*International Journal of Engineering & Technology*, 7 (4) (2018) 5874-5881

International Journal of Engineering & Technology

*Website: www.sciencepubco.com/index.php/IJET*
*doi: 10.14419/ijet.v7i4.19427*
***Research paper***



# A Bayesian approach for PAPR and MUI reduction in OFDM-based massive MIMO systems

**Budhaditya Bhattacharyya [1] *, Aritra Basu [1]**

*[1] School of Electronics and Communication Engineering, VIT University, Vellore, India*
**Corresponding author E-mail: budhaditya@vit.ac.in*

**Abstract**

Orthogonal frequency division multiplexing (OFDM) based massive multiple-input multiple-output (MIMO) downlink systems face the issues of high peak-to-average power ratio (PAPR) and multiuser interference (MUI) which significantly affect their performance. The solution lies in finding an OFDM-modulated signal that possesses a low PAPR and also enables MUI cancellation. In this paper, a comparative analysis has been performed based on a Bayesian PAPR reduction algorithm. This method models the problem into a hierarchical truncated Gaussian mixture prior (TGM) model which makes use of the redundant degrees of freedom of the transmission array. This leads to a low PAPR signal as most of the samples are concentrated on the boundaries. A variational expectation-maximization (EM) tactic is incorporated to obtain an estimate of the hyperparameters. This is followed by the implementation of the generalized approximate message passing (GAMP) algorithm to reduce the complexity of computation. MATLAB simulations show a significant improvement in PAPR reduction and MUI cancellation with this Bayesian approach leading to better power efficiency and system performance.

***Keywords***: *Bayesian Learning; EM; GAMP; Massive MIMO; MUI; OFDM; PAPR; TGM.*

## 1. Introduction

Massive MIMO has received a lot of attention in the field of wireless communication. It is a promising technology that can successfully meet the growing demands for higher throughput and improved quality of service (QoS) [1]. It makes use of multiple antennas at the transmitter and the receiver to improve the system performance under the condition that the number of transmit antennas at the base station (BS) is larger than the number of receive antennas. It aims to achieve a high spectral efficiency and link reliability.

However, the MIMO systems suffer from frequency selective fading. OFDM is a type of multicarrier modulation (MCM) technique that is used for combating the frequency-selective fading over dispersive channels. But, OFDM suffers from several issues such as high PAPR, which demands the use of expensive and power inefficient linear electronic components at the transmitter. Instead of making use of such high-resolution digital-to-analog converter (DAC) and linear power amplifier (LPA), one can opt for a precoding based PAPR reduction strategy at the transmitter to enable an inexpensive and power efficient solution [2]. Thus, with its potential to improve the energy efficiency and cost by enabling the use of inexpensive, low power components, it is expected that massive MIMO will bring radical changes to future wireless communication systems.

Many PAPR reduction techniques have been developed for single-input single-output (SISO) systems, the most efficient ones being clipping and filtering (CF), selected mapping (SLM) and partial transmission sequence (PTS) [3]. However, the extension of these schemes to multiuser MIMO systems is very complicated. The situation further worsens when the number of antennas is increased as the complexity of the signal processing at the receiver side is compounded with greater user distribution [4]. Therefore, it becomes absolutely necessary to formulate some methods to reduce the PAPR of OFDM-based massive MIMO systems.

A novel Bayesian approach was proposed for OFDM-based downlink massive MIMO systems in [5]. The MUI cancelation was designed as an underdetermined linear inverse problem that presents multiple solutions. To enable a low PAPR solution, a hierarchical truncated Gaussian mixture prior model (TGM) was used and assigned to the solution. This led to a quasi-constant magnitude solution which ensures that maximum entries lie on the truncated boundaries, ensuring a low PAPR. A variational expectation-maximization (EM) algorithm was developed to calculate the estimates of the hyperparameters associated with the TGM prior. Finally, a generalized approximate message passing (GAMP) technique was embedded to reduce the complexity of the algorithm [6].

In this paper, a comparative study has been carried out based on the EM-TGM-GAMP algorithm and the Zero-Forcing (ZF) precoding and the Maximum Ratio Combining (MRC) techniques. The paper is organized as follows: Section 2 reviews the related works, Section 3 describes the massive MIMO system modelling, Section 4 discusses the PAPR reduction technique, Section 5 describes the Bayesian modelling, Section 6 covers the theory of Bayesian inference, Section 7 discusses the mathematical modelling of the likelihood approximation, Section 8 gives a detailed explanation of the variational EM-GAMP framework, Section 9 covers the results of the simulations and Section 10 concludes the paper.

