Preprint

This is the submitted version of a paper published in *IEEE Transactions on Signal Processing*.

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Joint Channel and Clipping Level Estimation for OFDM in IoT-based Networks

Ehsan Olfat, *Student Member, IEEE,* Mats Bengtsson, *Senior Member, IEEE,*

Abstract—We consider scenarios such as IoT-based 5G or IoT-based Machine Type Communication (MTC), where a low-cost low-power transmitter communicates with a high-quality receiver. Then, digital pre-distortion of the non-linear power amplifier may be too expensive. In order to investigate the feasibility of receiver-side compensation of the transmitter RF impairments, we study joint maximum-likelihood estimation of channel and clipping level in multi-path fading OFDM systems. In particular, we propose an alternative optimization algorithm which uses frequency-domain block-type training symbols, and prove that this algorithm always converges, at least to a local optimum point. Then, we calculate the Cramér-Rao lower bound, and show that the proposed estimator attains it for high Signal-to-Noise Ratios. Finally, we perform numerical evaluations to illustrate the performance of the estimator, and show that iterative decoding can be done using the estimated channel and clipping level with almost the same performance as a genie-aided scenario, where the channel and clipping level are perfectly known.

Index Terms—OFDM, clipping, channel, estimation.

I. INTRODUCTION

ORTHOGONAL frequency division multiplexing (OFDM) is widely recognized as one of the key techniques for high data rate communications in wireless networks. Due to its capability to mitigate multipath effect and eliminate inter-symbol interference, OFDM has been utilized for current technologies, including Long Term Evolution (LTE) or LTE-advanced, local area networks (802.11a/g/n), digital broadcasting services like Digital Audio Broadcasting (DAB) or Terrestrial Digital Video Broadcasting (DVB-T) and internet access over copper wires like Asymmetric Digital Subscriber Line (ADSL) and more recently Very High Speed Digital Subscriber Line (VDSL). It is also considered as a promising candidate for future cellular networks such as 5G [1]. The massive interest in OFDM has resulted in tremendous research for making the real systems using OFDM more reliable and less expensive in practice.

Despite its impressive qualities, OFDM signals suffer from large envelope fluctuations producing a high Peak-to-Average Power Ratio (PAPR). This leads to an unwelcome tradeoff between linearity of the transmitted signal and the cost of the High Power Amplifier (HPA). A high PAPR causes the HPA to work in the saturation region which introduces nonlinear distortion at the transmitter output. One possible way to deal with the high power peaks of OFDM signals and the saturation of the HPA is to consider a high Input Back-Off (IBO).

Ehsan Olfat and Mats Bengtsson are with the Department of Signal Processing at Electrical Engineering School, KTH Royal Institute of Technology, Stockholm, Sweden, e-mail: {ehsan.olfat, mats.bengtsson}@ee.kth.se.

However, using a high IBO decreases the efficiency of HPA due to the fact that the working point of the power amplifier gets far from its saturation point. One important fact is that for a given supply voltage level, HPAs with a large linear dynamic range are less efficient than the ones with a smaller linear dynamic range [2]. As a result, OFDM systems need transmitting and receiving parts with a high dynamic range, which leads to expensive RF components. Indeed, the low efficiency and power utilization factor boost hardware cost as well as power consumption [2]. It is also one main source of concern for future wireless networks, when applying OFDM to millimeter Wave (mmWave) spectrum in Internet of Thing (IoT)-based 5G where constructing efficient HPAs are more difficult and costly at those frequencies [1], or to IoT-based Machine Type Communications (MTC) when devices need to be low cost and low power consuming [3]. For these reasons, several techniques to mitigate the high PAPR of OFDM signals have been proposed in the literature (see [4] and the references within).

The problem of dealing with the distortion introduced by the HPA nonlinearity has captivated a lot of attention during last twenty years and a myriad of works have been done on this topic. Almost all of these works fall into three categories: 1) Using pre-distorters at the transmitter side, 2) signal design to reduce PAPR and 3) distortion mitigating at the receiver side.

A common approach is linearizing a HPA by using a digital pre-distorter, deployed as a serial component prior to the HPA at the transmitter side. A pre-distorter attempts to pre-distort the signal in a way that the overall cascaded system of the pre-distorter and HPA behaves as a linear system. A good survey on these methods is provided in [5]–[11]. However, pre-distorting methods increase the complexity of the transmitter as well as its cost. The second approach is designing the transmitted signal by means of using coding and scrambling to reduce PAPR. A good survey on these methods is provided in [4]. However, using such methods also increases the complexity and cost at the transmitter side. The third approach to deal with the HPA's nonlinearity was first suggested in [12], [13] for OFDM signals. This technique aims to estimate the distortion created by the HPA, and mitigate it at the receiver side in an iterative manner. However, it requires perfect knowledge of the HPA model as well as its parameters at the receiver side which is practically difficult. The performance of this iterative approach integrated with HPA model estimation at the receiver side has been investigated [14]. A further extension was the introduction of joint cancellation of nonlinear distortion for multiple transmitters [15]. Moreover, extensions to SC-OFDM

was also studied [16] and [17]. To the best of our knowledge, there is no work in the literature which considers joint channel and nonlinearity parameter estimation.

In this paper, we propose algorithms to jointly estimate channel and Clipping Amplitude (CA), when a limiter (clipper) is deployed as the nonlinearity model. The importance of this model comes from the fact that it can be used in different scenarios. One way to reduce the PAPR at the transmitter side is the intentional clipping of the OFDM signal. To model this clipping, a limiter is used in the literature. Moreover, a limiter itself can be considered as a simplified yet useful model of nonlinear HPAs. Furthermore, even if a pre-distorter is used at the transmitter side the cascaded combination of that and the HPA often well approximated by a limiter. Considering low cost devices in IoT-based networks, clipping the high PAPR signal by a limiter can be a very promising approach to reduce the cost of having highly efficient HPAs. Furthermore, clipping the high PAPR signal increases the battery life, thanks to the resulting increased power efficiency of the HPA.

Estimating the nonlinear parameter at the receiver side is important, especially in IoT-based MTC because there are a lot of sensors that need to send some information occasionally. Then, at each time of transmission, the receiver (e.g. a fusion center or a BTS) needs to estimate the CA to have an updated value which will be deployed in the detection stage. Those CAs for different nodes may change during time because of environmental conditions such as temperature, humidity, and unstable power supply. Furthermore, for low cost devices, the manufacturing variations from unit to unit may be fairly large and it would be too expensive to measure and calibrate each unit separately. Moreover, electronic devices and components suffer from process of aging, so even if this nonlinearity parameter was known at the beginning, it would not later. Therefore, these nonlinearity parameters will become unknown to the transmitter during time and hence there is an essential need to estimate them at the receiver side occasionally. There are some works in the literature in which the estimation of nonlinear parameters is studied [18]–[21]. In all of these works the multi-path channel has been neglected or assumed to be perfectly known at the receiver side. In fact, the channel estimation in presence of unknown nonlinearity parameters is a challenging task. In this paper the CA is assumed to be a priory unknown to the transmitter (and the receiver), and the receiver uses an alternating optimization algorithm to give the jointly optimal estimate of the channel taps and the CA. Once the CA and the channel has been estimated, the system uses them to detect the transmitted symbols by deploying the iterative detection method proposed in [12].

The remainder of this paper is organised as follows. Section II provides the detailed description of the considered system model in this paper. Section III discusses the proposed algorithms of jointly estimating channel and clipping level. Section IV discusses the performance of the estimators and provides theoretical lower bounds on their performance. Section IV depicts numerical and simulation results and illustrates the performance of the proposed method. Finally, conclusions are drawn in Section V.

