# A NOVEL EM-GAMP TECHNIQUE FOR PAPR REDUCTION IN OFDM-BASED HUGE MIMO SYSTEMS

A. Jyothirmai [1], M. Rama Krishna [2] M.Tech.,

[1] *M.Tech Student Department of Electronics and Communication Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada, Andhra Pradesh 520008, INDIA*

[2] *Assistant Professor, Department of Electronics and Communication Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada, Andhra Pradesh 520008, INDIA*

*Jyothiaripi0206@gmail.com.*[1], *Mullaps@gmail.com.*[2]

***Abstract***— In this project the first module, problem of peak-to-average power ratio (PAPR) reduction in orthogonal frequency-division multiplexing (OFDM) based massive multiple-input multiple-output (MIMO) downlink systems. Mainly, a set of symbol vectors to be transmitted to K users, the problem is identified an OFDM-modulated signal that has a low PAPR and meanwhile enables multiuser interference (MUI) cancelation. In previous works that is Partial transmit sequence (PTS) technique requires "V" number of IFFT operations for each data block and bits of side information (SI). The PTS technique suffers from more PAPR problem and the complexity of searching for the optimum set of phase vector. The draw backs of PTS method one will be corrected by a new peak-to-average power ratio (PAPR) reduction approach for MIMO-OFDM is developed based on EM-GAMP algorithm. The sought-after signal is treated as a random vector with a Gaussian noise mixture, which has the potential to encourage a low PAPR signal with most of its samples concentrated on the boundaries. A variational expectation-maximization (EM) strategy is developed to obtain estimates of the hyper parameters associated with the prior model, along with the signal. In addition, the generalized approximate message passing (GAMP) is embedded into the variational EM framework, which results in a significant reduction in computational complexity of the proposed algorithm. Simulation results show our proposed algorithm provides a performance improvement over existing (PTS) methods in terms of both the PAPR reduction and computational complexity.

***Index Terms —*** Massive MIMO, OFDM, PAPR reduction, MUI, SER, PTS, EM-GAMP.

## I. INTRODUCTION

In wireless communication systems, the orthogonal frequency division multiplexing (OFDM) [1] [2] technique is a widely popular and attractive method for high-data-rate transmission because it can cope with frequency-selective fading channel. The modulators and demodulators of OFDM systems can be simply implemented by employing inverse fast Fourier transform (IFFT) and FFT to make the overall system efficient and effective. One of the main challenges of OFDM-based systems is the high peak-to-average power ratio (PAPR) of transmitted signals, resulting in signal distortion. The combination of multi-input multi-output (MIMO) and orthogonal frequency division multiplexing (OFDM) could exploit the spatial dimension capability to improve the system capacity by employing spatially separated antennas. In MIMO-OFDM systems, independent OFDM signals are transmitted from multiple transmit antennas. Therefore, MIMO-OFDM systems still suffer an inherent drawback of high PAPR. This phenomenon results from that in the time domain, an OFDM signal is the superposition of many narrowband subcarriers. At certain time instances, the peak amplitude of the signal is large and at the other times is small, that is, the peak power of the signal is substantially larger than the average power of the signal.

Massive multiple-input multiple-output (MIMO) antenna systems are regarded as one of the key technologies for next- generation (i.e., 5G) wireless communications systems due to their potential to improve data rate and link reliability, as well as to simplify the required signal processing. Massive multiple-input multiple-output (MIMO), also known as large-scale or very-large MIMO, is a promising technology to meet the ever growing demands for higher throughput and better quality-of-service of next-generation wireless communication systems [3]. Massive MIMO systems are those that are equipped with a large number of antennas at the base station (BS) simultaneously serving a much smaller number of single-antenna users sharing the same time-frequency bandwidth. In addition to higher throughput, massive MIMO systems also have the potential to improve the energy efficiency and enable the use of inexpensive, low- power components. Hence, it is expected that massive MIMO will bring radical changes to future wireless communication systems. In practice, broadband wireless communications may suffer from frequency-selective fading.

The multiple carrier frequency method is one of well-known PAPR reduction techniques to deal with frequency selective fading for OFDM systems. It has been described in [4]–[8] However, a major problem associated with the OFDM is a high peak-to-average power ratio (PAPR) owing to the independent phases of the sub-carriers [5]. To avoid out- of-band radiation and signal distortion, handling this high PAPR requires a high-resolution digital-to-analog converter (DAC) and a linear power amplifier (PA) at the transmitter, which is not only expensive but also power inefficient [4].

Many techniques have been developed for PAPR reduction in single-input single-output (SISO) OFDM wireless systems. The most methods are clipping method [11] [12], the clipping scheme is the simplest method to reduce the PAPR. However, the quality of its output signal is degraded by out-of band radiation and in-band distortion, Active constellation extension (ACE) [9] [10], Companding Transform and Filtering [13], selected mapping (SLM) [14], and others. Although these PAPR-reduction schemes can be extended to point-to- point MIMO systems easily, extension to the multi- user (MU) MIMO downlink is not straightforward, mainly because joint receiver-side signal processing is almost impossible in practice as the users are distributed. Recently, a new PAPR reduction method was developed for massive MIMO- OFDM systems i.e., partial transmission sequence (PTS) [1] the disadvantage of this scheme is the complexity, especially with an increase in V and W [16] [17]. Also, a large amount of memory is required to store the alternative transmit signals (if check performed in parallel) in order to compare them to find the one with the lowest peak value [18]. Alternatively the optimization can be performed in an iterative fashion where the current best transit signal is stored until a better one is found, at the cost of increased latency.

In this paper, we develop a novel Bayesian approach to address the joint PAPR reduction and MUI cancelation problem for downlink multi-user massive MIMO-OFDM systems. Specifically, MUI cancelation can be formulated as an underdetermined linear inverse problem which admits numerous solutions. This hierarchical prior has the potential to encourage a quasi-constant magnitude solution with as many entries as possible lying on the truncated boundaries, thus resulting in a low PAPR. A variational expectation-maximization (EM) algorithm is developed to obtain estimates of the hyperparameters associated with the prior model, along with the signal. In addition, the generalized approximate message passing (GAMP) technique [24] is employed to facilitate the algorithm development in the expectation step. This GAMP technique also helps significantly reduce the computational complexity of the pro- posed algorithm. Simulation results show that the proposed method presents a substantial improvement over the FITRA algorithm in terms of both PAPR reduction and computational complexity.

## II. SYSTEM MODEL

We first introduce the system model of OFDM based massive MIMO systems. Then we discuss some recent research on PAPR reduction for multi-user massive MIMO-OFDM systems.

### *A. System Model:*

The Block Diagram of OFDM is shown in below figure.1. A brife description of yhe model is provided below.

