A Bayesian Sparse Reconstruction Framework for Mitigation of Non-Linear Effects in OFDM Systems

Javier García, Jawad Munir, Amine Mezghani and Josef A. Nossek
Institute for Circuit Theory and Signal Processing
Technische Universität München, 80290 Munich, Germany
E-Mail: {javier.garcia, jawad.munir, amine.mezghani, josef.a.nossek}@tum.de

***Abstract*—The mitigation of nonlinear distortion caused by power amplifiers (PA) in Orthogonal Frequency Division Multiplexing (OFDM) systems is an essential issue to enable energy efficient operation. We propose a new algorithm for receiver-based clipping estimation in OFDM systems that combines the existing Iterative Hard Thresholding method with a novel Bayesian framework to estimate clipping parameters at the receiver. We avoid the use of pilots and formulate the recovery problem solely on reliably detected sub-carriers. We also develop a new criterion for selecting these reliable carriers that takes into account the channel code. Through simulations, we show that the proposed technique outperforms the existing methods both in terms of BER and speed.**

I. Introduction

With the recent leap from 3G to 4G, the power consumption of mobile devices has increased considerably. The 4G operates on Long Term Evolution (LTE), in which Orthogonal Frequency Division Multiplexing (OFDM) is used for the wireless communications. However, the use of multi-carrier modulations such as OFDM leads to a higher Peak-to-Average Power Ratio (PAPR) and thus higher power consumption in an RF Power Amplifier (PA).

Several PAPR reduction algorithms have been proposed to improve power efficiency [1]. Coding schemes [2] and Partial Transmission Sequences [3] come at the cost of throughput reduction. Companding transforms [4] and tone injection techniques [5] increase the complexity of the transmitter.

In some applications, such as the uplink of mobile communications systems, it is desirable to shift as much computational effort as possible to the receiver side, while keeping the transmitter simple. In these cases, an attractive PAPR reduction technique consists of just clipping the signal at the transmitter and then compensating for the clipping at the receiver.

A powerful distortion cancellation algorithm, called Scaled Coded Power Amplifier Nonlinearity Cancellation (sCPANC), was proposed in [6]. This method relies on iterative decoding, and suffers its associated complexity. It also requires knowledge of the input-output characteristic of the PA at the receiver.

The Support Agnostic Bayesian Matching Pursuit (SABMP) algorithm, proposed in [7], exploits the sparse nature of the clipping error while choosing a certain number of reliable subcarriers to avoid affecting the throughput of the system. Its main drawback is its high computational complexity.

We propose a new sparse reconstruction approach based on Weighted Iterative Hard Thresholding, which outperforms the existing clipping recovery techniques SABMP and sCPANC in both BER and execution speed terms, while still having no effect on the throughput of the channel.

Our paper is organized as follows. Section II describes the system model and formulates the recovery problem. In Sections III to V, the proposed Weighted Iterative Hard Thresholding algorithm and the developed Bayesian framework for parameter estimation are presented. Section VI compares the computational complexity of this technique with SABMP and sCPANC. In Section VII we provide some simulation results to reveal the usefulness of the method, and Section VIII gives conclusions and suggests some areas for future work.

II. System Model

Consider an OFDM system with Q-QAM modulation and N subcarriers. An IFFT is applied to each OFDM block $\boldsymbol{\mathcal{X}} \in \mathbb{C}^{N\times 1}$ to obtain the time-domain vector $\mathbf{x} = \mathbf{F}^H\boldsymbol{\mathcal{X}} \in \mathbb{C}^{N\times 1}$, where $\mathbf{F}$ is a $N \times N$ DFT matrix [8].

The time-domain signal $\mathbf{x}$ has high PAPR due to the addition of different frequency sub-carriers. To avoid saturating the power amplifier (PA), a soft clipping operation is applied to $\mathbf{x}$ before sending it to the PA. The clipped signal $\mathbf{x_p} \in \mathbb{C}^{N\times 1}$ is given by:

$$x_p[n] = g\left(x[n]\right) = \begin{cases} x[n], & |x[n]| \leq \tau \\ \tau e^{j\arg x[n]} & |x[n]| > \tau \end{cases} \tag{1}$$

where τ is the clipping threshold. We will model the clipping as an additive distortion $\mathbf{c}$ applied to the time-domain signal $\mathbf{x}$. The output $\mathbf{x_p}$ of the clipping process is then:

$$\mathbf{x_p} = \mathbf{x} + \mathbf{c} = \mathbf{F}^H\boldsymbol{\mathcal{X}} + \overline{\mathbf{c}}. \tag{2}$$