II. System Model

We consider an OFDM system with N number of sub-carriers, in which $\mathbf{s} = [s_0, \ldots, s_{N-1}]^\mathsf{T}$ is the frequency-domain symbol vector selected from a constellation such as QAM. The time-domain symbols are obtained by taking the Inverse Discrete Fourier Transform (IDFT) from the frequency-domain symbols as:

$$x_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s_k e^{j \frac{2\pi kn}{N}}, \quad n = 0, \ldots, N-1. \quad (1)$$

We can rewrite (1) in matrix form as:

$$\mathbf{x} = \mathbf{F}^\mathsf{H} \mathbf{s}, \quad (2)$$

where $\mathbf{F}$ is the $N \times N$ unitary Discrete Fourier Transform (DFT) matrix. The channel is slow-fading with $L + 1$ taps ($L \ll N$), denoted as $\mathbf{h} = [h_0, \ldots, h_L]^\mathsf{T}$. $g(.\,; A)$ is the limiter nonlinearity with the following amplitude modulation (AM)/AM and AM/phase modulation (PM) conversion characteristics:

$$F[r] = \begin{cases} r, & r \leq A \\ A, & r > A, \end{cases} \quad (3)$$

$$\Phi[r] = 0, \quad (4)$$

where r is the magnitude of the limiter input signal, and A is the CA. Combining AM/AM and AM/PM conversion characteristics, $g(u; A)$ as a function of complex scalar u (and parameterized by A) can be written as

$$g(u; A) = \begin{cases} u, & |u| \leq A \\ A e^{j \arg(u)}, & |u| > A. \end{cases} \quad (5)$$

The output of the limiter is $\mathbf{z} = g(\mathbf{x}; A)$, in which $g(.\,; A)$ is taken element-wise. However, it is difficult to directly work with the output of the limiter. To bypass this difficulty, we invoke the technique described in [22], where the output of the limiter can be represented in a linear fashion by introducing N augmented binary variables c_n for $n \in \{0, \ldots, N-1\}$ indicating whether the sample at time n has been clipped ($c_n = 1$) or not ($c_n = 0$), i.e.

$$c_n = \begin{cases} 1 & r_n > A \\ 0 & r_n \leq A, \end{cases} \quad (6)$$

which leads us to

$$z_n = (1 - c_n)x_n + A c_n e^{j\phi_n}, \quad n = 0, \ldots, N-1, \quad (7)$$

where $x_n = r_n \exp(j\phi_n)$. It is easier to represent (7) in vector form as:

$$\mathbf{z} = (\mathbf{1} - \mathbf{c}) \odot \mathbf{x} + A\mathbf{c} \odot e^{j\boldsymbol{\phi}}, \quad (8)$$

where $\mathbf{1}$ denotes an all one vector, $e^{j\boldsymbol{\phi}} = [e^{j\phi_0}, \ldots, e^{j\phi_{N-1}}]^\mathsf{T}$ denotes the phase vector of $\mathbf{x}$, and $\odot$ denotes the Hadamard product. Also, note that $\mathbf{x} = \mathbf{r} \odot e^{j\boldsymbol{\phi}}$.

To remove the Inter-Symbol Interference (ISI), a Cyclic Prefix (CP) with a length L_{cp} ($\geq L$) is pre-added to the time-domain symbols at the transmitter and is removed at the receiver. After the process of adding and removing CP,
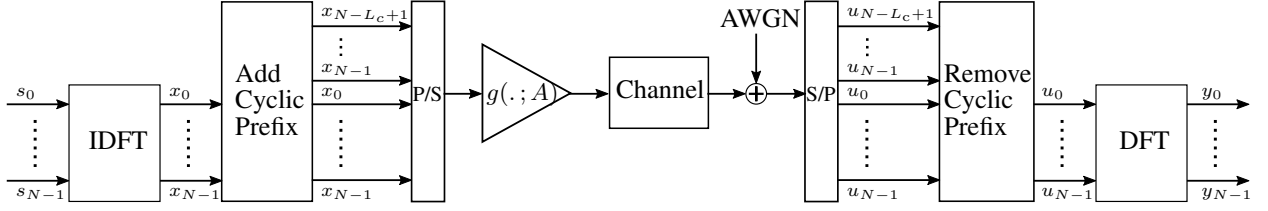
Fig. 1: The system model.

the matrix form time-domain representation of the OFDM transmission can be written as:

$$\mathbf{u} = \mathbf{H}\mathbf{z} + \mathbf{w}, \tag{9}$$

where $\mathbf{H}$ is an $N \times N$ circulant matrix whose first column is $[\mathbf{h}, \mathbf{0}_{N-L-1 \times 1}]$, which represents the circular convolution operator, and $\mathbf{w} = [w_0, \ldots, w_{N-1}]^\mathsf{T}$ is a zero-mean circularly symmetric complex Gaussian noise vector, i.e., $\mathbf{w} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$. Taking the DFT from both sides of (9), we obtain the frequency-domain representation of the OFDM transmission as:

$$\begin{aligned} \mathbf{y} &= \mathbf{F}\mathbf{u} \\ &= \sqrt{N}\mathbf{D_H}\mathbf{F}\mathbf{z} + \tilde{\mathbf{w}}, \end{aligned} \tag{10}$$

where $\mathbf{D_H}$ is a diagonal matrix with the N-point DFT of $\mathbf{h}$ as its diagonal elements, and $\tilde{\mathbf{w}} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$ is the DFT of time-domain noise vector. In (10), we have used the fact that the circulant matrix $\mathbf{H}$ can be diagonalized by pre- and post-multiplication with N-point DFT and IDFT matrices, i.e., $\mathbf{F}\mathbf{H}\mathbf{F}^\mathsf{H} = \sqrt{N}\mathbf{D_H}$. Let $\breve{\mathbf{F}}$ be a semi-unitary matrix formed by the first $L+1$ columns of $\mathbf{F}$, then $\mathbf{D_H} = \mathrm{diag}(\breve{\mathbf{F}}\mathbf{h})$.

III. JOINT CHANNEL AND CLIPPING AMPLITUDE ESTIMATION

In this section, we propose alternating optimization algorithms to jointly estimate channel and CA. To do this, we use frequency-domain block-type training symbols. Once we have the estimates, we can use them to detect the transmitted symbols in the subsequent OFDM blocks. The intuition behind this algorithm is that CA is a slowly time-varying parameter (much slower than channel variations). Moreover, wireless channels are usually slowly time varying, so the block-fading channel model, which remains the same during the transmission of several OFDM blocks is reasonable.

Using frequency-domain block-type training symbols, the problem of joint Maximum-Likelihood (ML) estimation of channel and CA can be formulated as the following Least-Squares (LS) using (10):

$$\min_{A>0, \mathbf{h}} \left\| \mathbf{y} - \sqrt{N}\mathbf{D_H}\mathbf{F}\mathbf{z} \right\|^2, \tag{11}$$

where $\| \cdot \|$ denotes the Euclidean norm.

A. Clipping Amplitude Estimation Given the Channel

We first solve (11) for A given $\mathbf{h}$. To do so, we first sort the elements in $\mathbf{r}$, and construct a new vector denoted as $\mathring{\mathbf{r}}$. Note that, $\mathring{\mathbf{r}} = \mathbf{P_s}\mathbf{r}$, in which $\mathbf{P_s}$ is the sorting permutation

matrix. Using the sorted vectors, the estimate of CA at the ith iteration can be written as

$$\hat{A} = \arg\min_{A>0} \left\| \mathbf{y} - \sqrt{N}\mathbf{D_H}^{(i)}\mathbf{F}\mathbf{P_s}^\mathsf{T}\mathring{\mathbf{z}} \right\|^2, \tag{12}$$

where

$$\mathring{\mathbf{z}} = \mathring{\mathbf{x}} \odot (\mathbf{1} - \mathring{\mathbf{c}}) + A e^{j\mathring{\boldsymbol{\phi}}} \odot \mathring{\mathbf{c}}, \tag{13}$$

in which $\mathring{\mathbf{x}} = \mathbf{P_s}\mathbf{x}$, $\mathring{\mathbf{c}} = \mathbf{P_s}\mathbf{c}$, and $\mathring{\boldsymbol{\phi}} = \mathbf{P_s}\boldsymbol{\phi}$. Note that $\mathring{\mathbf{c}}$ is constant within each interval $\mathring{r}_{k-1} \leq A \leq \mathring{r}_k$, $k = 0 \ldots N-1$, consequently (13) is piecewise affine in A, therefore the minimization problem (12) can be written as

$$\hat{A} = \arg\min_{A>0} J(A), \tag{14}$$

where

$$J(A) = \begin{cases} J_0(A), & 0 \leq A < \mathring{r}_0 \\ J_1(A), & \mathring{r}_0 \leq A < \mathring{r}_1 \\ \vdots & \vdots \\ J_{N-1}(A), & \mathring{r}_{N-2} \leq A < \mathring{r}_{N-1}, \end{cases} \tag{15}$$

in which

$$\begin{aligned} J_k(A) &= \|\mathbf{y} - \sqrt{N}\mathbf{D_H}\mathbf{F}\mathbf{P_s}^\mathsf{T}\mathring{\mathbf{z}}_k\|^2 \\ &= \|\mathbf{B}(\mathring{\mathbf{c}}_k \odot e^{j\mathring{\boldsymbol{\phi}}})\|^2 A^2 \\ &\quad - 2\Re\left\{ \left[\mathbf{y} - \mathbf{B}(\mathring{\mathbf{x}} \odot \bar{\mathring{\mathbf{c}}}_k) \right]^\mathsf{H} \mathbf{B}(\mathring{\mathbf{c}}_k \odot e^{j\mathring{\boldsymbol{\phi}}}) \right\} A \\ &\quad + \left\| \mathbf{y} - \mathbf{B}(\mathring{\mathbf{x}} \odot \bar{\mathring{\mathbf{c}}}_k) \right\|^2, \end{aligned} \tag{16}$$