## 2. Related works

A low-complexity SLM based PAPR reduction scheme was proposed in [8]. However, this is only applicable to space frequency block coded (SFBC) MIMO-OFDM systems as it is primarily

**Copyright © 2018 Budhaditya Bhattacharyya, Aritra Basu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.**

based on the linear property of SFBC. Also there is a trade-off between the number of sub-blocks and the PAPR. So, with the proposed scheme, although a large number of signal sets can be computed with only a few IFFTs, it may not achieve a significant PAPR reduction [9]. A PTS based scheme was proposed to resolve the high PAPR problem of OFDM systems in [10]. Although it has a low computational complexity, it fails to achieve any significant improvement in PAPR reduction. [11] deals with the issue of additional side information (SI) added due to SLM which reduces data throughput. Although it produces similar performance as the standard SLM method, it can only be applied to a flat fading channel. [12] discusses a hybrid SLM and PTS based PAPR reduction technique. SLM is applied first to choose the dataset with the least PAPR which is then fed to PTS to further reduce it. However this method involves intensive computation which reduces its efficiency. [13] makes use of a Bayesian approach to mitigate PAPR by utilizing the redundant degrees of freedom (DOF) of the transmit array. It achieves significant improvement in PAPR reduction as compared to the existing schemes, but suffers from a poor convergence rate. A joint PAPR reduction and MUI cancellation was proposed in [15] which also uses a Bayesian approach. It applies a Bernoulli Gaussian prior and seeks a minimum mean square error (MMSE) solution. Although it achieves a significant MUI cancellation, it fails to achieve a low PAPR. Recently, a linear constrained $l_\infty$ optimization based PAPR reduction method was developed for massive MIMO systems to achieve joint MUI cancelation and PAPR reduction [16]. However, it made use of a regularization parameter to strike a balance between the MUI cancellation and PAPR reduction. Another recent work was based on approximate message passing (AMP) based Bayesian inference [17]. This is very similar to the EM-TGM-GAMP algorithm, but the prior distributions are quite different, which does not lead to a quasi-constant magnitude solution.

## 3. System model

Here, an OFDM-based downlink massive MIMO system has been modelled with the BS having $M$ transmit antennas and serving $K$ independent single-antenna users (where, $K \ll M$ ). The total number of OFDM transmit signals is set equal to the number of transmit antennas $M$ , which is split into two sets $\tau$ and $\tau^c$ , with the tones in the first set being used for data transmission and the tones in its complementary set utilized as the guard band (unused tones). Hence, for each tone $n \in \tau$ , the corresponding $K \times 1$ vector $v_n$ comprises of the symbols for $K$ users. The data vector is normalized to satisfy $E[\| v_n \|^2] = 1$ . For each tone, $v_n$ is set as $v_n = 0_{K\times 1}$ so that no signal transmission occurs from the guard band.

To remove MUI, it is necessary to perform precoding at the BS. The signal vector $v_n$ on the $n^{th}$ tone is linearly precoded as (1) where $s_n \in C_{M\times 1}$ is the precoded vector that contains symbols transmitted on the $n^{th}$ sub-carrier through the $M$ antennas respectively, and $A_n \in C_{M\times K}$ represents the precoding matrix for the $n^{th}$ OFDM tone.

$$s_n = A_n v_n \tag{1}$$

Here, we compare the Bayesian approach with the classical ZF and MRC techniques. While ZF aims to achieve complete MUI cancellation, MRC attempts to strike a balance between noise reduction and MUI cancellation. The ZF precoding matrix is represented as (2) where $P_n \in C_{K\times M}$ denotes the MIMO channel matrix associated with the $n^{th}$ tone. After precoding, all precoded vectors are reordered based on the $M$ antenna positions for OFDM modulation as $[\alpha_1 ..... \alpha_M] = [s_1 ..... s_M]^T$ where $\alpha_n \in C_{M\times 1}$ represents the frequency-domain signal transmitted from the $n^{th}$ antenna. This is followed by the inverse discrete Fourier transform (IDFT) to obtain the time-domain signals $\hat{\alpha}_n$ . A cyclic prefix (CP) is then added to eliminate the intersymbol interference (ISI). Finally, these samples are converted to analog signals and transmitted via the frequency selective channel.

$$A_n = P_n^H (P_n P_n^H)^{-1} \tag{2}$$

At the receiver, the CPs removed and DFT is performed to obtain the frequency-domain signals. The received vector consists of $K$ user signals and can be denoted by (3) where $y_n \in C_{K\times 1}$ denotes the receive vector associated with the $n^{th}$ tone and $\varepsilon_n$ is the receiver noise that has independent and identically distributed (i.i.d) circularly symmetric complex Gaussian entries with zero mean and variance $N_o$ .