After addition and removal of a sufficiently long cyclic prefix (CP), the receiver applies an IFFT to the received signal:

$$\boldsymbol{\mathcal{Y}} = \mathbf{FHF}^H\boldsymbol{\mathcal{X}} + \mathbf{FH}\overline{\mathbf{c}} + \mathbf{Fz}, \tag{3}$$

where $\mathbf{z} \in \mathbb{C}^{N\times 1}$ is the AWGN noise with variance σ_z^2. The channel matrix $\mathbf{H} \in \mathbb{C}^{N\times N}$ is circulant due to the use of a CP. Therefore, the DFT matrices diagonalize it, such that $\boldsymbol{\Lambda} = \mathbf{FHF}^H$ is a diagonal matrix, and furthermore $\mathbf{FH} = \boldsymbol{\Lambda}\mathbf{F}$. This allows rewriting the frequency-domain received signal as:

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\Lambda}\boldsymbol{\mathcal{X}} + \boldsymbol{\Lambda}\mathbf{F}\overline{\mathbf{c}} + \mathbf{Fz}. \tag{4}$$

Equalization of the channel amounts to multiplying with the inverse of the diagonal matrix $\mathbf{\Lambda}$:

$$\overline{\boldsymbol{\mathcal{Y}}} = \mathbf{\Lambda}^{-1}\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{X}} + \mathbf{F}\overline{\mathbf{c}} + \mathbf{\Lambda}^{-1}\mathbf{F}\mathbf{z}. \tag{5}$$

In order to determine the unknown vector $\mathbf{c}$, the data term $\boldsymbol{\mathcal{X}}$ needs to be removed from (5). This could be done with the help of pilots, but there is a more efficient way that avoids bandwidth loss. A set $\mathcal{R}$ of M reliable carriers is chosen, in which there is high certainty that the distortion term $\mathbf{F}\mathbf{c} + \overline{\boldsymbol{\mathcal{Z}}}$ was not high enough to move the symbol to a different constellation region. This means that the demapped symbols in the reliable carriers are the same as the transmitted symbols:

$$\mathbf{J}\boldsymbol{\mathcal{X}} = \mathbf{J}\mathcal{Q}\left[\overline{\boldsymbol{\mathcal{Y}}}\right], \tag{6}$$

where $\mathbf{J} \in \{0,1\}^{M\times N}$ is a selection matrix that selects the reliable carriers, and $\mathcal{Q}\left[\overline{\boldsymbol{\mathcal{Y}}}\right]$ denotes the closest QAM constellation points to each element of $\overline{\boldsymbol{\mathcal{Y}}}$. The procedure for selection of the reliable carriers is addressed in Section III.

The receiver subtracts the demapped symbols from the reliable carriers:

$$\begin{aligned}\tilde{\boldsymbol{\mathcal{Y}}} = \mathbf{J}\overline{\boldsymbol{\mathcal{Y}}} - \mathbf{J}\mathcal{Q}\left[\overline{\boldsymbol{\mathcal{Y}}}\right] &= \mathbf{J}\boldsymbol{\mathcal{X}} - \mathbf{J}\mathcal{Q}\left[\overline{\boldsymbol{\mathcal{Y}}}\right] + \mathbf{J}\mathbf{F}\overline{\mathbf{c}} + \mathbf{J}\mathbf{\Lambda}^{-1}\mathbf{F}\mathbf{z} \\ &= \mathbf{J}\mathbf{F}\overline{\mathbf{c}} + \mathbf{J}\mathbf{\Lambda}^{-1}\mathbf{F}\mathbf{z}.\end{aligned} \tag{7}$$

Note also that, from (1), the clipping error $\overline{\mathbf{c}}$ always has the opposite phase to that of the clipped signal $\mathbf{x_p}$. The receiver can therefore estimate the received signal (see (5)):

$$\hat{\mathbf{x}}_{\mathbf{p}} = \mathbf{F}^H\overline{\boldsymbol{\mathcal{Y}}} = \mathbf{x} + \overline{\mathbf{c}} + \mathbf{F}^H\mathbf{\Lambda}^{-1}\mathbf{F}\mathbf{z}, \tag{8}$$

and estimate the phase of $\overline{\mathbf{c}}$ as $\boldsymbol{\theta} = \pi + \arg\{\hat{\mathbf{x}}_{\mathbf{p}}\}$. This allows expressing $\overline{\mathbf{c}}$ as:

$$\overline{\mathbf{c}} = \mathbf{\Theta}\mathbf{c}, \tag{9}$$

where $\mathbf{\Theta} = \operatorname{diag}\{\exp(j\boldsymbol{\theta})\} \in \mathbb{C}^{N\times N}$, and $\mathbf{c} = |\overline{\mathbf{c}}| \in \mathbb{R}^N$. With this treatment, now only the real-valued, positive amplitude $\mathbf{c}$ of $\overline{\mathbf{c}}$ needs to be estimated. By plugging (9), into (7), we obtain the final sparse recovery problem formulation:

$$\tilde{\boldsymbol{\mathcal{Y}}} = \mathbf{A}\mathbf{c} + \tilde{\boldsymbol{\mathcal{Z}}}, \tag{10}$$