where $\mathbf{B} = \sqrt{N}\mathbf{D_H}\mathbf{F}\mathbf{P_s}^\mathsf{T}$, $\mathring{\mathbf{c}}_k$ is an all one vector except its first k elements which are 0, and $\bar{\mathring{\mathbf{c}}}_k = \mathbf{1} - \mathring{\mathbf{c}}_k$. As we can see, for each interval, (16) is a convex quadratic function of A. Consequently, for each interval, we can easily find the minimizer by taking the first derivative of (16) and setting it to zero, and check if the solution lies in that interval or not. If it lies in that interval we have already found the minimizer, otherwise the minimum takes place at the interval's end point with the smaller cost function. Mathematically, for each interval $(\mathring{r}_{k-1}, \mathring{r}_k]$, the minimizer is

$$\hat{A}_k = \begin{cases} \breve{A}_k, & \mathring{r}_{k-1} \leq \breve{A}_k \leq \mathring{r}_k \\ \mathring{r}_{k-1}, & J_k(\mathring{r}_{k-1}) < J_k(\mathring{r}_k) \\ \mathring{r}_k, & \text{o.w.} \end{cases} \tag{17}$$

where

$$\breve{A}_k = \frac{\Re\left\{ \left[\mathbf{y} - \mathbf{B}(\mathring{\mathbf{x}} \odot \bar{\mathring{\mathbf{c}}}_k) \right]^\mathsf{H} \mathbf{B}(\mathring{\mathbf{c}}_k \odot e^{j\mathring{\boldsymbol{\phi}}}) \right\}}{\|\mathbf{B}(\mathring{\mathbf{c}}_k \odot e^{j\mathring{\boldsymbol{\phi}}})\|^2}. \tag{18}$$

To find the global minimizer, we find the minimizer for each interval using (17), and then find the minimizer among

those minimizers. Since $J(A)$ is a continuous function, in each interval, we need just check the point at which the derivative is zero and the rightmost corner point. Therefore, we have Algorithm 1.

Algorithm 1 Estimation of Clipping Amplitude (A)

1: **Inputs:**
$\mathbf{y}$, $\mathbf{r}$ and $\exp(j\phi)$
$\mathbf{F}$ and $\mathbf{D_H}$
$\mathring{\mathbf{c}}_k$ for $k = 0, \ldots, N-1$

2: **Initialize:**
$[\mathring{\mathbf{r}}, \mathbf{P}_s] = \text{sort}(\mathbf{r})$
$e^{j\mathring{\phi}} = \mathbf{P}_s e^{j\phi}$
$\mathbf{B} = \sqrt{N}\mathbf{D_H F P}_s^\mathsf{T}$
$\mathring{r}_{-1} = 0$

3: **for** $k = 0$ to $N-1$ **do**

4: $\quad \breve{A}_k = \dfrac{\Re\left\{\left[\mathbf{y} - \mathbf{B}(\mathring{\mathbf{x}} \odot \bar{\mathring{\mathbf{c}}}_k)\right]^\mathsf{H} \mathbf{B}(\mathring{\mathbf{c}}_k \odot e^{j\mathring{\phi}})\right\}}{\|\mathbf{B}(\mathring{\mathbf{c}}_k \odot e^{j\mathring{\phi}})\|^2}$

5: $\quad$ **if** $\mathring{r}_{k-1} \leq \breve{A}_k \leq \mathring{r}_k$ **then**

6: $\quad\quad \hat{A}_k = \breve{A}_k$

7: $\quad$ **else**

8: $\quad\quad \hat{A}_k = \mathring{r}_k$

9: $\quad$ **end if**

10: $\quad \hat{J}_k = J_k(\hat{A}_k)$

11: **end for**

12: $\hat{k} = \arg \min\limits_{0 \leq k \leq N-1} \hat{J}_k$

13: $\hat{A} = \hat{A}_{\hat{k}}$

B. Channel Estimation given Clipping Amplitude

Here, we solve (11) for $\mathbf{h}$ given A. We can rewrite (11) as:

$$\min_{\mathbf{h}} \left\| \mathbf{y} - \sqrt{N}\,\text{diag}(\mathbf{Fz})\breve{\mathbf{F}}\mathbf{h} \right\|^2, \tag{19}$$

therefore, the LS channel estimate can be computed as:

$$\hat{\mathbf{h}} = \frac{1}{\sqrt{N}}\left(\breve{\mathbf{F}}^\mathsf{H}\text{diag}(\mathbf{Fz})^\mathsf{H}\text{diag}(\mathbf{Fz})\breve{\mathbf{F}} \right)^\dagger \breve{\mathbf{F}}^\mathsf{H}\text{diag}(\mathbf{Fz})^\mathsf{H}\mathbf{y}, \tag{20}$$

where $(\cdot)^\dagger$ denotes the Moore–Penrose pseudo–inverse.

C. Initialization of the Alternating Algorithm

We propose two alternative initialization strategies to the alternating optimization algorithm as follows:

1) *Initializing by the Channel:* Since the value of A is unknown at the beginning, channel can be estimated using the un-clipped version of transmitted time-domain symbols. It is equivalent to putting $\mathbf{x}$ instead of $\mathbf{z}$ in (20), and using $\mathbf{s} = \mathbf{Fx}$. Therefore, the initializing channel estimate is:

$$\hat{\mathbf{h}}^{(0)} = \frac{1}{\sqrt{N}}\left(\breve{\mathbf{F}}^\mathsf{H}\text{diag}(\mathbf{s})^\mathsf{H}\text{diag}(\mathbf{s})\breve{\mathbf{F}} \right)^{-1} \breve{\mathbf{F}}^\mathsf{H}\text{diag}(\mathbf{s})^\mathsf{H}\mathbf{y}. \tag{21}$$

2) *Initializing by the Clipping Amplitude:* By substituting (20) into the cost function in (19), the resulting LS error for A is:

$$\mathcal{E}(A) = \mathbf{y}^\mathsf{H}(\mathbf{I} - \mathbf{T})\mathbf{y}, \tag{22}$$

where

$$\begin{aligned}
\mathbf{T} &= \text{diag}(\mathbf{Fz})\breve{\mathbf{F}}\left(\breve{\mathbf{F}}^\mathsf{H}\text{diag}(\mathbf{Fz})^\mathsf{H}\text{diag}(\mathbf{Fz})\breve{\mathbf{F}}\right)^{-1} \\
&\quad \times \breve{\mathbf{F}}^\mathsf{H}\text{diag}(\mathbf{Fz})^\mathsf{H}.
\end{aligned}$$

By minimizing (22) over A, the CA can be estimated. However, (22) is a complicated function of A and is not convex. Nevertheless, one can find its approximate solution by using grid search methods. Therefore, we can find the initializing CA by solving

$$\hat{A}^{(0)} = \arg \min_{A > 0} \mathcal{E}(A). \tag{23}$$

Due to the non-convexity of $\mathcal{E}(A)$, it is highly possible that even more advanced optimization methods than grid search just find a local minimum.

