although we tested basis pursuit in our design and found it to perform just as well as a SOCP, we preferred not adopting it as the simplest recovery algorithm due to theoretical reasons related to convergence.

REFERENCES

[1] E. J. Candes, J. Romberg and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[2] A.K. Fletcher, S. Rangan, and V. Goyal, "Necessary and Sufficient Conditions on Sparsity Pattern Recovery," arXiv:0804.1839v1 [cs.IT], Apr. 2008. Submitted to *IEEE Trans. Inform. Theory*.
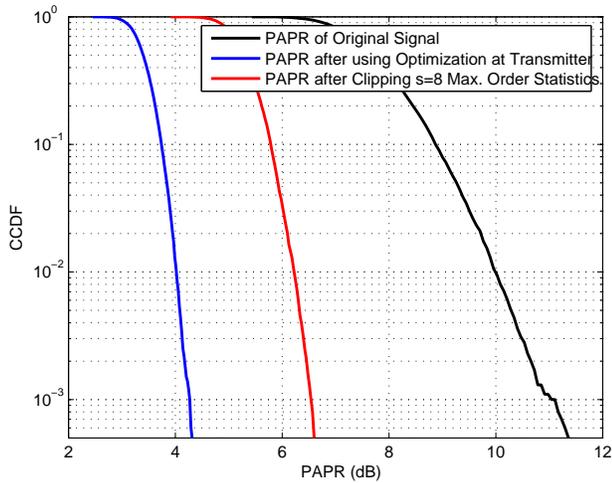
Fig. 1. Comparison of Peak-to-Average-Power Ratio, $N = 256$, $m = 20\%$ of N, SNR = 30 dB, M = 8
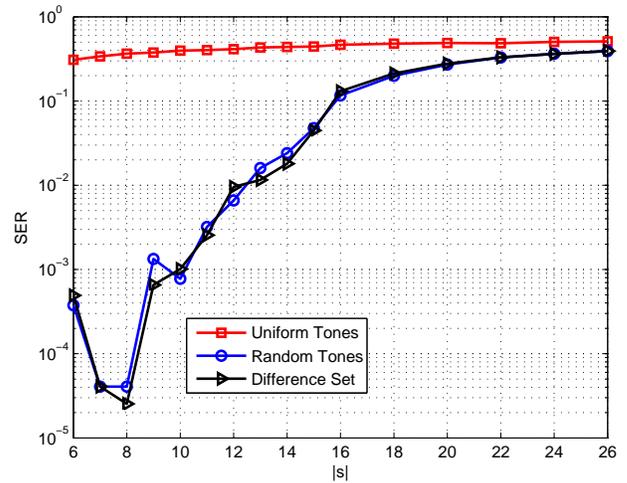
Fig. 3. Symbol Error Rate vs $|s|$ for different tone positioning schemes. $N = 197$, $m = 49$
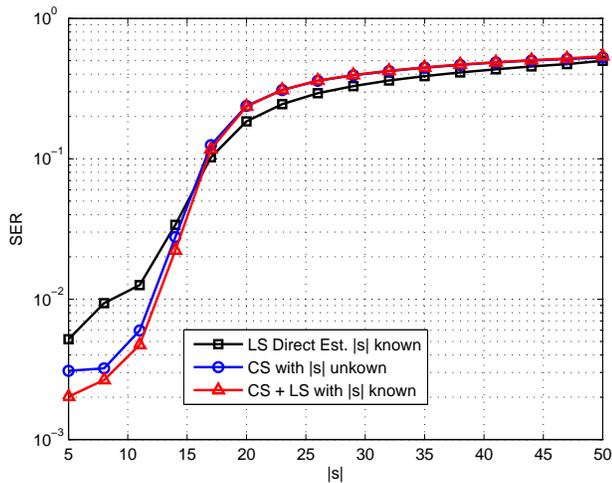
Fig. 2. Symbol Error Rate vs $|s|$ for (8), (13), and (14)
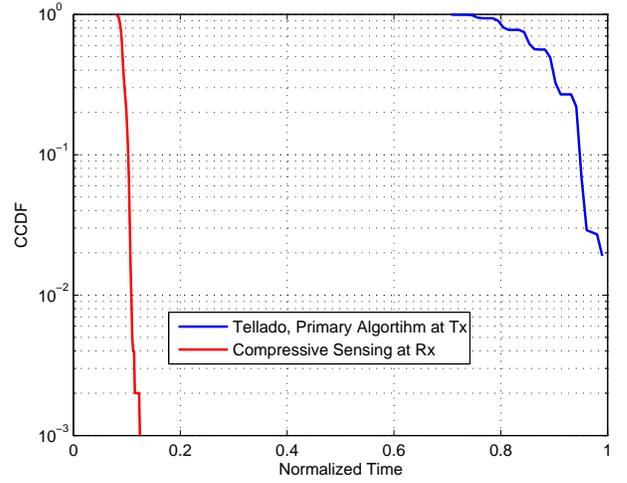
Fig. 4. Normalized CCDF, $N = 256$, $m = 20\%$ of N, $\varepsilon = 10^{-3}$

[3] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections:universal encoding strategies?," *Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406-5425, Dec. 2006.

[4] M. Wainwright, "Sharp thresholds for high-dimensional and noisy recovery of sparsity," arXiv:math/0605740v1 [math.ST], May 2006.

[5] J. Tellado, *Multicarrier Modulation with Low PAR Applications to DSL and Wireless*. Kluwer Academic Publishers, Norwell 2000.

[6] 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. on Inform. Theory (ISIT)*, July 2008.

[7] J. A. Tropp, "On the conditioning of random subdictionaries," *Appl. Comput. Harmonic Anal.*, vol. 25, no. 1, pp. 1-24, 2008

[8] 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.

[9] S. Wei, D.L. Goeckel, and P.E. Kelly, "A modern extreme value theory approach to calculating the distribution of the peak-to-average power ratio in OFDM systems," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, 2002.

[10] B. Arnold, N. Balakrishnan, and H.N. Nagaraja, *A First Course in Order Statistics*, Society for Industrial and Applied Mathematics, Classics in Applied Mathematics, vol. 54, Philadelphia, 2008.

[11] P. Xia, S. Zhou, and G.B Giannakis, "Achieving the Welch bound with difference sets," *Trans. Inform. Theory*, vol. 51, no. 5, pp. 1900-1907, May 2005.

[12] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming (web page and software). Available: http://stanford.edu/ boyd/cvx, February 2009.

[13] P. Růžiéka and J. Wiedermann, "On the lower bound for minimum comparison selection". *Applications of Mathematics*, vol. 23, no. 1, pp. 18, 1978.

[14] D. Gordon, La Jolla Difference Set Repository [Online]. Available: http: //www.ccrwest.org/diffsets/diff_sets/baumert.html, last date of access 29 March 2009.