$$y_n = P_n s_n + \varepsilon_n \tag{3}$$

When the ZF scheme is used, the received signal simplifies to (4) signifying that MUI is completely eliminated. However, for the MRC scheme, the matrix is given by (5) and the received signal is denoted by (6).

$$y_n = v_n + \varepsilon_n \tag{4}$$

$$A_n = \frac{P_n^H}{\| P_n \|^2} \tag{5}$$

$$y_n = \frac{P_n P_n^H}{\| P_n \|^2} s_n + \varepsilon_n \tag{6}$$

## 4. PAPR reduction

In OFDM, the phases of the sub-carriers are independent of each other leading to either constructive or destructive interference. To limit the signal distortion and out-of-band radiation, DACs and LPAs are required at the transmission side to deal with the large peaks generated in the OFDM signal. This results in an expensive and power-inefficient remedy.

PAPR is defined as the ratio of the peak power to the average power of the signal. For the $n^{th}$ transmit antenna, it is given by (7). When the number of transmit antennas is significantly greater than the number of users, multiple precoding matrices are available leading to a set of precoded signals $s = [s_1^T ..... s_N^T]^T$ which can achieve complete MUI elimination. The objective now is to select a suitable candidate $s$ from this set corresponding to a low PAPR solution [19]. Thus, instead of designing a precoding matrix, a signal $s$ can be searched for which can lead to joint PAPR reduction and MUI cancelation.

$$PAPR = \frac{2N \| \hat{a}_n \|_\infty^2}{\| \hat{a}_n \|_2^2} \tag{7}$$

For MUI cancellation, the precoded vectors $s_n$ must satisfy (8) and (9).

$$v_n = P_n s_n , \; n \in \tau \tag{8}$$

$$0_{M\times 1} = s_n , \; n \in \tau^c \tag{9}$$

This can now be modelled as an underdetermined linear inverse problem given by (10) where $\bar{v} \in C_{MK\times 1}$ represents the combined

form of all vectors denoted by $v_n$ and $0_{M\times 1}$ and $P$ represents a block diagonal matrix with its diagonal blocks equal to $P_n$ $\forall n \in \tau$ and $I_M$ $\forall n \in \tau^c$. The re-ordering operation is now equivalent to a linear transformation $a = Tw$ where $a = [a_1^T \ldots\ldots a_M^T]$ and $T$ is a permutation matrix for assigning each precoded vector to the $M$ antennas. With $\hat{a}_m = IDFT(a_m)$, we now have (11) where $\bar{F} = I_M \otimes F_N$ denotes the Kronecker product and $\hat{a} = [\hat{a}_1^T \ldots\ldots \hat{a}_M^T]$. Given a symbol vector $v$, the objective is to search for a signal $\hat{a}$ having a low PAPR. This is a form of the minimax problem, which aims to minimize the maximum PAPR among all antennas. To simplify this issue, a constrained optimization $\min \|\hat{a}\|_\infty$ is used subject to (11). This can be further simplified to a real-valued problem as $\min \|x\|_\infty$ subject to (12). Thus, by minimizing the largest magnitude of entries, one can reduce PAPR of every transmit antenna [16].

$$\bar{v} = \bar{P}s \tag{10}$$

$$\bar{v} = \bar{P}T^T \bar{F}\hat{a} \tag{11}$$

$$y = Ax \tag{12}$$

Where

$$y = \begin{bmatrix} \Re\{\hat{v}\} \\ \Im\{\hat{v}\} \end{bmatrix},\ A = \begin{pmatrix} \Re\{\bar{P}T^T\bar{F}\} & -\Im\{\bar{P}T^T\bar{F}\} \\ \Im\{\bar{P}T^T\bar{F}\} & \Re\{\bar{P}T^T\bar{F}\} \end{pmatrix} \text{ and } x = \begin{bmatrix} \Re\{\hat{a}\} \\ \Im\{\hat{a}\} \end{bmatrix}$$

## 5. Bayesian model

The system is modelled as (13), where $\varepsilon$ denotes the noise vector which is assumed to comprise of i.i.d. Gaussian random variables with zero-mean and unknown variance $\beta^{-1}$. Here $\beta$ is considered to be unknown as the Bayesian framework allows for the determination of the model parameters.

$$y = Ax + \varepsilon \tag{13}$$

To facilitate PAPR reduction, a quasi-constant magnitude solution is required for the underdetermined linear system. Although a constant magnitude signal can lead to minimum PAPR, it is not a practically feasible scenario. Thus, we seek a signal with maximum entries located on the boundary points of an interval $[-v\ v]$ with the remaining entries bounded within the interval to meet the MUI cancelation constraint.