where $\mathbf{A} = \mathbf{J}\mathbf{F}\mathbf{\Theta} \in \mathbb{C}^{M\times N}$ is the sensing matrix, and $\tilde{\boldsymbol{\mathcal{Y}}} = \mathbf{J}\left(\overline{\boldsymbol{\mathcal{Y}}} - \mathcal{Q}\left[\overline{\boldsymbol{\mathcal{Y}}}\right]\right) \in \mathbb{C}^{M\times 1}$ is the observation vector. The noise $\tilde{\boldsymbol{\mathcal{Z}}} = \mathbf{J}\mathbf{\Lambda}^{-1}\mathbf{F}\mathbf{z} \in \mathbb{C}^{N\times 1}$ has diagonal covariance matrix:

$$\mathbf{R}_{\tilde{\boldsymbol{\mathcal{Z}}}\tilde{\boldsymbol{\mathcal{Z}}}} = \sigma_z^2\mathbf{J}\mathbf{\Lambda}^{-1}\left(\mathbf{\Lambda}^{-1}\right)^H\mathbf{J}^H \in \mathbb{C}^{M\times M}. \tag{11}$$

The receiver needs to estimate $\mathbf{c}$ and then subtract $\mathbf{F}\mathbf{c}$ from (5) to obtain the undistorted symbols $\boldsymbol{\mathcal{X}}$.

III. Selection of reliable carriers

The choice of reliable carriers in (6) is done by computing a reliability measure $R[n]$ for each subcarrier n, and choosing the M subcarriers with highest $R[n]$.

If the channel code is not taken into account, the optimal reliability measure has already been proposed in [7], and computes the a posteriori probability of the symbol coming from the closest constellation point, $p\left(\mathcal{Q}\left[\overline{\mathcal{Y}}[n]\right] = \mathcal{X}[n] \mid \overline{\mathcal{Y}}[n]\right)$:

$$R_{\mathrm{LR}}[n] = \frac{e^{-\frac{\left|\overline{\mathcal{Y}}[n] - \mathcal{Q}\left[\overline{\mathcal{Y}}[n]\right]\right|^2}{\sigma_z^2}}}{\sum_{q=1}^{Q} e^{-\frac{\left|\overline{\mathcal{Y}}[n] - \mathcal{Q}_q\right|^2}{\sigma_z^2}}}, \tag{12}$$

where $\mathcal{Q}_q$ denotes the q-th constellation point.

In this work, we propose a more accurate reliability measure that takes the channel code into account. If an a posteriori probability (APP) decoder is used [9], it can output the log-likelihood ratios (LLRs) of the encoded bits:

$$\mathrm{LLR}_{n,i} \triangleq \log\frac{p\left(b_{n,i} = 1\right)}{p\left(b_{n,i} = 0\right)}, \tag{13}$$

where $b_{n,i}, n \in \{1,\ldots,N\}, i \in \{1,\ldots,\log_2 Q\}$ denotes the i-th bit of the n-th subcarrier, in the encoded domain. The probability that this bit is correct can then be computed as:

$$p(\hat{b}_{n,i} = b_{n,i}) = \frac{e^{|\mathrm{LLR}_{n,i}|}}{1 + e^{|\mathrm{LLR}_{n,i}|}}. \tag{14}$$

The product of these probabilities for all bits in one subcarrier gives the probability that the subcarrier symbol is correct, which is our proposed reliability measure. In logarithmic units:

$$R_{\mathrm{CC}}[n] = \sum_{i=1}^{\log_2 Q} \log\frac{e^{|\mathrm{LLR}_{n,i}|}}{1 + e^{|\mathrm{LLR}_{n,i}|}}. \tag{15}$$

This measure uses information about the channel code, and can correctly select reliable carriers with a higher distortion term than the ones chosen by (12). This larger distortion term is better suited for the subsequent distortion estimation.

IV. Sparse Reconstruction of the Clipping

A. Problem formulation

Once the reliable carriers are selected, the receiver needs to obtain an estimate of $\mathbf{c}$ by solving the sparse reconstruction problem in (10). Due to the large amount of subcarriers in a typical OFDM scenario (256-2048), the ℓ_1 minimization techniques such as Basis Pursuit Denoising [10] might be too slow for real-time applications. Therefore, we focus on greedy approaches to the compressed sensing problem:

$$\hat{\mathbf{c}} = \arg\min\left[\left\|\tilde{\boldsymbol{\mathcal{Y}}} - \mathbf{A}\mathbf{c}\right\|_2^2\right] \text{ s.t. } \|\mathbf{c}\|_0 < K, \tag{16}$$

i.e. the aim is to minimize the Mean Squared Error (MSE) of the transformed clip signal $\tilde{\boldsymbol{\mathcal{Y}}}$, subject to a maximum allowed number of nonzero taps K in $\mathbf{c}$. The estimation of this K is a key issue and is dealt with in Section V.

B. Weighted Iterative Hard Thresholding

The Support Agnostic Bayesian Matching Pursuit (SABMP) technique proposed in [7] is very robust in a wide range of scenarios, but its complexity is extremely high (see Section VI). Therefore, we propose an alternative much faster method.