D. Alternating Optimization Algorithm

Depending on which initialization is used, we have two alternating algorithms denoted as Algorithms 2 and 3:

Algorithm 2 Alternating Optimization with Initializing Channel

---

1: **Inputs:**
$\mathbf{y}$, $\mathbf{r}$ and $\exp(j\phi)$
$\mathbf{F}$
$\mathring{\mathbf{c}}_k$ for $k = 0, \ldots, N-1$

2: **Initialize:**
$[\mathring{\mathbf{r}}, \mathbf{P}_s] = \text{sort}(\mathbf{r})$
$e^{j\mathring{\phi}} = \mathbf{P}_s e^{j\phi}$
$\hat{\mathbf{D}}_\mathbf{H}^{(0)} = \text{diag}(\breve{\mathbf{F}}\hat{\mathbf{h}}^{(0)})$, $\hat{\mathbf{h}}^{(0)}$ is given by (21)
$\mathring{r}_{-1} = 0$
$i = 1$

3: **while** (convergence criteria not met) **do**

4: $\quad \mathbf{B} = \sqrt{N}\hat{\mathbf{D}}_\mathbf{H}^{(i-1)}\mathbf{F P}_s^\mathsf{T}$

5: $\quad$ calculate $\hat{A}^{(i-1)}$ and $\hat{k}^{(i-1)}$ using Algorithm 1

6: $\quad \mathring{\mathbf{z}} = \mathring{\mathbf{x}} \odot (\mathbf{1} - \mathring{\mathbf{c}}_{\hat{k}^{(i-1)}}) + \hat{A}^{(i-1)}e^{j\mathring{\phi}} \odot \mathring{\mathbf{c}}_{\hat{k}^{(i-1)}}$

7: $\quad \mathbf{V} = \text{diag}(\mathbf{F P}_s^\mathsf{T}\mathring{\mathbf{z}})\breve{\mathbf{F}}$

8: $\quad \hat{\mathbf{h}}^{(i)} = (\mathbf{V}^\mathsf{H}\mathbf{V})^\dagger \mathbf{V}^\mathsf{H}\mathbf{y}$

9: $\quad \hat{\mathbf{D}}_\mathbf{H}^{(i)} = \text{diag}(\breve{\mathbf{F}}\hat{\mathbf{h}}^{(i)})$

10: $\quad i = i + 1$

11: **end while**

E. Convergence

Note that both algorithms of alternating optimization are guaranteed to converge to a local optimum because in every iteration we can find the unique optimal solution of the optimization problem [23]. In particular, we have the following proposition:

Proposition 1: *Both algorithms 2 and 3 converge to a local optimum point of the cost function given in* (11) *(or* (12)*).*

Proof: let us denote the cost function in (11) by $J(\mathbf{h}, A)$, then for Algorithm 2, we have

$$\hat{A}^{(i-1)} = \arg \min_{A > 0} J(\hat{\mathbf{h}}^{(i-1)}, A). \tag{24}$$

Algorithm 3 Alternating Optimization with Initializing CA

1: **Inputs:**
   $\mathbf{y}$, $\mathbf{r}$ and $\exp(j\boldsymbol{\phi})$
   $\mathbf{F}$
   $\mathring{\mathbf{c}}_k$ for $k = 0, \ldots, N-1$

2: **Initialize:**
   $[\mathring{\mathbf{r}}, \mathbf{P}_\mathrm{s}] = \mathrm{sort}(\mathbf{r})$
   $e^{j\hat{\phi}} = \mathbf{P}_\mathrm{s} e^{j\phi}$
   $\hat{A}^{(0)}$ by solving (23) using grid search
   $\mathring{r}_{-1} = 0$
   $i = 1$

3: **while** (convergence criteria not met) **do**

4:     $\mathbf{V} = \mathrm{diag}\big(\mathbf{F}\mathbf{P}_\mathrm{s}^\mathsf{T}\mathring{\mathbf{z}}^{(i-1)}\big)\check{\mathbf{F}}$

5:     $\hat{\mathbf{h}}^{(i-1)} = (\mathbf{V}^\mathsf{H}\mathbf{V})^\dagger \mathbf{V}^\mathsf{H}\mathbf{y}$

6:     $\hat{\mathbf{D}}_\mathbf{H}^{(i-1)} = \mathrm{diag}(\check{\mathbf{F}}\hat{\mathbf{h}}^{(i-1)})$

7:     $\mathbf{B} = \hat{\mathbf{D}}_\mathbf{H}^{(i-1)}\mathbf{F}\mathbf{P}_\mathrm{s}^\mathsf{T}$

8:     calculate $\hat{A}^{(i)}$ and $\hat{k}^{(i)}$ using Algorithm 1

9:     $i = i+1$

10: **end while**

Therefore

$$J\big(\hat{\mathbf{h}}^{(i-1)}, \hat{A}^{(i-1)}\big) \leq J\big(\hat{\mathbf{h}}^{(i-1)}, A\big), \ \forall A > 0. \quad (25)$$

Also, note that

$$\hat{\mathbf{h}}^{(i)} = \arg\min_\mathbf{h} J\big(\mathbf{h}, \hat{A}^{(i-1)}\big), \quad (26)$$

hence

$$J\big(\hat{\mathbf{h}}^{(i)}, \hat{A}^{(i-1)}\big) \leq J\big(\mathbf{h}, \hat{A}^{(i-1)}\big), \ \forall \mathbf{h} \in \mathbb{C}^{L+1}. \quad (27)$$

where $\mathbb{C}^{L+1}$ denotes the $(L+1)$–dimensional complex vector space. From (24)–(27), we can conclude

$$J\big(\hat{\mathbf{h}}^{(i)}, \hat{A}^{(i)}\big) \leq J\big(\hat{\mathbf{h}}^{(i-1)}, \hat{A}^{(i-1)}\big), \ i = 1, 2, \ldots \quad (28)$$

Therefore, the objective function is decreasing at each iteration and eventually converges to a local minimum. The same argument is also valid for Algorithm 3. □

F. Computational Complexity

We first compute the worse case computational complexity of Algorithm 1 using big–$\mathcal{O}$ notation. Note that the complexity order of complex-valued and real-valued operations are asymptotically equivalent. The complexity of initialization part is dominated by the matrix multiplication which has complexity of order $\mathcal{O}(N^3)$. Lines 4 and 10 at each iteration have complexity of order $\mathcal{O}(N^2)$, hence the total complexity of N iterations is of order $\mathcal{O}(N^3)$. The min function at line 12 has complexity of order $\mathcal{O}(N)$, therefore Algorithm 1 has complexity of order $\mathcal{O}(N^3)$.

However, the computational complexity of Algorithm 1 can be reduced by not forming $\mathbf{B}$ explicitly in the initialization section, and instead computing (18) in line 4 as the following:

$$\check{A}_k = \frac{\Re\left\{\left[\mathbf{y} - \sqrt{N}\mathbf{D}_\mathbf{H}\mathrm{FFT}\{\mathbf{x} \odot \bar{\mathbf{c}}_k\}\right]^\mathsf{H} \mathbf{D}_\mathbf{H}\mathrm{FFT}\{\mathbf{c}_k \odot e^{j\phi}\}\right\}}{\sqrt{N}\|\mathbf{D}_\mathbf{H}\mathrm{FFT}\{\mathbf{c}_k \odot e^{j\phi}\}\|^2} \quad (29)$$

where $\mathbf{c}_k = \mathbf{P}_\mathrm{s}^\mathsf{T}\mathring{\mathbf{c}}_k$, $\bar{\mathbf{c}}_k = \mathbf{1} - \mathbf{c}_k$, and $\mathrm{FFT}\{\cdot\}$ denotes the unitary N–point Fast Fourier Transform (FFT). This will reduce the complexity of Algorithm 1 from $\mathcal{O}(N^3)$ to $\mathcal{O}(N^2 \log_2 N)$, since the FFT with complexity $\mathcal{O}(N \log_2 N)$ is repeated N times.

In Algorithm 2, the complexity of the initialization part is dominated by (21), which has complexity of order $\mathcal{O}(N^2)$. On the other hand, the complexity of the initialization part in Algorithm 3 is dominated by the grid search algorithm which has complexity of order $\mathcal{O}(N^2)$ for function evaluation for each grid point multiplied by the number of grid points (N_GS) plus the complexity of the min function. Therefore, overall, the complexity of the initialization part of Algorithm 3 is of order $\mathcal{O}(N_\mathrm{GS}N^2) = \mathcal{O}(N^2)$. Both Algorithms 2 and 3 have the same order of complexity per iteration, which is dominated by Algorithm 1 which has complexity of order $\mathcal{O}(N^2 \log_2 N)$. According to this, both Algorithms 2 and 3 have the same worse case asymptotic order of complexity.