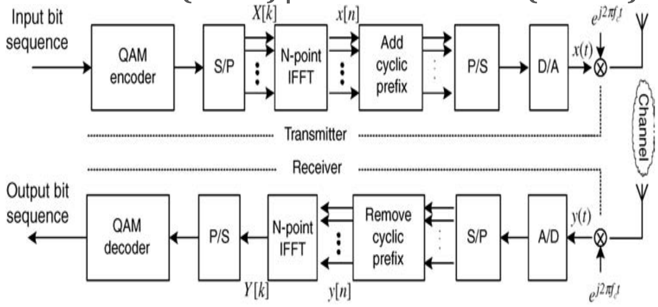



Fig. 1.The block diagram of OFDM based massive MIMO system

The system model of the OFDM-based massive MIMO downlink scenario is shown in Fig. 1, where the BS is assumed to have M transmit antennas and serve K independent single-antenna users (K ≪ M), and the total number of OFDM tones is N. In practice, the set of tones available are divided into two sets $\tau$ and$\tau_c$, where the tones in set $\tau$ are used for data transmission and the tones in its complementary set $T_c$ are used for guard band (unused tones at both ends of the spectrum). Hence, for each tone n ∈ $\tau$, the corresponding K ×1 vector $s_n$ comprises the symbols for K users. We normalize the data vector to satisfy E{$\|s_n\|_2^2$}=1. For each tone n ∈ $\tau_c$, we set $s_n$=$0_{k*1}$ that means no signal is transmitted in the guard band.

Precoding must be performed at the BS to remove multi-user interference (MUI). Usually, the signal vector on the $n_{th}$ tone is linearly precoded as

$$w_{n} = p_n s_n \qquad (1)$$

Where $w_n \in C^{M*1}$ is the precoded vector that contains symbols to be transmitted on the nth sub-carrier through the M antennas respectively and $p_n \in C^{M*K}$ represents the precoding matrix for the $n^{th}$ OFDM tone.

Zero-forcing (ZF) precoding and minimum-mean square-error (MMSE) precoding are two classical precoding schemes. The former aims at removing MUI completely, while the latter tries to achieve balance between the MUI cancellation and the noise enhancement.

In this paper, we consider the ZF precoding scheme. Note that since K ≪ M, the ZF precoding matrix has an infinite number of forms, among which the most widely used is

$$p_n^{zf} = H_n^H \, (H_n \, H_n^H)^{-1} \qquad (2)$$

Where $H_n \in C^{M*K}$ denotes the MIMO channel matrix associated with the nth tone.

After precoding, all precoded vectors $w_n$ are reordered to M antennas for OFDM modulation,

$$[a_1 \ldots\ldots a_M] = [w_1 \ldots \ldots w_N]^T \qquad (3)$$

where $a_m \in C^{N*1}$ represents the frequency domain signal to be transmitted from the $m^{th}$ antenna. The inverse Fast Fourier transform (IFFT), is used to convert frequency domain signal to time domain signal i.e., $\widehat{a_m} = F_N^H \, a_m$ , ∀m. Then, a cyclic prefix (CP) is added to the time-domain samples of each antenna to eliminate inter symbol interference (ISI). And we don't know delay spread exactly; the hardware doesn't allow block spaces because it needs to send out signals continuously. To avoid the block spaces, coping the tail part of the symbol and past at blocked spaces. Finally, these time samples are converted to analog signals and transmitted through the frequency- selective channel.

At the receivers, after removing the CPs of the received signals, the FFT is reconstructing the frequency-domain signals. The receive vector consisting of K users' signals can be described as

$$r_n = H_n\, w_n + e_n,\ \forall n \qquad (4)$$

Where $r_n \in C^{k*1}$ denotes the receive vector associated with the nth tone, and $e_n \in C^{k*1}$ is the receiver noise.

*B. Peak-to-Average Power Ratio (PAPR) Reduction:*

The main challenges of OFDM-based systems are the high peak-to-average power ratio (PAPR) of transmitted signals, resulting in signal distortion. OFDM modulation typically a large dynamic range because the phases of the sub-carriers are independent of each other. To avoid out-of-band radiation and signal distortion by using high-resolution DACs and linear power amplifier at the transmitter to accommodate the large peaks of OFDM signals, which leads to more expensive and power-inefficient.

PAPR is defined as the ratio of the peak power of the signal to its average power. Specifically, the PAPR at the $m^{th}$ transmit antenna is defined as

$$\text{PAPR}=\frac{maX_t|x(t)|^2}{E_t[|x(t)|^2]} \qquad (5)$$

Let E[x(t)] · denote the mathematical expectation and the amplitude, $||x||_2$ denote the second norm of a vector.

When the number of transmit antennas is larger than the number of users, numerous ZF precoding matrices are available. By using this ZF precoding process, the MUI is removed and reduce the PAPR also.

In this paper, instead of designing the precoding matrix we directly search 'w' signal for reducing the PAPR and MUI cancelation.

## III. EXISTING METHOD

Partial transmit sequence (PTS):

The partial transmit sequence (PTS) technique partitions an input data block of N symbols into V disjoint sub blocks as follows:

$$X = [X^0; X^1; X^2; \dots ; X^{V-1}] \qquad (6)$$

Where $X^i$ are the sub blocks that are consecutively located and also are of equal size. Unlike the SLM technique in which scrambling is applied to all subcarriers, scrambling (rotating its phase independently) is applied to each sub block in the PTS technique (see Figure 2). Then each partitioned sub block is multiplied by a corresponding complex phase factor $b^v = e^{j\phi v}$, v=1, 2 …V, subsequently taking its IFFT to yield.

$$x = IFFT(\sum_{v=1}^{V} b^v X^v) = \sum_{v=1}^{V} b^v . IFFT\{X^v\} = \sum_{v=1}^{V} b^v X^v \qquad (7)$$

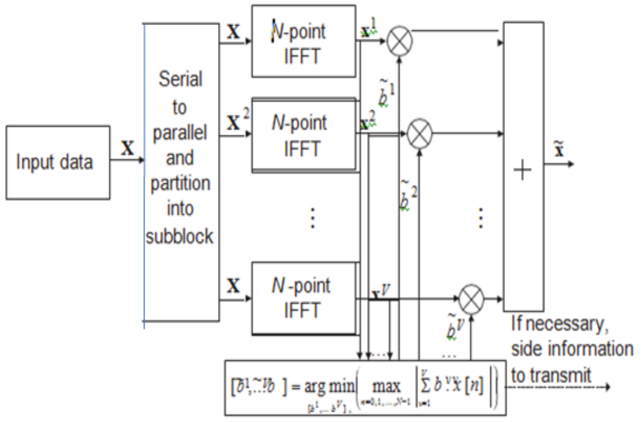



Fig. 2 partial transmit sequence (PTS) technique for PAPR reduction.

Where $\{X^v\}$ is referred to as a partial transmit sequence (PTS). The phase vector is chosen so that the PAPR can be minimized, which is shown as

$$[\tilde{b}^1, \dots \tilde{b}^V] = \arg\ min\, [b^1, \dots, b^V] (\sum_{n=0}^{N-1} |\sum_{v=1}^{V} b^v x^v\, [n]|) \qquad (8)$$

Then, the corresponding time-domain signal with the lowest PAPR vector can be expressed as

$$\tilde{x} = \sum_{v=1}^{V} \tilde{b}^v x^v \qquad (9)$$

In general, the selection of the phase factors $\{\sum_{v=1}^{V} b^v\}$ is limited to a set of elements to reduce the search complexity. As the set of allowed phase factors is

$$b = \left\{e^{\frac{j2\pi}{W}}\right\}, \quad i=0,1,2,\dots\dots W-1 \qquad (10)$$

$W^{v-1}$ Sets of phase factors should be searched to find the optimum set of phase vectors. Therefore, the search complexity increases exponentially with the number of subblocks.