Our proposed method, Weighted Iterative Hard Thresholding (WIHT), has three inputs: the sensing matrix $\mathbf{A}$, the observation vector $\tilde{\boldsymbol{\mathcal{Y}}}$, and a weighting vector $\mathbf{w} \in \mathbb{R}^{N\times 1}$. This

Algorithm 1 Weighted Iterative Hard Thresholding (WIHT)

Input: $\tilde{\boldsymbol{\mathcal{Y}}}$, $\mathbf{A}$, $\mathbf{w}$, K

Support set: $\hat{\mathcal{S}} = \text{supp}\left\{H_K\left(\text{diag}\left\{\mathbf{w}\right\}\mathbf{A}^H\tilde{\boldsymbol{\mathcal{Y}}}\right)\right\}$

BLUE Estimate $\hat{\mathbf{c}} = \mathbf{J}_{\hat{\mathcal{S}}}^H(\mathbf{A}_{\hat{\mathcal{S}}}^{\mathrm{H}}\mathbf{R}_{\tilde{\mathcal{Z}}\tilde{\mathcal{Z}}}^{-1}\mathbf{A}_{\hat{\mathcal{S}}})^{-1}\mathbf{A}_{\hat{\mathcal{S}}}^{\mathrm{H}}\mathbf{R}_{\tilde{\mathcal{Z}}\tilde{\mathcal{Z}}}^{-1}\tilde{\boldsymbol{\mathcal{Y}}}$

Output: $\hat{\mathbf{c}}$

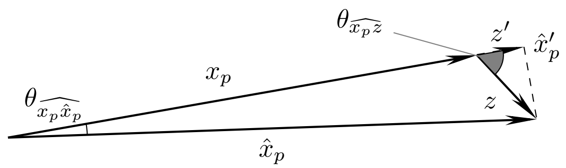

Figure 1. Approximation of $|\hat{x}_p|$ by $\left|\hat{x}'_p\right| = \left||x_p| + |z|\cos\theta_{\widehat{x_p z}}\right|$.

weighting vector is an estimate of the a priori probability of each tap of the solution vector to be active (nonzero), and in Section V we propose an accurate estimator for it.

The proposed algorithm is based on Iterative Hard Thresholding (IHT) [11], which consists of a Steepest Descent approach in which a hard thresholding operator $H_K(\cdot)$ is applied after each step. This operator keeps the K maximum samples of its argument and sets all other taps to 0.

The proposed WIHT method runs a single iteration of standard IHT but applying the weighting $\mathbf{w}$ before thresholding:

$$\hat{\mathbf{c}}^{(1)} = H_K\left(\text{diag}\left\{\mathbf{w}\right\}\mathbf{A}^H\tilde{\boldsymbol{\mathcal{Y}}}\right). \tag{17}$$

Our experiments showed that, if the weighting $\mathbf{w}$ is accurate enough, the support of $\hat{\mathbf{c}}^{(1)}$ after only one iteration usually coincides with that of $\mathbf{c}$. Therefore, WIHT takes this support set $\hat{\mathcal{S}} = \text{supp}\left\{\hat{\mathbf{c}}^{(1)}\right\}$ and performs a Best Linear Unbiased Estimate (BLUE) over it:

$$\hat{\mathbf{c}} = \mathbf{J}_{\hat{\mathcal{S}}}^H(\mathbf{A}_{\hat{\mathcal{S}}}^{\mathrm{H}}\mathbf{R}_{\tilde{\mathcal{Z}}\tilde{\mathcal{Z}}}^{-1}\mathbf{A}_{\hat{\mathcal{S}}})^{-1}\mathbf{A}_{\hat{\mathcal{S}}}^{\mathrm{H}}\mathbf{R}_{\tilde{\mathcal{Z}}\tilde{\mathcal{Z}}}^{-1}\tilde{\boldsymbol{\mathcal{Y}}}, \tag{18}$$

where $\mathbf{J}_{\hat{\mathcal{S}}} \in \{0,1\}^{K\times N}$ selects the support set, and $\mathbf{A}_{\hat{\mathcal{S}}} = \mathbf{A}\mathbf{J}_{\hat{\mathcal{S}}}^H$ contains the columns of $\mathbf{A}$ corresponding to the support set. The noise covariance matrix $\mathbf{R}_{\tilde{\mathcal{Z}}\tilde{\mathcal{Z}}}$ is given by (11).

V. Clip parameter estimation

The WIHT algorithm is very fast, but its performance depends highly on the accuracy of the weighting function $\mathbf{w}$, and of the number of active taps K. We have developed an accurate receiver-based Bayesian estimator for these parameters.