IV. CRAMÉR-RAO LOWER BOUND

The Cramér Rao Lower Bound (CRLB) expresses a lower bound on the achievable variance of unbiased estimators. The ML estimator asymptotically achieves the CRLB under some regularity conditions [24]. Here, the estimation parameters comprise of complex-valued channel taps and a real-valued CA. Therefore, to calculate the CRLB, we denote the parameter vector as $\boldsymbol{\theta} = [\mathbf{h}^\mathsf{T}, \mathbf{h}^\mathsf{H}, A]^\mathsf{T}$, which is used to derive the complex Fisher Information Matrix (FIM). The logarithm of the Probability Density Function (PDF) of the frequency-domain observation vector $\mathbf{y}$ given in (10) can be written as:

$$\log p(\mathbf{y}; \boldsymbol{\theta}) = -\frac{1}{\sigma^2}\big\|\mathbf{y} - \sqrt{N}\mathbf{D}_\mathbf{H}\mathbf{F}\mathbf{z}\big\|^2 + \mathrm{const.}, \quad (30)$$

where $\mathrm{const.}$ comprises the terms which are independent of the estimation parameter vector $\boldsymbol{\theta}$. The log PDF (30) is a differentiable function of $\mathbf{h}$ in the whole parameter space. On the other hand, the log PDF is a smooth function of A in $\mathbb{R}^+ = [0, \infty]$ except at the points $\{r_0, \ldots, r_{N-1}\}$. In this case, the regularity conditions just hold for the set $\mathbb{S}_A^+ = \mathbb{R}^+ \setminus \{r_0, \ldots, r_{N-1}\}$, and not for the whole parameter space. To solve this problem, we limit our attention just to the set $\mathbb{S}_A = \mathbb{S}_A^+ \setminus \{0\}$, partition it into disjoint intervals $(\mathring{r}_{k-1}, \mathring{r}_k)$, $k = 0 \ldots N-1$, and calculate the constrained CRLB for each interval. To do this, we use the following lemma:

Lemma 1. *The constrained CRLB is identical to the unconstrained CRLB at the regular points [25].*

Using Lemma 1, we can simply find the derivative of $\mathbf{z}$ given in (8) with respect to A at any regular point of the parameter space, which will be used in the CRLB derivation, as follows:

$$\frac{\partial \mathbf{z}}{\partial A} = \mathbf{c} \odot e^{j\phi}, \ \forall A \in \mathbb{S}_A \quad (31)$$

Now, we can use the following lemma to calculate the CRLB for all the regular points belonging to set $\mathbb{S}_A$:

Lemma 2. *For the mixed complex-valued and real-valued estimation parameters $\boldsymbol{\theta} = [\boldsymbol{\theta}_1^\mathsf{T}, \boldsymbol{\theta}_1^\mathsf{H}, \boldsymbol{\theta}_2^\mathsf{T}]^\mathsf{T}$, where $\boldsymbol{\theta}_1 \in \mathbb{C}^{n_1}$,*

$\theta_2 \in \mathbb{R}^{n_2}$, *the unconstrained complex FIM and its corresponding CRLB have the following structure [26, Lemma 1], [27]:*

$$\mathbf{I}(\boldsymbol{\theta}) = \left[\begin{array}{cc|c} \mathbf{C}_1 & \mathbf{C}_2^* & \mathbf{P} \\ \mathbf{C}_2 & \mathbf{C}_1^* & \mathbf{P}^* \\ \hline \mathbf{P}^{\mathsf{H}} & \mathbf{P}^{\mathsf{T}} & \mathbf{Q} \end{array} \right], \qquad (32)$$

where

$$\mathbf{C}_1 = \mathbb{E}\left\{ \frac{\partial \log p(\mathbf{y};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1^*} \frac{\partial \log p(\mathbf{y};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1^*}^{\mathsf{H}} \right\}, \qquad (33)$$

$$\mathbf{C}_2 = \mathbb{E}\left\{ \frac{\partial \log p(\mathbf{y};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1} \frac{\partial \log p(\mathbf{y};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1}^{\mathsf{T}} \right\}, \qquad (34)$$

$$\mathbf{P} = \mathbb{E}\left\{ \frac{\partial \log p(\mathbf{y};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1^*} \frac{\partial \log p(\mathbf{y};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2}^{\mathsf{T}} \right\}, \qquad (35)$$

$$\mathbf{Q} = \mathbb{E}\left\{ \frac{\partial \log p(\mathbf{y};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2} \frac{\partial \log p(\mathbf{y};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_2}^{\mathsf{T}} \right\}. \qquad (36)$$

Then

$$\mathrm{CRLB}(\boldsymbol{\theta}_2) = \left[\mathbf{Q} - 2\Re\left\{ \mathbf{P}^{\mathsf{H}}\boldsymbol{\Delta}\mathbf{P} + \mathbf{P}^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{P} \right\} \right]^{-1} \quad (37)$$

$$\mathrm{CRLB}(\boldsymbol{\theta}_1) = \boldsymbol{\Delta} + (\boldsymbol{\Delta}\mathbf{P} + \boldsymbol{\Delta}^*\mathbf{P}^*)\,\mathrm{CRLB}(\boldsymbol{\theta}_2)$$
$$\times \left(\mathbf{P}^{\mathsf{H}}\boldsymbol{\Delta}^{\mathsf{H}} + \mathbf{P}^{\mathsf{T}}\boldsymbol{\Delta}^{\mathsf{T}} \right), \quad (38)$$

where

$$\boldsymbol{\Delta} = \left(\mathbf{C}_1 - \mathbf{C}_2^*\mathbf{C}_1^{*-1}\mathbf{C}_2 \right)^{-1}, \qquad (39)$$

$$\boldsymbol{\Lambda} = -\mathbf{C}_1^{*-1}\mathbf{C}_2\mathbf{C}_1^{-1}. \qquad (40)$$

In our case, $\boldsymbol{\theta}_1 = \mathbf{h}$, and $\boldsymbol{\theta}_2 = \theta_2 = A$, and we have:

$$\frac{\partial \log p(\mathbf{y};\boldsymbol{\theta})}{\partial \mathbf{h}^*} = \frac{\sqrt{N}}{\sigma^2} \check{\mathbf{F}}^{\mathsf{H}} \mathrm{diag}(\mathbf{Fz})^{\mathsf{H}} (\mathbf{y} - \sqrt{N}\mathbf{D_H}\mathbf{Fz}). \quad (41)$$

Therefore, using Lemma 2, we find

$$\mathbf{C}_1 = \frac{N}{\sigma^2} \check{\mathbf{F}}^{\mathsf{H}} \mathrm{diag}(\mathbf{Fz})^{\mathsf{H}} \mathrm{diag}(\mathbf{Fz}) \check{\mathbf{F}}, \qquad (42)$$

$$\mathbf{C}_2 = \mathbf{0}, \qquad (43)$$

$$\mathbf{P} = \mathbf{p} = \frac{N}{\sigma^2} \check{\mathbf{F}}^{\mathsf{H}} \mathrm{diag}(\mathbf{Fz})^{\mathsf{H}} \mathbf{D_H}\mathbf{F}(\mathbf{c} \odot e^{j\phi}), \quad (44)$$

$$\mathbf{Q} = q = \frac{2N}{\sigma^2} \left\| \mathbf{D_H}\mathbf{F}(\mathbf{c} \odot e^{j\phi}) \right\|^2, \qquad (45)$$

then

$$\boldsymbol{\Delta} = \mathbf{C}_1^{-1}, \qquad (46)$$

$$\boldsymbol{\Lambda} = \mathbf{0}. \qquad (47)$$

Therefore

$$\mathrm{CRLB}(A) = \frac{1}{q - 2\mathbf{p}^{\mathsf{H}}\mathbf{C}_1^{-1}\mathbf{p}}, \qquad (48)$$

and

$$\mathrm{CRLB}(\mathbf{h}) = \mathbf{C}_1^{-1} + \frac{1}{q - 2\mathbf{p}^{\mathsf{H}}\mathbf{C}_1^{-1}\mathbf{p}} \mathbf{C}_1^{-1}\mathbf{p}\mathbf{p}^{\mathsf{H}}\mathbf{C}_1^{-\mathsf{H}}. \quad (49)$$

V. SIMULATION RESULTS

In this section, the performance of the proposed estimation algorithm is investigated by computer simulation. The number of sub-carriers (N) considered here are 128, 256 and 512. Moreover, the subcarrier modulation is 16QAM. The Rayleigh fading channel model used here is the exponential channel model (IEEE 802.11b model) [28], which has $L+1$ complex Gaussian distributed taps h_l having a mean power of $\sigma_l^2 = \mathbb{E}\{|h_l|^2\} = \sigma_0^2 e^{-T_s/\tau_{\mathrm{rms}}l}$ for $l = 0,\dots,L$ in which σ_0 is chosen such that $\sum_{l=0}^{L} \sigma_l^2 = 1$, $T_s = 50\,ns$ is the sampling rate, and $\tau_{\mathrm{rms}} = 30\,ns$ and $150\,ns$ are the RMS delays leading to channel lengths $L + 1 = 7$ and 31, respectively. Our model follows the slow-time varying condition in which the channel is fixed during several OFDM blocks. We generate N symbols drawn randomly from the 16QAM constellation to construct the block-type pilots, since it is not needed to design the pilot symbols. Furthermore, in all simulations we keep the average energy of the input signal to the limiter equal to 1. Therefore, the clipping level is defined as $\mathrm{CL} = 20\log_{10}(A)[\mathrm{dB}]$. Moreover, to simulate the effect of the receiver noise, we keep the average energy of the input signal to the limiter fixed, and decrease the variance of the noise. For Symbol Error Rate (SER) simulations, we use the iterative detection algorithm introduced in [12], but with the estimated channel and CA as follows:

Algorithm 4 Iterative Detection Algorithm [12]

1: **Inputs:**
    $\mathbf{y}$, $\mathbf{F}$, and N_q
    $\hat{\mathbf{h}}$ and $\hat{A}$ calculated by Algorithm 2 (or 3)

2: **Initialize:**
    $\hat{\mathbf{D}}_{\mathbf{H}} = \mathrm{diag}(\check{\mathbf{F}}\hat{\mathbf{h}})$
    $\kappa = 1 - e^{-\hat{A}^2} + \hat{A}\sqrt{\frac{\pi}{2}}\mathrm{erfc}(\hat{A})$
    $\mathbf{d}^{(0)} = \mathbf{0}$

3: **for** $i = 1$ to N_q **do**

4:     $\hat{\mathbf{s}}^{(i)} = \left\langle \frac{1}{\kappa}\left(\frac{1}{\sqrt{N}}\hat{\mathbf{D}}_{\mathbf{H}}^{-1}\mathbf{y} - \mathbf{d}^{(i-1)} \right) \right\rangle$

5:     $\hat{\mathbf{x}}^{(i)} = \mathbf{F}^{\mathsf{H}}\hat{\mathbf{s}}^{(i)}$

6:     $\mathbf{d}^{(i)} = \mathbf{F}\left(g(\hat{\mathbf{x}}^{(i)}; \hat{A}) - \kappa\hat{\mathbf{x}}^{(i)} \right)$

7: **end for**

---

where N_q is the maximum number of algorithm iterations, κ is the optimal linear gain for the limiter (calculated, e.g., in [29]), $\mathrm{erfc}(u) = 2/\sqrt{\pi}\int_u^\infty e^{-t^2}\,\mathrm{d}t$, and $\langle \cdot \rangle$ denotes element-wise hard detecting function.

Figs. 2 and 3 show the Normalized Mean Square Error (NMSE) performance of estimating CA (A) and channel taps ($\mathbf{h}$), when $L + 1 = 7$, $N = 128$ and $\mathrm{CL} = 1\,\mathrm{dB}$, respectively. As we can see from the figures the NMSEs are decreasing by increasing the SNR and both of them reach CRLB in a medium SNR regime. In Figs. 4 and 5, we have increased the number of sub-carriers to $N = 512$, and also the clipping level to $\mathrm{CL} = 3\,\mathrm{dB}$ which corresponds to a HPA with higher dynamic range. Again, in these figures the NMSEs are decreasing and both of them reach CRLB in a medium SNR regime. By comparing all these figures, we can see that for higher number of sub-carriers the NMSE reaches its CRLB faster. This can be justified by the fact that for a given channel length, the higher the number of sub-carriers, the higher the number of clipped observations, which results in more accurate estimates.

Figs. 6 and 7 show the SER performance of Algorithms 2 and 3, integrated with the iterative detection method (Algorithm 4), when $L + 1 = 7$, and $\mathrm{CL} = 1\,\mathrm{dB}$ for $N = 128$ and $N = 512$, respectively. We have used two maximum number
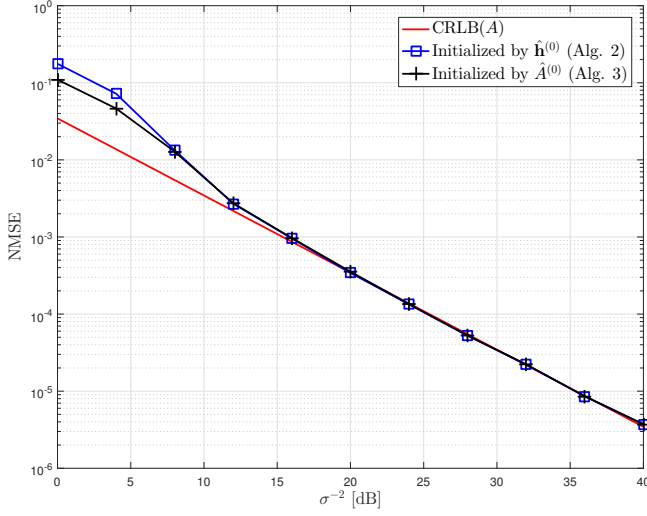
Fig. 2: NMSE performance of the CA estimation using Algorithms 2 and 3, when $L + 1 = 7$, $N = 128$ and CL = 1 dB.
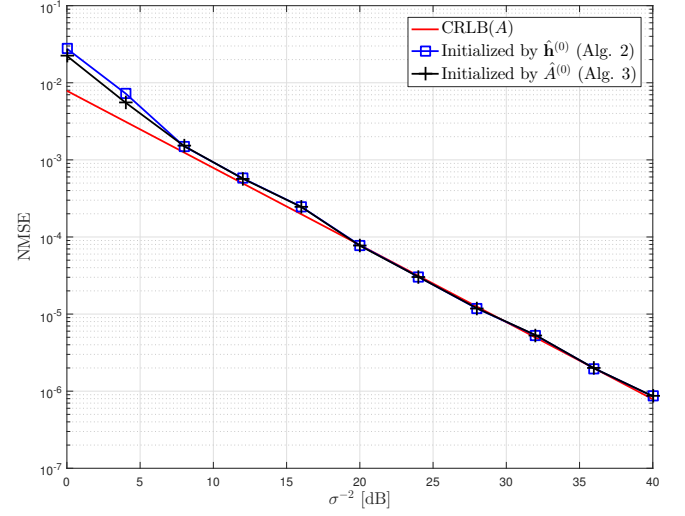

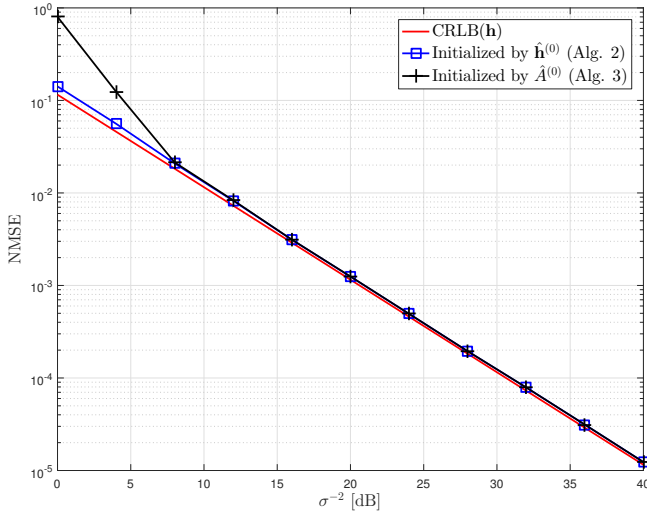Fig. 4: NMSE performance of the CA estimation using Algorithms 2 and 3, when $L + 1 = 7$, $N = 512$ and CL = 3 dB.

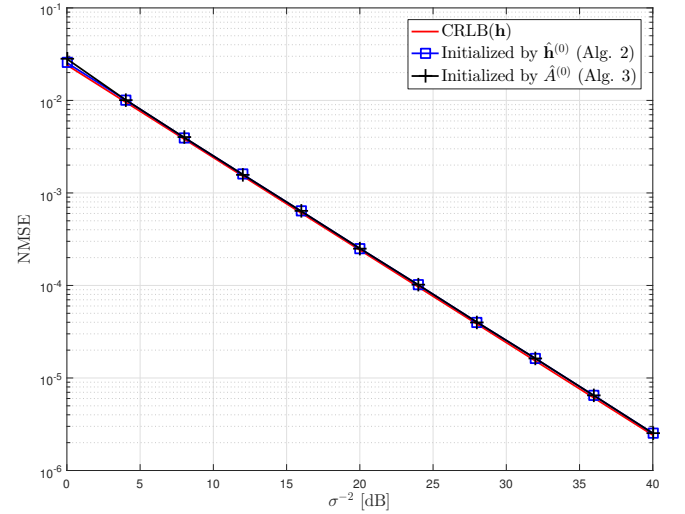Fig. 3: NMSE performance of the channel estimation using Algorithms 2 and 3, when $L+1 = 7$, $N = 128$ and CL = 1dB.

Fig. 5: NMSE performance of the channel estimation using Algorithms 2 and 3, when $L+1 = 7$, $N = 512$ and CL = 3dB.

of iterations for Algorithm 4 as $N_q = 2$ and 3. As illustrated in the figures the SER performance of different schemes when we estimate the CA and the channel taps using Algorithms 2 and 3 almost perfectly coincide with a genie-aided scenario, where the channel and CA are perfectly known at the receiver side.