The PTS technique requires V IFFT operations for each data block and $log_2 W^V$ bits of side information. The PAPR performance of the PTS technique is affected by not only the number of subblocks, V, and the number of the allowed phase factors, W, but also the subblock partitioning.

In fact, there are three different kinds of the subblock partitioning schemes: adjacent, interleaved, and pseudo-random. Among these, the pseudo-random one has been known to provide the best performance.

As discussed above, the PTS technique suffers from the complexity of searching for the optimum set of phase vector, especially when the number of sub blocks increases. In the literature various schemes have been proposed to reduce this complexity. One particular example is a suboptimal combination algorithm, which uses the binary phase factors of {1, -1}.

---

Algorithm 1: Partial transmit sequence

---

The Algorithm steps for PTS method is explain below.

1. Partition the input data block into V sub blocks

$$X = [X^0; X^1; X^2; \dots ; X^{V-1}]$$

2. Set all the complex phase factors $b^v = e^{j\phi v}$, multiplied with each sub block.

$b^v = 1$ for v = 1: V, find PAPR of equation (7) and set it as PAPR_min.

$$x = IFFT(\sum_{v=1}^{V} b^v X^v) = \sum_{v=1}^{V} b^v . IFFT\{X^v\} = \sum_{v=1}^{V} b^v X^v$$

3. Find PAPR with $b^{-1} = (-1)$.

$$x = IFFT(\sum_{v=1}^{V} b^v X^v) = \sum_{v=1}^{V} b^v . IFFT\{X^v\} = \sum_{v=1}^{V} b^v X^v$$

4. If PAPR> PAPR_min, switch $b^v$ back to 1. Otherwise, update PAPR_min = PAPR.
5. If v=V, increment v by one and go back to step 3. Otherwise, exit this process with the set of optimal phase factors.

---

The number of computations for Equation (7) in this suboptimal combination algorithm is V, which is much fewer than that required by the original PTS technique (i.e. <<$W^v$).

PTS requires side information to be sent to the receiver to inform it of the phase rotation used so the data can be decoded. Reference [1], [16] noted that the number of angles should be kept low to keep the side information to a minimum.

If each phase rotation is chosen from a set of W admissible angles then the required number of bits for side information. In order to reduce complexity phase angles should be restricted to {±1, ±j}, i.e. W=4, this allows multiplications to be performed with sign changes. That increasing the number of allowed phase angles has a minimum impact of PAPR reduction. Reference [1] noted that explicit side information can be avoided if differential encoding is used for the modulation across the subcarriers within each subblock.

In this case only the block partitioning need to be known at the receiver and one subcarrier in each subblock must be left unmodulated as a reference carrier.

PTS is flexible as the number of blocks and phase rotations can be increased providing more alternative transmit signals to choose from.

The disadvantage of this scheme is the complexity, especially with an increase in V and W. Also, a large amount of memory is required to store the alternative transmit signals (if check performed in parallel) in order to compare them to find the one with the lowest peak value. Alternatively the optimization can be performed in an iterative fashion where the current best transit signal is stored until a better one is found, at the cost of increased latency.

## IV. PROPOSED METHOD

### *A. EM-GAMP Introduction:*

In this project the first module, problem of peak-to-average power ratio (PAPR) reduction in orthogonal frequency-division multiplexing (OFDM) based massive multiple-input multiple-output (MIMO) downlink systems. Mainly, a set of symbol vectors to be transmitted to K users, the problem is identified an OFDM-modulated signal that has a low PAPR and meanwhile enables multiuser interference (MUI) cancelation. The EM-GAMP algorithm is one of the best solutions to overcome this PAPR problem in OFDM signal. The sought-after signal is treated as a random vector with a Gaussian noise mixture, which has the potential to encourage a low PAPR signal with most of its samples concentrated on the boundaries. A variational expectation-maximization (EM) strategy is developed to obtain estimates of the hyper parameters associated with the prior model, along with the signal. In addition, the generalized approximate message passing (GAMP) is embedded into the variational EM framework, which results in a significant reduction in computational complexity of the proposed algorithm.

To facilitate our algorithm development, we introduce a noise term to model the mismatch between y and Ax, i.e.

$$y = Ax + \epsilon \quad (11)$$

Where ϵ denotes the noise vector and its entries are assumed to be i.i.d. Gaussian random variables with zero-mean and unknown variance β−1. Here we treat β as an unknown parameter because the Bayesian framework allows an automatic determination of its model parameters and usually provides a reasonable balance between the data fitting error and the desired characteristics of the solution. In case that there is a pre-specified tolerance value for the MUI, we can also set an appropriate value for β instead of treating it as unknown.

To reduce the PAPR associated with each transmit antenna, we aim to find a quasi-constant magnitude solution to the above underdetermined linear system. Note that a constant magnitude signal achieves a minimum PAPR. Ideally we hope to find a solution with all of its entries having a constant magnitude. Nevertheless, it is highly unlikely that there exists such a solution to satisfy (or approximately satisfy with a tolerable error) the MUI cancelation equality, i.e. (8). Therefore we, alternatively, seek a quasi-constant magnitude solution with as many entries as possible located on the boundary points of an interval [−v, v], whereas the rest entries bounded within [−v, v] but not restricted to lie on the boundary points in order to meet the MUI cancelation constraint.

To encourage a quasi-constant magnitude solution, we propose a hierarchical truncated Gaussian mixture prior for the signal x. In the first layer, coefficients of x are assumed independent of each other and each entry $x_i$ is assigned a truncated Gaussian mixture distribution:

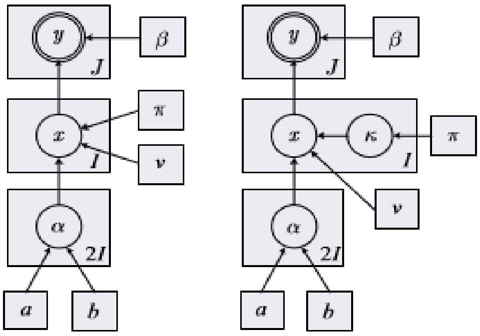
$$P(x_i) = \int_0^{\pi \frac{N(x_i ; v, \alpha_{i1}^{-1})}{\eta_{i1}} + (1-\pi)\, \pi \frac{N(x_i ; -v, \alpha_{i2}^{-1})}{\eta_{i2}}} \quad \begin{matrix} if\ x_{i \in [-v,v]}, \\ otherwise \end{matrix} . \quad (12)$$

Where the first component of (12) is characterized by a truncated Gaussian distribution with its mean and variance given by v and $\alpha_{i1}^{-1}$ , respectively; the second component is characterized by a truncated Gaussian distribution with its mean and variance given by −v and $\alpha_{i2}^{-1}$ , respectively;

The prior distributions with different model hyper parameters$\alpha_{i1}$, $\alpha_{i2}$ are illustrated in Fig.7.1, where π and v are both set to 0.5. We can see that the prior distribution defined in (22) resembles the shape of a bowl. Thus the prior has the potential to push the entries of the solution toward its boundaries. In addition, the use of the Gamma hyper prior allows the posterior mean of the precision to become arbitrarily large.