A. Statistical properties of the received signal

The proposed method uses the estimated clipped signal $\hat{\mathbf{x}}_\mathbf{p}$ from (8) as an input. We neglect the correlations of the noise vector $\mathbf{F}^H\boldsymbol{\Lambda}^{-1}\mathbf{F}\mathbf{z}$ and consider each received sample $\hat{x}_p[n]$ separately. This is a suboptimal approach, but still greatly outperforms the heuristic used so far. For the remainder of this section, we drop the index $[n]$ for convenience, and denote the current sample of $\mathbf{F}^H\boldsymbol{\Lambda}^{-1}\mathbf{F}\mathbf{z}$ by z. Note that, although correlated, these samples are still circularly symmetric. The magnitude of each tap $\hat{x}_p$ of the estimated clipped signal is then (note that x_p includes the clipping (2)):

$$|\hat{x}_p| = |x_p + z| = \sqrt{|x_p|^2 + 2|x_p||z|\cos\left(\theta_{\widehat{x_p z}}\right) + |z|^2}. \tag{19}$$

We now approximate the magnitude of $\hat{x}_p$ by that of its projection x'_p in the direction of x_p, as depicted in Figure 1:

$$|\hat{x}_p| \approx \left|x'_p\right| = \left||x_p| + |z|\cos\theta_{\widehat{x_p z}}\right|. \tag{20}$$

This is equivalent to replacing $|z|^2$ with $|z|^2\cos^2\theta_{\widehat{x_p z}}$ in (19), and holds as long as $|x_p| \gg |z|$. We note that, if this does not hold, it means that $|x_p|$ is small and thus the probability of clipping is low. Our proposed weighting function (36) will still give a value close to 0 in this case, and therefore the approximation error is not relevant for our purpose.

The unclipped signal x is the sum of a large number N of i.i.d. symbols, and therefore can be assumed Gaussian with variance σ_x^2. Its magnitude is therefore Rayleigh:

$$p_{|x|}(u) = \frac{2u}{\sigma_x^2}e^{-\frac{u^2}{\sigma_x^2}}, \quad u \geq 0, \tag{21}$$

where $u \triangleq |x|$. The probability density function (PDF) of the magnitude of the clipped signal $u_p \triangleq |x_p|$ is easy to obtain: it is the same as that of $|x|$ in the unclipped region ($u < \tau$), and concentrates all the probability mass for $|x| > \tau$ (computed as $\int_\tau^\infty p_{|x|}(u)\,\mathrm{d}u$) into a Dirac delta at $u_p = \tau$:

$$p_{|x_p|}(u_p) = \begin{cases} \frac{2u_p}{\sigma_x^2}e^{-\frac{u_p^2}{\sigma_x^2}}, & 0 \leq u_p < \tau, \\ e^{-\frac{\tau^2}{\sigma_x^2}}\delta\left(u_p - \tau\right) & u_p \geq \tau. \end{cases} \tag{22}$$

Now, we obtain the PDF of $x'_p = |x_p| + |z|\cos\theta_{\widehat{x_p z}}$. Due to the circular symmetry of z, the PDF of $z' = |z|\cos\theta_{\widehat{x_p z}}$ is the same as that of the real part of the noise:

$$p_{z'}(z') = \frac{1}{\sigma_z\sqrt{\pi}}e^{-\frac{z'^2}{\sigma_z^2}}. \tag{23}$$

The PDF of x'_p is the convolution of the PDFs of x_p and z':

$$p_{x'_p}(x'_p) = \int_0^\tau p_{x_p}(x_p)p_{z'}(x'_p - x_p)\,\mathrm{d}x_p \tag{24}$$

Then, the PDF of $\hat{u}_p \triangleq \left|x'_p\right| \approx |\hat{x}_p|$ is obtained by summing the positive and negative parts of $p_{x'_p}(x'_p)$:

$$p_{|\hat{x}_p|}(\hat{u}_p) \approx p_{x'_p}(\hat{u}_p) + p_{x'_p}(-\hat{u}_p). \tag{25}$$

The result of applying (24) and (25) to (22) is a long closed expression that can be written as:

$$p_{|\hat{x}_p|}(\hat{u}_p) \approx f_c(\hat{u}_p) + f_{\overline{c}}(\hat{u}_p), \tag{26}$$

where $f_c(\hat{u}_p) \triangleq p\left(\text{“clip”}\right)p\left(\hat{u}_p \,\middle|\, \text{“clip”}\right)$ is given by:

$$f_c(\hat{u}_p) = \frac{1}{\sigma_z\sqrt{\pi}}e^{-\frac{(\hat{u}_p-\tau)^2}{\sigma_z^2}-\frac{\tau^2}{\sigma_x^2}}\left(1 - e^{-\frac{4\hat{u}_p\tau}{\sigma_z^2}}\right), \tag{27}$$

and $f_{\overline{c}}(\hat{u}_p) \triangleq p\left(\overline{\text{“clip”}}\right) p\left(\hat{u}_p \mid \overline{\text{“clip”}}\right)$ is computed as:

$$f_{\overline{c}}(\hat{u}_p) = \frac{1}{\sigma_x^2+\sigma_z^2}\left\{\frac{\sigma_z}{\sqrt{\pi}}\left[2e^{-\frac{\hat{u}_p^2}{\sigma_z^2}} - \right.\right.$$
$$\left. -e^{-\frac{(\hat{u}_p-\tau)^2}{\sigma_z^2}-\frac{\tau^2}{\sigma_x^2}}\left(1-e^{-\frac{4\hat{u}_p\tau}{\sigma_z^2}}\right)\right] + \frac{2\hat{u}_p\sigma_x}{\sqrt{\sigma_x^2+\sigma_z^2}}e^{-\frac{\hat{u}_p^2}{\sigma_x^2+\sigma_z^2}}\left[\right.$$
$$\Phi\left(\frac{\sqrt{2}}{\sigma_z}\left(\tau\frac{\sqrt{\sigma_x^2+\sigma_z^2}}{\sigma_x} - \hat{u}_p\frac{\sigma_x}{\sqrt{\sigma_x^2+\sigma_z^2}}\right)\right) + 2\Phi\left(\frac{\sqrt{2}\hat{u}_p\sigma_x}{\sigma_z\sqrt{\sigma_x^2+\sigma_z^2}}\right) -$$
$$\left.\left. -\Phi\left(\frac{\sqrt{2}}{\sigma_z}\left(\tau\frac{\sqrt{\sigma_x^2+\sigma_z^2}}{\sigma_x} + \hat{u}_p\frac{\sigma_x}{\sqrt{\sigma_x^2+\sigma_z^2}}\right)\right) - 1\right]\right\}, \tag{28}$$

where $\Phi(v) \triangleq \int_{-\infty}^{v} \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2}}\, \mathrm{d}\theta$ denotes the cumulative Gaussian distribution function.

B. Proposed estimator

From (22) and (23) we can derive closed form expressions for the second and fourth order moments of $|\hat{x}_p|$:

$$\mu_2 = \mathrm{E}\left\{|x_p|^2\right\} + \mathrm{E}\left\{|z'|^2\right\}, \tag{29}$$

$$\mu_4 = \mathrm{E}\left\{|x_p|^4\right\} + 4\mathrm{E}\left\{|x_p|^2\right\}\mathrm{E}\left\{|z'|^2\right\} + \mathrm{E}\left\{|z'|^4\right\}, \tag{30}$$

yielding:

$$\begin{cases}\mu_2 = \sigma_x^2\left(1-e^{-\alpha}\right) + \sigma_z^2, \\ \mu_4 = 2\sigma_x^4\left[1-(1+\alpha)e^{-\alpha}\right] + 2\sigma_z^2\left(2\mu_2 - \sigma_z^2\right),\end{cases} \tag{31}$$

where $\alpha \triangleq \tau^2/\sigma_x^2$. This is a nonlinear system of two equations with two unknowns α and σ_x. By computing:

$$\beta \triangleq \frac{\mu_4 + 2\sigma_z^2\left(\sigma_z^2 - 2\mu_2\right)}{2\left(\mu_2 - \sigma_z^2\right)} = \frac{1-(1+\alpha)e^{-\alpha}}{1-e^{-\alpha}}, \tag{32}$$

we can solve for the two α variables inside $e^{-\alpha}$ terms in the right-hand side of (32), while treating the other α as a constant:

$$\alpha = \ln\left(\frac{2\beta}{2\beta - 1 - \alpha + \sqrt{(1+\alpha)^2 - 4\alpha\beta}}\right). \tag{33}$$

This fixed-point equation is solved iteratively for α to obtain the estimated input signal variance $\sigma_x^2 = \left(\mu_2 - \sigma_z^2\right)/\left(1-e^{-\alpha}\right)$ and clipping threshold $\tau = \sigma_x\sqrt{\alpha}$, as shown in Algorithm 2.

From (22), the probability of clipping is given by $\rho = e^{-\alpha}$. Therefore, the number of active taps K can be estimated as:

$$K = Ne^{-\alpha}, \tag{34}$$

and then increased by a certain amount to account for the fact that overestimating K is much less harmful to the result of WIHT than underestimating it.

Finally, the most important parameter, the weighting function, can be obtained by applying Bayes rule:

$$w[n] = p(\text{“clip”} \mid |\hat{x}_p[n]|) = \frac{p\left(\text{“clip”}\right) p\left(|\hat{x}_p[n]| \mid \text{“clip”}\right)}{p(|\hat{x}_p[n]|)} \tag{35}$$

$$w[n] = \frac{f_c(|\hat{x}_p[n]|)}{f_c(|\hat{x}_p[n]|) + f_{\overline{c}}(|\hat{x}_p[n]|)}, \tag{36}$$

with $f_c(\cdot)$ and $f_{\overline{c}}(\cdot)$ given by (27) and (28). We note that, even though this is a lengthy expression, its impact on execution speed is small because it does not involve matrix operations.