The results depicted in Fig. 8 shows the SER performance of Algorithms 2 and 3, integrated with the iterative detection method (Algorithm 4), when $L + 1 = 7$, and CL = 3 dB for $N = 256$, where the maximum number of iterations for Algorithm 4 is set to $N_q = 2$. Here, since the CL is higher, the result for N_q's higher than 2 is almost the same as the case when $N_q = 2$, so by increasing N_q, we do not earn any tangible gain. Again, we can see from the figures that the SER of different schemes when we estimate the CA and the channel taps using Algorithms 2 and 3 almost perfectly coincide with the genie-aided case.

Fig. 9 shows the SER performance of Algorithms 2 and 3, integrated with the iterative detection method (Algorithm 4) with $N_q = 3$, when $L + 1 = 31$, CL = 1dB for N = 128. Note that the channel length $L + 1 = 31$ is the maximum allowed length for $L_{cp} = 32$ which is used conventionally for an OFDM system with $N = 128$ number of sub-carriers. In this case the channel frequency selectivity is higher. As can be seen from the figure, both estimation algorithms provide really good results.

Figs. 10 and 11 show the SER performance of Algorithms 2 and 3, integrated with the iterative detection method (Algorithm 4), versus CL, when $L + 1 = 7$, SNR = 20 dB for $N = 128$ and $N = 256$, respectively. Here, the maximum number of iterations for Algorithm 4 is set to $N_q = 2$. As we can see from these figures, the linear and the clipping without compensation systems provide lower and upper bounds for the iterative algorithm. All the curves converge together as CL
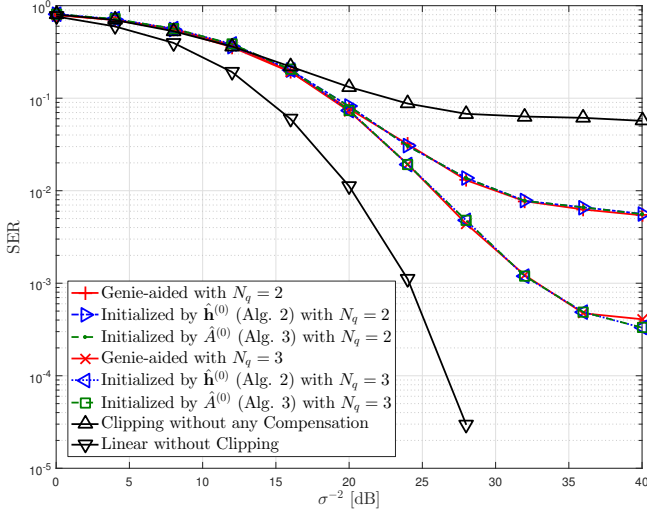
Fig. 6: SER performance of the iterative detection (Algorithm 4) vs. SNR, when $L+1=7$, $N=128$ and CL = 1 dB.

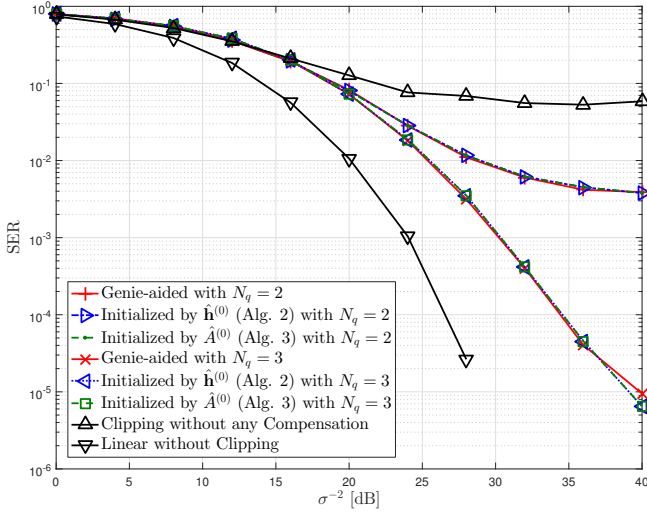Fig. 8: SER performance of the iterative detection (Algorithm 4) vs. SNR, when $L+1=7$, $N=256$ and CL = 3 dB.

Fig. 7: SER performance of the iterative detection (Algorithm 4) vs. SNR, when $L+1=7$, $N=512$ and CL = 1 dB.

Fig. 9: SER performance of the iterative detection (Algorithm 4) with $N_q=3$ vs. SNR, when $L+1=31$, $N=128$ and CL = 1 dB.

increases. Moreover, as illustrated the results of Algorithm 2 and 3, and the result of the genie-aided knowledge are on top of each other.

Fig. 12 shows a contour plot for the SER performance versus CL and SNR using Algorithms 2 and 3, when $L+1=7$, and $N=128$. The performance of Algorithms 2 and 3 are really close together which makes their difference negligible. This figure shows for a given PAPR (or number of sub-carriers) what CL and SNR regions provide a good performance. For instance, to achieve SER of 10^{-2}, we need to look at the corresponding contour curve which shows the locus of different SNRs and CLs. As can be seen from the figure, for CL = 0.6 dB, SER = 10^{-2} can be achieved by SNR = 30 dB, and for CL = 3 dB, this SER is achieved by a lower SNR of 22 dB.

It should be mentioned that, the comparisons here may seem unfair since the amount of training overhead, and thereby the
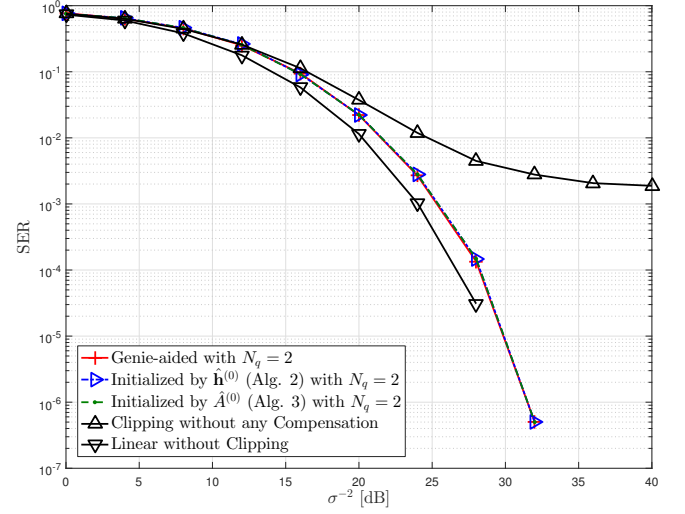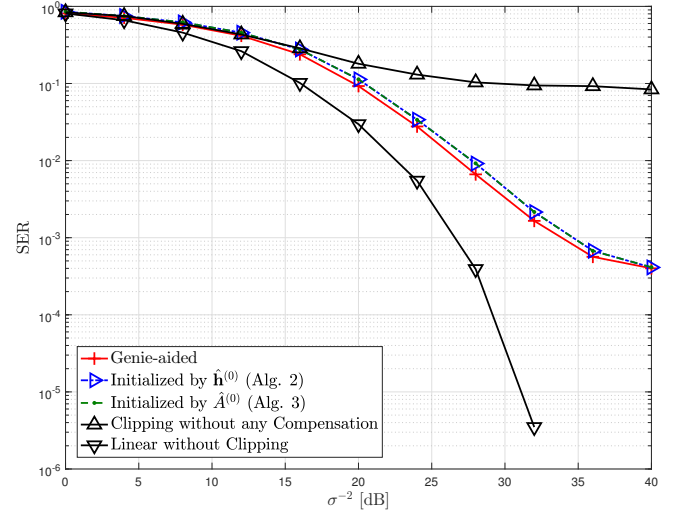
spectral efficiency (or energy per bit) is not the same for the different schemes. However, to make a completely fair comparison, we need to make additional assumptions on how often the CL estimation and the channel estimation is done. In fact, the main point of these plots is to show that the SER performance is virtually unaffected by the estimation errors.