As a result, the associated entries $x_i$ will eventually lie on one of the two boundary points, leading to a quasi-constant magnitude solution.

The graphical model of the proposed hierarchical is presented in Fig. 3(a). In general, Bayesian inference requires computing the logarithm of the prior. In this regard, (24) is a inconvenient form for inference.




(a) (b)

Fig.3. Graphical models for low-PAPR signal priors. (a) Original prior, (b) Modified prior

Here circles denoting hidden variables, double circles denoting observed variables and squares representing model parameters.

To address this issue, we turn the prior into an exponential form by introducing a binary latent variable $k_i$ indicating which component is selected for $x_i$, i.e., $k_i = 1$ indicates the first component is selected while $k_i = 0$ corresponds to the second component.

*B. EM-GAMP Algorithm:*

To search for a low PAPR solution, a hierarchical truncated Gaussian mixture prior model is proposed and assigned to the unknown signal (i.e. solution). This hierarchical prior has the potential to encourage a quasi-constant magnitude solution with as many entries as possible lying on the truncated boundaries, thus resulting in a low PAPR. A variational expectation-maximization (EM) algorithm is developed to obtain estimates of the hyper parameters associated with the prior model, along with the signal. In addition, the generalized approximate message passing (GAMP) technique [22] is employed to facilitate the algorithm development in the expectation step. This GAMP technique also helps significantly reduce the computational complexity of the proposed algorithm. Simulation results show that the proposed method presents a substantial improvement over the PTS algorithm in terms of both PAPR reduction and computational complexity.

i) Variational Bayesian Methodology:

We now proceed to perform Bayesian inference for the proposed hierarchical model. A variational expectation maximization (EM) strategy is employed for the Bayesian inference. In our model, z ≜ {x, α1, α2, κ} are treated as hidden variables. The noise variance β and the boundary parameter v are unknown deterministic parameters, i.e. θ ≜ {β, v}. Before proceeding, we provide a brief review of the variational EM algorithm.

Consider a probabilistic model with observed data y, hidden variables z and unknown deterministic parameters θ. It is straightforward to show that the marginal probability of the observed data can be decomposed into two terms

$$\ln(p(y;\theta)) = F(q,\theta) + KL(q\|p) \quad (13)$$

Where

$$F(q,\theta) = \int q(z) \ln\left(\frac{p(y,z;\theta)}{q(z)}\right) dz \quad (14)$$

and

$$KL(q\|p) = -\int q(z) \ln\left(\frac{p(z|y;\theta)}{q(z)}\right) dz \quad (15)$$

Where q(z) is any probability density function, KL(q‖p) is the Kullback-Leibler divergence between p(z|y;θ) and q(z). Since KL (q‖p) ≥ 0, it follows that F (q, θ) is a lower bound of ln (p(y; θ)), with the equality holds only when KL (q‖p) = 0, which implies p(z|y;θ) = q(z). The EM algorithm can be viewed as an iterative algorithm which iteratively maximizes the lower bound F (q, θ) with respect to the distribution q (z) and the parameters θ.

Assume that the current estimate of the parameters is $\theta^{OLD}$. The EM algorithm evaluates $q^{NEW}$(z) by maximizing F(q,$\theta^{OLD}$) with respect to q(z) in the E-step, and then find new parameter estimate $\theta^{NEW}$ by maximizing F($q^{NEW}$, θ) with respect to θ in the M-step. It is easy to see that when $q^{NEW}$(z)=p(z|y;$\theta^{OLD}$), the lower bound F(q, $\theta^{OLD}$) is maximized. Nevertheless, in practice, the posterior distribution p(z|y; $\theta^{OLD}$) is usually computationally intractable. To address this difficulty, we could assume q (z) has some specific parameterized functional form.

Then in the M-step, a new estimate of θ is obtained by maximizing the Q-function

$$Q(\theta, \theta^{OLD}) = \langle \ln(p(y,z;\theta)) \rangle_{q(Z)} \quad (16)$$

ii) Likelihood Function Approximation via GAMP:

Let z = {x, α1, α2, κ} denote all hidden variables appearing in our hierarchical model, and θ ≜{β, v} denote the unknown deterministic parameters. As discussed in the previous subsection, the posterior of z can be approximated by a factorized form as follows

$$p(x,\alpha_1,\alpha_2,\kappa|y;\beta,v) \approx q(x,\alpha_1,\alpha_2,\kappa) = q(x)q(\alpha_1)q(\alpha_2)q(\kappa) \quad (17)$$

The approximate posteriors can be obtained as

$$\ln q(x) = \langle \ln p(y,x,,\alpha_1,\alpha_2,\kappa;\beta,v) \rangle_{q(\alpha1)q(\alpha_2)q(\kappa)} + const, \ldots \ldots\ldots(18.A)$$

$$\ln q(\alpha_1) = \langle \ln p(y,x,\alpha_1,,\alpha_2,\kappa;\beta,v) \rangle_{q(x)q(\alpha_2)q(\kappa)} + const, \ldots\ldots\ldots.(18.B)$$

$$\ln(\alpha_2) = \langle \ln p(y,x,\alpha_1,,\alpha_2,\kappa;\beta,v) \rangle_{q(x)q(\alpha_1)q(\kappa)} + const \ldots\ldots..(18.C)$$

$$\ln q(\kappa) = \langle \ln p(y,x,\alpha_1,,\alpha_2,\kappa;\beta,v) \rangle_{q(x)q(\alpha_1)q(,\alpha_2)} + const \ldots\ldots(18.D)$$

iii) Generalized approximate message passing (GAMP):

A variational expectation-maximization (EM) algorithm is developed to obtain estimates of the hyper parameters associated with the prior model, along with the signal. In addition, the generalized approximate message passing (GAMP) technique [22] is employed to facilitate the algorithm development in the expectation step. This GAMP technique also helps significantly reduce the computational complexity of the pro- posed algorithm.

GAMP is a very-low-complexity Bayesian iterative technique recently developed in [23] for obtaining approximate marginal posteriors and likelihoods. It therefore can be naturally embedded within the EM framework to provide an approximate posterior distribution of x and reduce the computational complexity, as shown in [20]. Specifically, the EM-GAMP framework of [25] proceeds in a double-loop manner: the outer loop (EM) computes the Q-function using the approximate posterior distribution of x, and maximizes the Q-function to update the model parameters (e.g. α1, α2, κ); the inner loop (GAMP) utilizes the newly estimated parameters to obtain a new approximation of the posterior distribution of x.