Algorithm 2 Clip parameter estimation

Input: $|\hat{\mathbf{x}}_\mathbf{p}|$, σ_z
$\mu_2 = \frac{1}{N}\sum_{n=0}^{N-1}|\hat{x}_p[n]|^2$; $\mu_4 = \frac{1}{N}\sum_{n=0}^{N-1}|\hat{x}_p[n]|^4$
$\beta = \left(\mu_4 + 2\sigma_z^2\left(\sigma_z^2 - 2\mu_2\right)\right) / \left(2\left(\mu_2 - \sigma_z^2\right)\right)$
Initialize: $\hat{\alpha}^{(0)} = \frac{\max\{|\hat{\mathbf{x}}_\mathbf{p}|\}^2}{\mathrm{E}\left[|\hat{\mathbf{x}}_\mathbf{p}|^2\right]-\sigma_z^2}$, $k = 0$
while $k < k_{\max}$ and $\frac{\left|\alpha^{(k+1)}-\alpha^{(k)}\right|}{\alpha^{(k+1)}} > \mathrm{tol}$ **do**
$\hat{\alpha}^{(k+1)} = \ln\left(\frac{2\beta}{2\beta-1-\hat{\alpha}^{(k)}+\sqrt{\left(1+\hat{\alpha}^{(k)}\right)^2-4\hat{\alpha}^{(k)}\beta}}\right)$
$k := k+1$
end while
$\hat{\sigma}_x = \frac{\mu_2-\sigma_z^2}{1-e^{-\hat{\alpha}^{(k)}}}$; $\hat{\tau} = \hat{\sigma}_x\sqrt{\hat{\alpha}^{(k)}}$; $\hat{\rho} = e^{-\hat{\alpha}^{(k)}}$; $\mathbf{w}$ from (36)
Output: $\hat{\sigma}_x$, $\hat{\tau}$, $\hat{\rho}$, $\mathbf{w}$

VI. Computational Complexity

We assume the use of a Viterbi decoder and a convolutional code with constraint length C.

The computational complexity of the proposed Weighted IHT algorithm comes mainly from the calculation of the pseudo-inverse in (18) and the need to apply channel decoding once in order to compute the code-based reliability measure. This yields a complexity of $\mathcal{O}\left(MK^2 + N2^C\right)$.

The SABMP algorithm needs to compute this pseudo-inverse several times, but the implementation in Section IV of [7] reduces its complexity to $\mathcal{O}(MK)$ by exploiting the previous results. Even then, the calculation needs to be done N times for each possible support size up to K, making the overall complexity of this algorithm $\mathcal{O}\left(NMK^2\right)$.

Note that, even if the complexity of WIHT has an exponential term 2^C, the constraint length of the code is a fixed term. For a typical value of $C = 7$ and $N = 512$, even with a small sparsity rate of 5%, we have $K = 25$, and both SABMP and WIHT need more reliable carriers M than active taps K. Even a minimalistic choice of $M = 25$ makes $MK^2 = 31250 \gg 128 = 2^C$. Therefore, in practical scenarios, the execution time of SABMP is at least two orders of magnitude longer than that of the proposed WIHT.

Finally, the sCPANC technique performs an IFFT of order N, channel decoding, channel encoding and an FFT of order N in each one of the I iterations. The complexity of these operations amounts to $\mathcal{O}\left(IN\log N + IN2^C\right)$.

VII. Simulation Results

In this section, we use simulations to compare our proposed WIHT technique with the existing methods SABMP [7] and sCPANC [6]. Two reference bounds are also given for comparison: the unrecovered case (no clip removal) and the oracle-LS case. The latter corresponds to the receiver perfectly knowing the support set, and applying a least-squares solution over it.

All experiments simulate an OFDM system with $N = 512$ subcarriers and 16-QAM modulation. The number of reliable

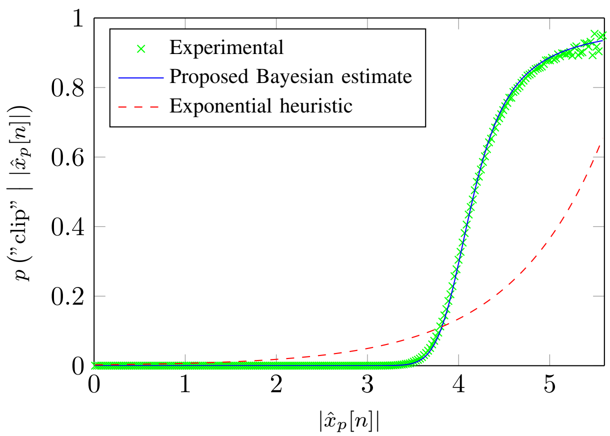

Figure 2. Comparison of estimators of the priors $w[n] = p(\text{“clip”} \mid |\hat{x}_p[n]|)$.