VI. CONCLUSION

In this paper, we have studied joint maximum-likelihood estimation of channel and clipping level at the receiver side in future IoT-based OFDM networks, where there are lots of low-cost low-power nodes transmitting to and receiving from more complex nodes such as a BTS. In particular, we have proposed two alternating optimization algorithms, in which we have optimally solved a non-smooth non-convex optimization problem. We have also computed the theoretical lower bounds
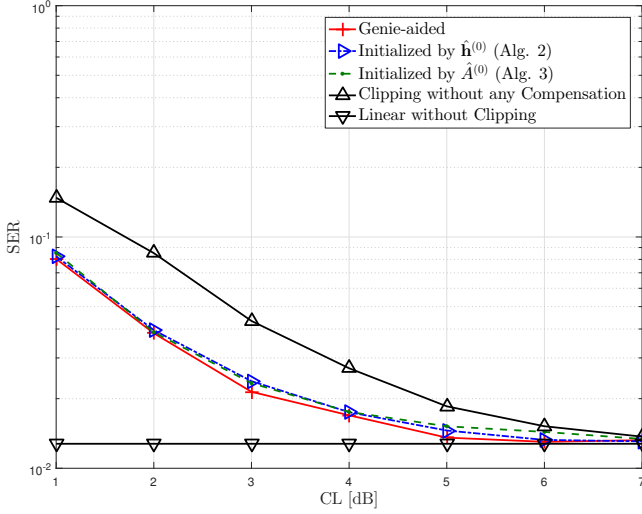
Fig. 10: SER performance of the iterative detection (Algorithm 4) vs. CL, when $L + 1 = 7$, $N = 128$ and SNR $= 20$ dB.

Fig. 12: Contour plot of SER versus CL and SNR using Algorithms 2 and 3, when $N = 128$.
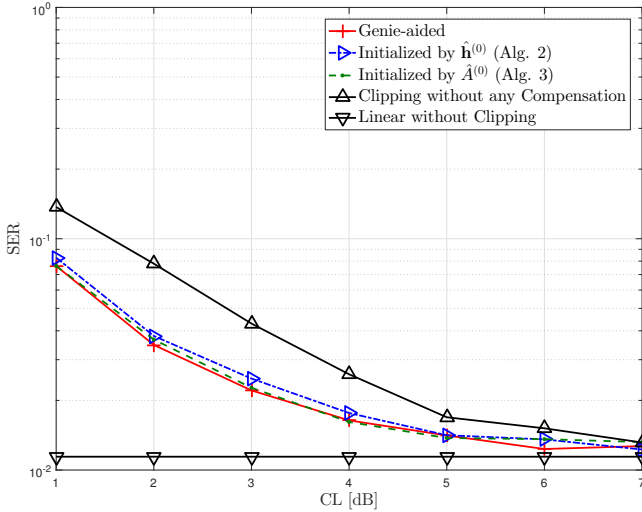

Fig. 11: SER performance of the iterative detection (Algorithm 4) vs. CL, when $L + 1 = 7$, $N = 256$ and SNR $= 20$ dB.

(CRLB) on the performance of these estimators, and showed that they attain these lower bounds. Next, we have combined the channel and the CA estimates with the iterative detection method from [12] to perform symbol detection at the receiver. Finally, we have showed by simulations that the performance of the iterative detection method using the proposed algorithms is almost the same as the one of the case that the receiver has genie-aided knowledge of the channel and CA.

REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[2] S. C. Cripps, *RF Power Amplifiers for Wireless Communications*. Artech House, 1999.

[3] D. Astely, E. Dahlman, G. Fodor, S. Parkvall, and J. Sachs, "LTE release 12 and beyond," *IEEE Commun. Mag.*, vol. 51, no. 7, pp. 154–160, July 2013.
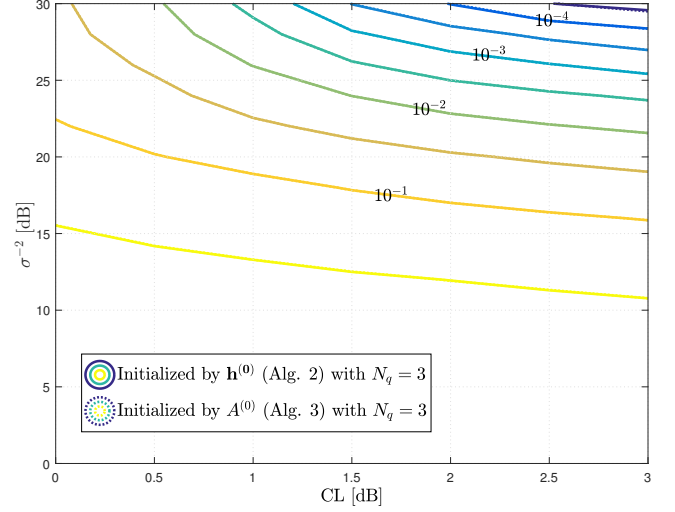
[4] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1567–1592, Mar. 2013.

[5] J. Joung, C. K. Ho, K. Adachi, and S. Sun, "A survey on power-amplifier-centric techniques for spectrum- and energy- efficient wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 315–333, Aug. 2015.

[6] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "Memory polynomial predistorter based on the indirect learning architecture," in *Proc. IEEE Global Telecomm. Conf*, no. 967-971, Nov. 2002.

[7] J. Kim and K. Konstdntinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electronics Letters*, vol. 37, no. 23, pp. 1417–1418, Nov. 2001.

[8] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, , and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159–165, Jan. 2004.

[9] C. Eun and E. J. Powers, "A new volterra predistorter based on the indirect learning architecture," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 223–227, Jan. 1997.

[10] A. Zhu, P. J. Draxler, J. J. Yan, T. J. Brazil, D. F. Kimball, and P. M. Asbeck, "Open-loop digital predistorter for RF power amplifiers using dynamic deviation reduction-based volterra series," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 7, pp. 1524–1534, July 2008.

[11] F. M. Ghannouchi and O. Hammi, "Behavioural modeling and predistortion," *IEEE Microw. Mag.*, vol. 10, no. 7, pp. 52–64, Dec. 2009.

[12] J.Tellado, L. M. C. Hoo, and J. M. Cioffi, "Maximum-likelihood detection of nonlinearly distorted multicarrier symbols by iterative decoding," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 218–228, Feb. 2003.

[13] ——, "Maximum-likelihood detection of nonlinearly distorted multicarrier symbols by iterative decoding," in *Proc. IEEE GlobeCom*, May 1999, pp. 2493–2498.

[14] F. H. Gregorio, S. Werner, J. Cousseau, J. Figueroa, and R. Wichman, "Receiver-side nonlinearities mitigation using an extended iterative decision-based technique," *Signal Processing*, vol. 91, no. 8, pp. 2042–2056, Mar. 2011.

[15] F. Gregorio, T. I. Laakso, and J. E. Cousseau, "Receiver cancellation of nonlinear power amplifier distortion in SDMA-OFDM systems," in *Proc. ICASSP*, vol. 4, 2006, pp. 325–328.

[16] A. S. Tehrani, H. Cao, A. Behravan, T. Eriksson, and C. Fager, "Successive cancellation of power amplifier distortion for multiuser detection," in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, Sept. 2010, pp. 1–5.

[17] J. Gazda, P. Drotar, M. Deumal, D. Kocur, and P. Galajda, "Simple iterative cancellation of nonlinear distortion in OFDMA systems," in *14th International OFDM Workshop*, 2009, pp. 133–137.

[18] E. Olfat and M. Bengtsson, "Estimation of the clipping level in OFDM

systems," in *49th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA*, 2015, pp. 1169–1173.

[19] J. Dohl and G. Fettweis, "Iterative blind estimation of nonlinear channels," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, May 2014, pp. 3923–3927.

[20] ——, "Blind estimation of memoryless AM/PM nonlinearities in OFDM systems," in *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th*, Sept. 2013, pp. 1–5.

[21] ——, "Blind estimation and mitigation of nonlinear channels," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2011 IEEE 22nd International Symposium on*, Sept. 2011, pp. 829–833.

[22] D. Declercq and G. B. Giannakis, "Recovering clipped OFDM symbols with bayesian inference," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, vol. 1, 2000, pp. 157–160 vol.1.

[23] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *Journal of Optimization Theory and Applications*, vol. 109, no. 3, pp. 475–494, 2001.

[24] S. M. Kay, *Fundamentals of statistical signal processing, volume I: estimation theory*. Prentice Hall, 1993.

[25] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1285–1301, Nov. 1990.

[26] V. Nagesha and S. Kay, "Cramer Rao lower bounds for complex parameters," "http://www.ele.uri.edu/faculty/kay/New%20web/downloadable%20files/Nageha_complex%20CRLB.pdf", [Available Online].

[27] S. T. Smith, "Statistical resolution limits and the complexified cramér-rao bound," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 1597–1609, May 2005.

[28] K. H. S. Halford and M. Webster, *Evaluating the performance of HRb proposals in the presence of multipath*, doc: IEEE 802.11-00/282r2, Sept. 2000.

[29] M. Friese, "On the degradation of OFDM-signals due to peak-clipping in optimally predistorted power amplifiers," in *Global Telecommunications Conference, 1998. GLOBECOM 1998. The Bridge to Global Integration. IEEE*, Nov. 1998, pp. 939–944.