However, this procedure is not suitable for our variational EM framework, because from the GAMP's point of view, the hyper parameters {α1,α2,κ} need to be known and fixed in order to compute an approximate posterior distribution of x, while the variational EM treats the model parameters (e.g. α1,α2,κ) as latent variables.

Therefore, instead of computing the approximate posterior distribution of x, in our variational EM framework, the GAMP is simply used to obtain an amiable approximation of the likelihood function p(y|x;β), and this approximation involves no latent variables {α1,α2,κ}.

Besides, unlike the EM-GAMP framework where the inner loop (GAMP) is implemented in an iterative way, in our proposed variational EM-GAMP framework, as detailed in Algorithm 1, the GAMP only needs to go through one iteration to obtain an approximation of the likelihood function. In fact, the GAMP algorithm described here is a simplified version of the original GAMP algorithm by retaining only its first three steps and skipping its iterative procedure. Note that the original GAMP algorithm involves a four-step iterative process, in which the fourth step computes the posterior of x by using the approximate likelihood function obtained from the first three steps.

GAMP is known to work well for A with i.i.d zero-mean sub-Gaussian entries, but may fail for a rank-deficient A. One may refer to the method [21] to improve the stability of the GAMP against the ill-condition of the matrix A. Nevertheless, GAMP is expected to perform well in wireless communication scenarios since indoor and urban outdoor environments are typically rich in scattering and entries of MIMO channel matrices are usually assumed to be i.i.d Gaussian [23].

*C. EM-GAMP Procedure:*

To search for a low PAPR solution, a hierarchical truncated Gaussian mixture prior model is proposed and assigned to the unknown signal (i.e. solution). This hierarchical prior has the potential to encourage a quasi-constant magnitude solution with as many entries as possible lying on the truncated boundaries, thus resulting in a low PAPR. A variational expectation-maximization (EM) algorithm is developed to obtain estimates of the hyper parameters associated with the prior model, along with the signal.

In addition, the generalized approximate message passing (GAMP) technique [22] is employed to facilitate the algorithm development in the expectation step. This GAMP technique also helps significantly reduce the computational complexity of the proposed algorithm.

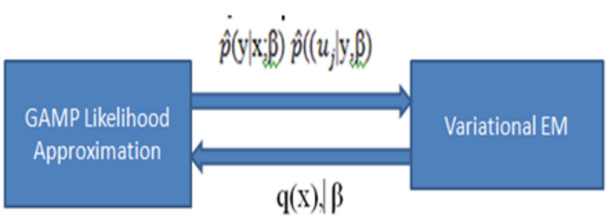



Fig.4 Variational EM-GAMP framework.

GAMP is a simplification of loopy BP, and can be used to compute approximate marginal posteriors and likelihoods. Here we approximate the joint likelihood function p(y|x;β) as a product of approximate marginal likelihoods computed via the GAMP, i.e.

$$p(y|x;\beta) \approx \hat{p}(y|x;\beta) \propto \prod_{i=1}^{I} N(x_i|\widehat{r_\imath, \tau_i^r}) \qquad (19)$$

Where $N(x_i|\widehat{r_\imath, \tau_i^r})$ is the approximate marginal likelihood obtained by the GAMP algorithm. To calculate $r_i$ and $\tau_i^r$, an estimate of the posterior q(x) and β is required as inputs to the GAMP algorithm (see the details of the GAMP algorithm provided below). Hence the GAMP algorithm can be embedded in the variational EM framework: given an estimate of q(x) and β, use the GAMP to obtain an approximation of the likelihood function p(y|x;β); with the approximation ˆ p(y|x;β), the variational EM proceeds to yield a new estimate of q(x) and β, along with estimates of other deterministic parameters (e.g. v) and posterior distributions for the other hidden variables (e.g. α1,α2,κ). This iterative procedure is illustrated in Figure 4.

Note that besides the approximation $\hat{p}$ (y|$x_i$;β), GAMP also produces approximations for the marginal posteriors of the noiseless output u = $[u_1, u_2\ldots, \ldots u_j]^T \triangleq$ Ax, which are given by

$$p(u_j|y,\beta) \approx \hat{p}((u_j|y,\beta) \propto p((y_j|(u_j;\beta)N((u_j|\widehat{p_\jmath},\tau_j^p) \qquad (20)$$

Where $\widehat{p_\jmath}$ and $\tau_j^p$ are quantities obtained from the GAMP algorithm.

GAMP is a very-low-complexity Bayesian iterative technique recently developed in [24] for obtaining approximate marginal posteriors and likelihoods. It therefore can be naturally embedded within the EM framework to provide an approximate posterior distribution of x and reduce the computational complexity, as shown in [24]. Specifically, the EM-GAMP framework of [24] proceeds in a double-loop manner: the outer loop (EM) computes the Q-function using the approximate posterior distribution of x, and maximizes the Q-function to update the model parameters (e.g. α1, α2, κ); the inner loop (GAMP) utilizes the newly estimated parameters to obtain a new approximation of the posterior distribution of x.

However, this procedure is not suitable for our variational EM framework, because from the GAMP's point of view, the hyper parameters {$\alpha_1$, $\alpha_{2,}$κ} need to be known and fixed in order to compute an approximate posterior distribution of x, while the variational EM treats the model parameters (e.g. α1,α2,κ) as latent variables. Therefore, instead of computing the approximate posterior distribution of x, in our variational EM framework, the GAMP is simply used to obtain an amiable approximation of the likelihood function p(y|x;β), and this approximation involves no latent variables {α1,α2,κ}.

Besides, unlike the EM-GAMP framework where the inner loop (GAMP) is implemented in an iterative way, in our proposed variational EM-GAMP framework, the GAMP only needs to go one iteration to obtain an approximation of the likelihood function. In fact, the GAMP algorithm described here is a simplified version of the original GAMP algorithm by retaining only its first three steps and skipping its iterative procedure. Note that the original GAMP algorithm involves a four-step iterative process, in which the fourth step computes the posterior of x by using the approximate likelihood function obtained from the first three steps.

. Note that we can also treat {α1, α2, κ} as deterministic parameters and resort to the EM-GAMP framework for Bayesian inference. Nevertheless, in this case, we need to estimate a set of binary parameters $\{k_i\}$ in the M-step. This is essentially a combinatorial search problem and the binary estimation may cause the algorithm to get stuck in undesirable local minima.

i) E-Step: Update of Hidden Variables:

Update of q(x): As discussed above, p(y|x;β) is approximated as a factorized form of independent scalar likelihoods, which enables the computation of q(x) (18.A). Specifically, using (18.A) (19) can be simplified as

$\ln q(x)=\langle \ln p\,(y, x, , \alpha_1, \alpha_2, \kappa; \beta, v)\rangle_{q(\alpha1)q(\alpha_2)q(\kappa)}$ +const (20)

lnq(x) = −∞ otherwise. It can be seen that lnq(x) has a factorized form, which implies that hidden variables {xi} have independent posterior distributions. Also, it can be readily verified that the posterior q(xi) follows a truncated Gaussian distribution

$$q\,(x_i) = \int \begin{matrix} \frac{N\left(x_i|\mu_{i,}\sigma_i^2\right)}{\phi_i} & \text{if } x_i \in[-v,v], \\ 0 & otherwise \end{matrix} \quad \text{......................} (21)$$