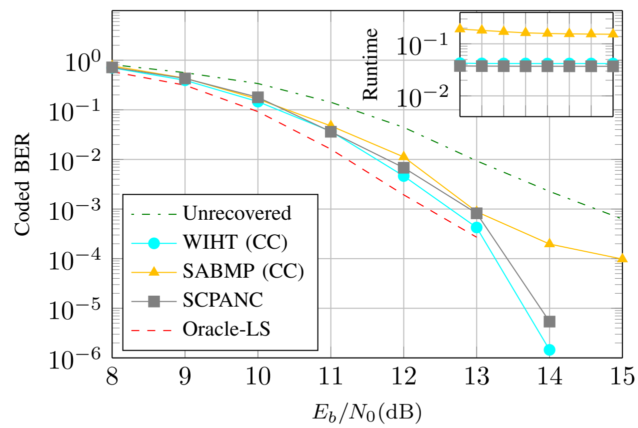

Figure 3. Coded BER performance of recovery techniques (16-QAM, $N = 512$ carriers, $M = 350$ reliable carriers, threshold $\tau = 1.4\sigma_x$, CC rate 3/4).

carriers was empirically chosen to be $M = 350$. Gaussian channels with unit variance and length 4 were generated, and the E_b/N_0 parameter is defined as:

$$\frac{E_b}{N_0} = 10 \log \left(\frac{\mathrm{E}\left[\|\mathbf{H}\mathbf{x}_\mathbf{p}\|_2^2 \right]}{\sigma_z^2} \frac{1}{R_{cc} \log_2 Q} \right), \tag{37}$$

where R_{cc} is the convolutional code rate.

A. Experiment I: Accuracy of the estimation of priors

First, a system with clipping threshold $\tau = 1.4\sigma_x$ was simulated. The proposed Bayesian estimate of the prior function $w[n] = p(\text{“clip”} \mid |\hat{x}_p[n]|)$ was compared to the exponential weighting from [7] and to the experimental probability function (obtained by counting how many of the received symbols had actually been clipped in each amplitude bin). Figure 2 shows the results. The proposed Bayesian estimator agrees almost perfectly with the experimental probability, while the exponential heuristic only provides a very coarse estimate.

B. Experiment II: Comparison of clipping recovery techniques

For the second experiment, the clipping threshold was again $\tau = 1.4\sigma_x$. The three considered techniques were compared in terms of bit error rate (BER) over E_b/N_0 on a coded system with CC rate $3/4$ and constraint length $C = 7$. The sCPANC algorithm was run for $I = 3$ iterations, to keep its execution time in the same order of magnitude as WIHT. For the SABMP technique, the exponential weighting and the LR reliability measure were used, as in [7], while the WIHT technique was run with the proposed code-based reliability measure and Bayesian weighting function.

The results are given in Figure 3. The runtime subgraph shows that SABMP is even slower than the iterative channel decoding performed by sCPANC. WIHT is the fastest technique. Furthermore, thanks to the code-based reliability measure and the Bayesian estimation framework proposed in this work, WIHT also outperforms both sCPANC and SABMP in BER terms. This makes our proposed WIHT technique the best option for clipping estimation in OFDM signals.

VIII. Conclusion

A Weighted Iterative Hard Thresholding algorithm for mitigating the nonlinear effects of power amplifiers is proposed. Our method achieves better performance and faster execution speed than the existing techniques SABMP and sCPANC. Additionally, unlike sCPANC, our technique does not require knowledge of the PA model at the receiver, which is not readily available in the uplink of a mobile communications system.

The extension of the model to a multi-user scenario with per-subcarrier bit and power allocation, as well as to generic Filter Bank Multicarrier modulations, are left for future work.

References

[1] T. Jiang and Y. Wu, “An overview: Peak-to-average power ratio reduction techniques for OFDM signals,” *IEEE Transactions on Broadcasting*, vol. 54, no. 2, pp. 257–268, June 2008.

[2] T. Jiang and G. Zhu, “OFDM peak-to-average power ratio reduction by complement block coding scheme and its modified version,” in *The 60th IEEE Vehicular Technology Conference 2004 Fall*, September 2004, pp. 448–451.

[3] S. H. Muller and J. B. Huber, “OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences,” *IEEE Electronics Letters*, vol. 33, no. 5, pp. 36–69, February 1997.

[4] X. B. Wang, T. T. Tjhung, , and C. S. Ng, “Reduction of peak-to-average power ratio of OFDM system using a companding technique,” *IEEE Transactions on Broadcasting*, vol. 45, no. 3, pp. 303–307, September 1999.

[5] P. of 1999 Asilomar Conference, Ed., *Peak power reduction in OFDM and DMT via active channel modification*, 1999.

[6] L. G. Baltar, S. Dierks, F. H. Gregorio, J. E. Cousseau, and J. A. Nossek, “OFDM receivers with iterative nonlinear distortion cancellation,” in *Signal Processing Advances in Wireless Communications (SPAWC), 2010 IEEE Eleventh International Workshop on*, June 2010, pp. 1–5.

[7] A. Ali, A. Al-Rabah, M. Masood, and T. Al-Naffouri, “Receiver-based recovery of clipped OFDM signals for PAPR reduction: A Bayesian approach,” *IEEE Access*, vol. 2, Oct 2014.

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

[9] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollarab, *A Soft-Input Soft-Output Maximum A Posteriori (MAP) Module to Decode Parallel and Serial Concatenated Codes*. NASA, November 1996.

[10] E. J. Candes and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Communications on Pure and Applied Mathematics*, 2006.

[11] T. Blumensath and M. E. Davies, “Iterative hard thresholding for compressed sensing,” *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.