Where $\sigma_i^2$= variance

$\mu_{i,}$ = Mean

$\emptyset_i$= Normalization constant

Update of q ( $\alpha_1$ ) :

Keeping only the terms that depend on $\alpha_1$, the variational optimization of q ($\alpha_1$) yields

$\ln q(\alpha_1)= \langle \ln p\,( x|\alpha_1, \alpha_2, \kappa; v)p(\alpha_1)\rangle_{q(x)q(\alpha_2)q(\kappa)}$ +const

$= \sum_{i=1}^{I}.\ \langle \ln p\,( x_i|\alpha_{i1}, \alpha_{i2}, k_i; v)p(\alpha_{i1})\rangle_{q(x)q(\alpha_2)q(\kappa)}$ +const

Therefore q($\alpha_1$) follows a Gamma distribution

$$q\,(\alpha_1) = \text{Gamma}\,(\alpha_{i1}|\ \widehat{a_{\iota 1}}, \widehat{b_{\iota 1}}) \quad (22)$$

with

$$\widehat{a_{\iota 1}} = a + \frac{1}{2}\langle k_i \rangle$$

$$\widehat{b_{\iota 1}} = b + \frac{1}{2}\langle k_i \rangle\ \langle (x_i - v)^2 \rangle$$

Update of q ($\alpha_2$):

Following a procedure similar to the derivation of q ($\alpha_1$), we have

$$q\,(\alpha_2) = \text{Gamma}\,(\alpha_{i2}|\ \widehat{a_{\iota 2}}, \widehat{b_{\iota 2}}) \quad (23)$$

with

$$\widehat{a_{\iota 2}} = a + \frac{1}{2}(1 - \langle k_i \rangle$$

$$\widehat{b_{\iota 2}} = b + \frac{1}{2}(1 - \langle k_i \rangle)\ \langle (x_i + v)^2 \rangle$$

ii. M-step: Update of deterministic Parameters:

As indicated earlier, in the variational EM framework, the deterministic parameters $\theta = \{\beta, v\}$ are estimated by maximizing the Q-function, i.e.

$$\theta^{NEW} = {}^{max}_{\theta} Q\ (\theta, \theta^{old}) \quad (24)$$

Update of β:

We first discuss the update of the parameter β the inverse of the noise variance. Since the GAMP algorithm provides an approximate posterior distribution for the noise less output u ≜ Ax., we can simply treat u as hidden variables when computing the Q-function, i.e.

$$Q(\beta, \beta^{(t)}) = \sum_{j=1}^{J} \langle \ln p\big(y_j|u_j; \beta\big)\rangle_{p(\widehat{u_j}|y,\beta)} + \text{const} \quad (25)$$

The new estimate of data is obtained by maximizing the Q-function, which can be solved by setting the derivative of Q($\beta, \beta^{(t)}$) with respect to $\beta$ to zero.

Update of v:

We now discuss how to update the boundary parameter v. The boundary parameter v can be updated by maximizing the Q-function with respect to v. Nevertheless, the optimization is complex since the Q-function involves computing the expectation of the normalization terms $\eta_{il}$, i = 1, 2, …..I, l = 1, 2 with respect to the posterior distributions p ($\alpha_{il}$). The basic idea is to find an appropriate value of v such that the mismatch ‖y−A$\hat{x}$‖2 is minimized, where ˆ x denotes the estimated signal which is chosen as the mean of the posterior distribution q(x).

Note that when the boundary parameter v is small, the mismatch could be large since there may not exist a solution to satisfy the constraint y = Ax. Therefore we can firstly set a small value of v, then gradually increase v by a step-size such that the mismatch keeps decreasing and eventually becomes negligible.

**Algorithm 2: EM-GAMP:**

EM-TGM-GAMP Initialization: β(0) = $10^3$, $v^{(0)}$ = ‖y‖∞/‖A‖∞, initialize the means of q(x), q(α1), q(α2), q(κ) as 0, 1, 1, $\frac{1}{2}$1 respectively, set the variance of q(x) as 1, and set iteration number t = 0.

Repeat the following steps until t ≥ tMAX

1. Based on the mean and variance of q(x) and β(t), calculate the approximate distributions ˆ p(y|x;β(t))and ˆ p(uj|y,β(t)), j = 1,...,J,
2. Using the approximate likelihood ˆ p(y|x;β(t)), update the posteriors of hidden variables: q(x), q(α1), q(α2) and q(κ)
3. Compute the new estimate β(t+1) and obtain the v(t+1)
4. Increase t = t + 1 and return to step 1.

## V. SIMULATION RESULT

In this section, some simulations are employed to demonstrate PAPR reduction performance and computational complexity comparison between the proposed scheme and the original PTS scheme. The OFDM symbol of each antenna channel contains 1024 subcarriers, and for simplicity, we expect all N subcarriers. . We compare our approach method( EM-GAMP) with the PTS [1] [16], the zero-forcing (ZF) precoding scheme, and the amplitude clipping scheme [11] in which the ZF is first employed and then the peaks of the resulting signal are clipped with a specified threshold.

In our simulations, we consider a MIMO system which has M = 16 number of bits per QAM symbol and alphabet size at the BS and serves K = 10 single antenna users. A 16-QAM constellation is considered, and the number of OFDM tones is set to N = 1024, in which only |T| = 840 tones are used for data transmission and remaining 93 tones are guard symbols.

The complementary cumulative distribution function (CCD- F) is used to evaluate the PAPR reduction performance. The CCDF denotes the probability that the

PAPR of the estimated signal exceeds a given threshold PAPR0,

i.e. CCDF(PAPR0) = Pr(PAPR > PAPR0). (31)

Also, to evaluate the multiuser interference of the transmit signals. The out-of-band (power) ratio (OBR) is introduced to measure the out-of-band radiation of the solution, which is defined as

$$\text{OBR} = \frac{|\mathcal{T}|\ \sum_{n\in\mathcal{T}^c}\|w_n\|_2^2}{|\mathcal{T}^c|\ \sum_{n\in\mathcal{T}}\|w_n\|_2^2} \quad (32)$$

.The signals estimated by respective schemes. In the (a), (b), of Fig. 5, we depict the real part of the first transmit antenna's time-domain signal (i.e.$\hat{a}_1$) estimated by respective schemes (the imaginary part behaves similarly). Time domain signals representation for PTS and EM-GAMP methods. The PAPR for EM-GAMP is 1.2015 dB, and the MUI (Multi User Interference) is -79.98 dB. The PAPR for PTS is 9.393 (Approximately).

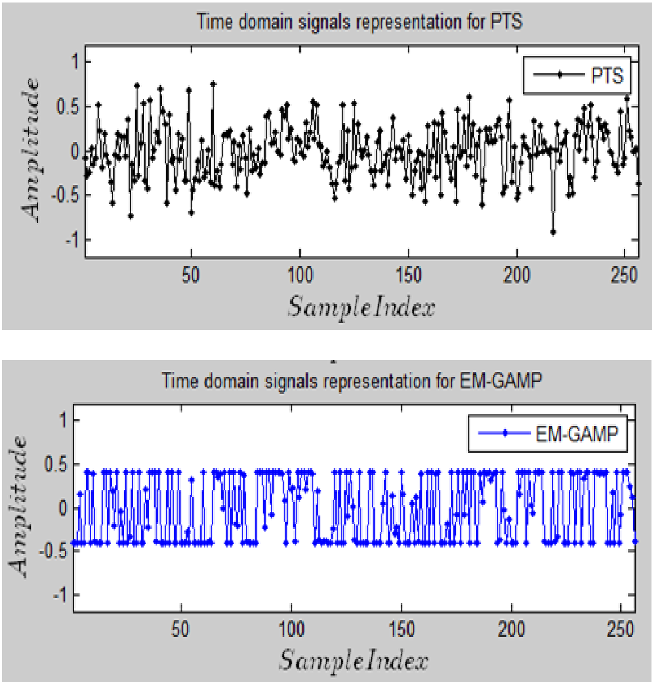



Fig. 5 Time domain signals representation for different PTS and EM-GAMP

Fig 6(a) is an OFDM signal using PTS method with subblocks 1, 2, 4. Here the signal is sub divided for reducing the PAPR loss like subblock 1, subblock 2, subblock 4. If increases the subblocks than decreases the PAPR loss. In this PTS method the PAPR loss is depends on subblocks i.e. PAPR loss is inversely proportional to subblocks

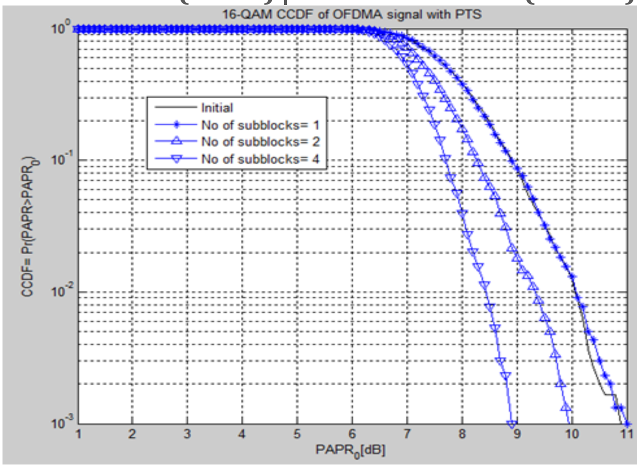



Fig:6(a) PAPR reduction using PTS method in OFDM signal with subblocks 1,2,4

Fig 6(b) is an OFDM signal using PTS method with subblocks 1, 2, 4. Here the signal is sub divided for reducing the PAPR loss like subblock 8, subblock 16. If increases the subblocks than decreases the PAPR loss. In this PTS method the PAPR loss is depends on subblocks i.e. PAPR loss is inversely proportional to subblocks.

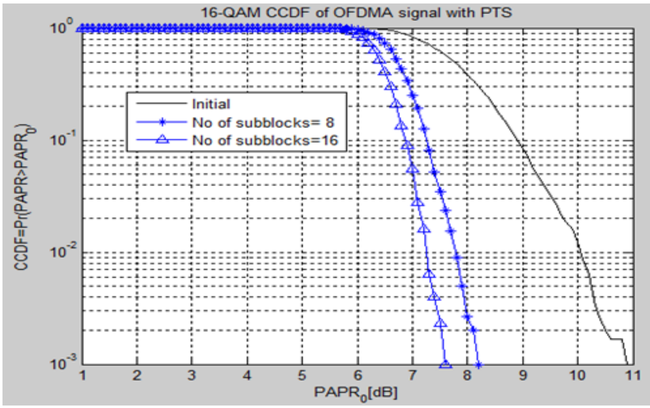



Fig: 6(b) PAPR reduction using PTS method in OFDM signal with subblocks 8, 16.

To better evaluate the PAPR reduction performance, we plot the CCDF of the PAPR for EM-GAMP in Fig.7 (a). The number of trials is chosen to be 1024 in our experiments. Note that PAPRs associated with all M antennas are taken in account in calculating the empirical CCDF. We also include the results of our proposed algorithm obtained at the 200$^{th}$ iteration achieves a substantial PAPR reduction: it reduces the PAPR by more than 4dB compared to the ZF scheme (at CCDF(PAPR) = 1%), by about 2dB compared to the PTS algorithm with 2000 iterations.

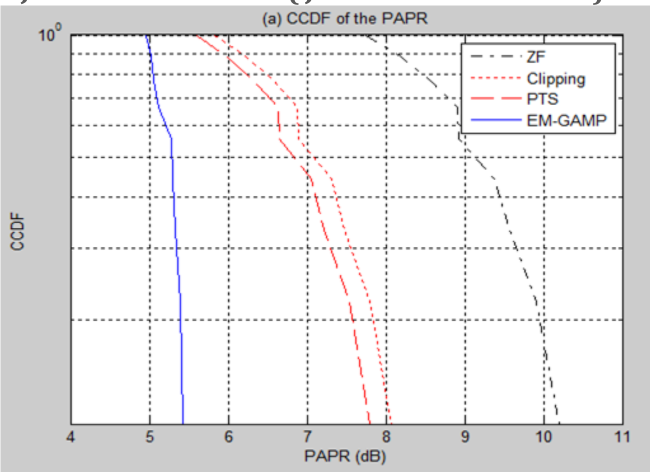



Fig: 7 (a) PAPR reduction using EM-GAMP

The SER performance of respective schemes is shown in Fig. 7(b), where the signal-to-noise ratio (SNR) is defined as $\text{SNR} = \| x \|_2^2 \; MN_0$, $N_0$ denotes the variance of the receiver noise.

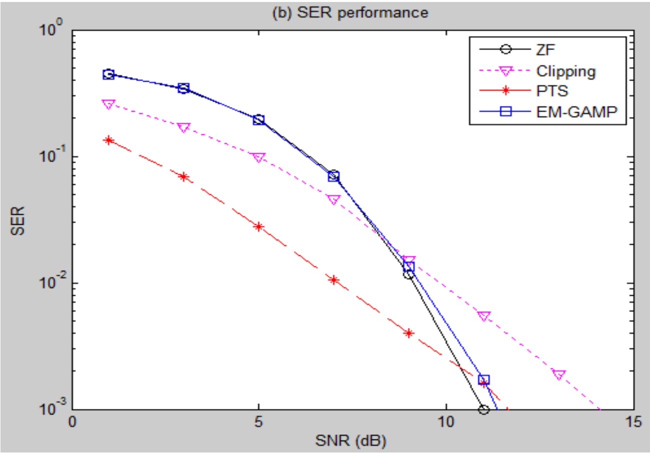



Fig: 7(b) SER performance in different schemes.

TABLE 1

| Method | PTS | EM-GAMP |
|---|---|---|
| FFT Size | 1024 | 1024 |
| Iterations | 3000 | 200 |
| PAPR | 8dB | 5dB |
| Complexity | High | Low |

## VI. CONCLUSION

We considered the problem of joint PAPR reduction and multiuser interference (MUI) cancelation in OFDM based massive MIMO downlink systems. A hierarchical truncated Gaussian mixture prior model was proposed to encourage a low PAPR solution/signal. A variational EM algorithm was developed to obtain estimates of the hyperparameters associated with the prior model, as well as the signal. Specifically, the GAMP technique was embedded into the variational EM framework to facilitate the algorithm development. The proposed algorithm only involves simple matrix- vector multiplications at each iteration, and thus has a low computational complexity. Simulation results show that the proposed algorithm achieves notable improvement in PAPR reduction as compared with the PTS algorithm [14], and meanwhile renders better MUI cancelation and lower out-of- band radiation. The proposed algorithm also demonstrates a fast convergence rate, which makes it attractive for practical real-time systems.

## REFERENCES


[1] Young-Jeon Cho, Kee-Hoon Kim, Jun-Young Woo, Kang-Seok Lee, Jong-Seon No, Fellow, IEEE, and Dong-Joon Shin, Senior Member, IEEE, “Low-Complexity PTS Schemes Using Dominant Time-Domain Samples in OFDM Systems,” IEEE Trans. Broad casting, , pp 1-6, 2017.

[2] R. van Nee and R. Prasad, OFDM for Wireless Multimedia Communications, Boston: Artech House, 2000.

[3] H. Yang, “A road to future broadband wireless access: MIMO-OFDM based air interface,” Communications Magazine, IEEE, vol. 43, no. 1, pp. 53– 60, 2005.

[4] T. Jiangand Y. Wu, “An overview: Peak-to-average power ratio reduction techniques for OFDM signals,” IEEE Trans. Broad casting, vol. 54, no. 2, pp. 257–268, June 2008.

[5] G. Wunder, R. F. Fischer, H. Boche, S. Litsyn, and J. No, “The PAPR problem in OFDM transmission: New directions for a long-lasting problem,” IEEE Signal Process. Mag., vol. 30, no. 6, pp. 130–144, Jan. 2014.

[6] S. H. Han and J. H. Lee, “An overview of peak-to-average power ratio reduction techniques for multicarrier transmission,” IEEE Wireless Commun., vol. 12, no. 2, pp. 56–65, Apr. 2005.

[7] J. Tellado, “Peak to average power reduction for multicarrier modulation,” PhD thesis (Stanford University, 2000).

[8] H. Prabhu, O. Edfors, J. Rodrigues, L. Liu, and F. Rusek, “A low- complex peak-to-average power reduction scheme for OFDM based massive MIMO systems,” in Communications, Control and Signal Processing (ISCCSP), 2014 6th International Symposium on, Athens, Greece, 2014.

[9] T. Tsiligkaridis and D. L. Jones, “PAPR reduction performance by active constellation extension for diversity MIMO-OFDM systems, ” J. Electrical and Computer Eng. , Sept. 2010.

[10] B. S. Krongold and D. L. Jones, “PAR reduction in OFDM via active constellation extension, ” IEEE Trans. Broadcasting, vol. 49, no. 3, pp. 258–268, Sept. 2003.

[11] G. Hill, "Comparison of Low complexity clipping algorithms for OFDM," in *PIMRC*, Portugal, 2002.

[12] X. Li and L. J. Cimini, "Effects of clipping and filtering in the performance of OFDM," *IEEE Communications Letters,* vol. 2, pp. 131-133, May 1998.

[13] Yong Wang, Chao Yang, and Bo Ai, Senior Member, IEEE, “Iterative Companding Transform and Filtering for Reducing PAPR of OFDM Signal”, IEEE Transactions on Consumer Electronics, Vol. 61, No. 2, PP. 144-150, May 2015

[14] R. W. B¨auml, R. F. Fischer, and J. B. Huber, “Reducing the peak-to- average power ratio of multicarrier modulation by selected mapping,” IEE Elec. Letters, vol. 32, no. 22, pp. 2056–2057, Oct. 1996.

[15] R. F. H. Fischer and M. Hoch, “Directed selected mapping for peak- to-average power ratio reduction in

MIMO OFDM," IEE Elec. Letters, vol. 42, no. 2, pp. 1289–1290, Oct. 2006.
[16] S. H. M¨uller and J. B. Huber, "OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences," IEE Elec. Letters, vol. 33, no. 5, pp. 368–369, Feb. 1997.
[17] Hyun-Seung Joo, Kee-Hoon Kim, Jong-Seon No, and Dong-Joon Shin, "New PTS Schemes for PAPR Reduction of OFDM Signals Without Side Information," IEEE Trans. Broad casting, , pp.1-9, 2017.
[18] S. G. Kang, J. G. Kim, and E. K. Joo, "A novel subblock partition scheme for partial transmit sequence OFDM," *IEEE Transactions on Broadcasting,* vol. 45, pp. 333-338, 1999.
[19] S. H. Müller and J. B. Huber, "A comparison of Peak Power Reduction Schemes for OFDM," in *Globecom '97*, Phoenix, USA, 1997
[20] C. Tellambura, "Computation of the continuous time PAR of an OFDM signal with BPSK subcarriers, " IEEE Commun. Let. vol. 5, no. 5, pp. 185–187, May 2001.
[21] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for Bayesian inference," IEEE Signal Process. Mag., vol. 25, no. 6, pp. 131–146, Jan. 2008.
[22] J. Vila and P. Schniter, "Expectation-maximization Gaussian-mixture approximate message passing," IEEE Trans. Signal Process., vol. 61, no. 19, pp. 4658–4672, Oct 2013.
[23] Q. Guo, D. Huang, S. Nordholm, J. Xi, , and Y. Yu, "Iterative frequency domain equalization with generalized approximate message passing," IEEE Signal Process. Lett., vol. 20, no. 6, pp. 559–562, June 2013.

[24] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," IEEE International Symposium, Information Theory Proceedings (ISIT), Full version at arXiv:1010.5141, pp. 2168- 2172, Aug. 2011.
[25] J. Chen, C. Wang, K. Wong, and C. Wen, "Low-complexity precoding design for massive multiuser MIMO systems using approximate message passing," IEEE Trans. Vehicular Technology, vol. PP, no. 99, pp. 1–8, Jul. 2015.

AUTHORS

A. Jyothirmai received her B.Tech degree in Electronics & Communication Engineering. in 2016 from Jawaharlal Nehru Technological University (JNTUK), Kakinada, Andhra Pradesh, India and Now she studying M.Tech in Digital Electronics and Communication system in Andhra Loyola Institute of Engineering & Technology (ALIET), JNTUK University, Kakinada, Andhra Pradesh, India. Her research interest includes OFDM system and various PAPR Reduction schemes.

Dr. M. Rama Krishna received B.E. degree in Electronics & Communication Engineering in 1997 from Andhra University, Visakhapatnam, India and M.E. from Osmania University, Hyderabad, India. He has teaching experience of more than 20 years in various technical universities in India. Currently working with Andhra Loyola Institute of Engineering & Technology (ALIET), A.P, India. His research interests on wireless sensor networks and OFDM system.