A hierarchical truncated Gaussian mixture (TGM) prior is required to achieve this quasi-constant magnitude solution for a signal $x$. The coefficients of $x$ are assumed to be independent of each other and are assigned to a TGM distribution as (14) $\forall x_i \in [-v\ v]$ where the first component is characterized by a TGM distribution with mean $v$ and variance $\alpha_{i1}^{-1}$ and the second component is characterized by the same with mean $v$ and variance $\alpha_{i2}^{-1}$. $\pi \in [0, 1]$ denotes the mixing coefficient, $\eta_{i1}$ and $\eta_{i2}$ are the normalization constants given by (15) and (16) while $\alpha_1$ and $\alpha_2$ are the precision parameters [20].

$$p(x_i) = \pi \frac{\mathrm{N}(x_i; v, \alpha_{i1}^{-1})}{\eta_{i1}} + (1-\pi)\frac{\mathrm{N}(x_i; v, \alpha_{i2}^{-1})}{\eta_{i2}} \tag{14}$$

$$\eta_{i1} = \frac{1}{2} - \phi(-2v\sqrt{\alpha_{i1}}) \tag{15}$$

$$\eta_{i2} = \phi(2v\sqrt{\alpha_{i2}}) - \frac{1}{2} \tag{16}$$

Bayesian inference involves computation of the logarithm of the prior, which is a cumbersome process. So, a binary latent variable $\kappa_i$ is introduced such that $\kappa_i = 1$ signifies that the first component is selected and $\kappa_i = 0$ implies the second component. The equivalent prior is given by (17) $\forall x_i \in [-v\ v]$ and the distribution of $\kappa_i$ is written as (18).

$$p(x_i \mid \alpha_{i1}, \alpha_{i2}, \kappa_i; v) = \left(\frac{\mathrm{N}(x_i; v, \alpha_{i1}^{-1})}{\eta_{i1}}\right)^{\kappa_i} \left(\frac{\mathrm{N}(x_i; v, \alpha_{i1}^{-1})}{\eta_{i1}}\right)^{(1-\kappa_i)} \tag{17}$$

$$p(\kappa_i, \pi) = (\pi)^{\kappa_i}(1-\pi)^{(1-\kappa_i)} \tag{18}$$

## 6. Bayesian inference

We make use of variational EM strategy for the Bayesian inference. Here, we are dealing with a probabilistic model that comprises of observed data $y$, hidden variables $z$ and unknown deterministic parameters $\theta$ as shown in Fig 1. $Z = \{x, \alpha_1, \alpha_2, \kappa\}$ are considered to be hidden variables, whereas the noise variance $\beta$ and the boundary parameter $v$ are treated as unknown deterministic parameters $\theta = \{\beta, v\}$. The marginal probability of the observed data can be denoted by (19) where the two terms can be further expressed as (20) and (21), where $q(z)$ is the probability density function and $KL(q \| p)$ is the Kullback-Leibler divergence between $p(z \mid y; \theta)$ and $q(z)$ [21].

$$\ln p(y; \theta) = F(q, \theta) + KL(q \| p) \tag{19}$$

$$F(q, \theta) = \int q(z) \ln\left(\frac{p(y, z; \theta)}{q(z)}\right) dz \tag{20}$$

$$KL(q \| p) = -\int q(z) \ln\left(\frac{p(z \mid y; \theta)}{q(z)}\right) dz \tag{21}$$

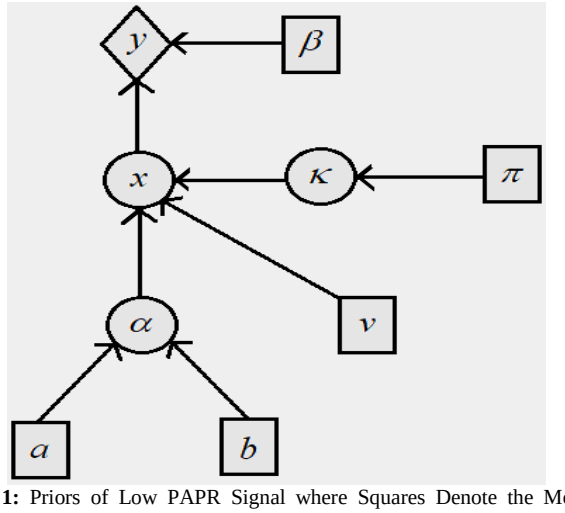



**Fig. 1:** Priors of Low PAPR Signal where Squares Denote the Model Parameters and Circles Represent the Hidden Variables.

## 7. Likelihood approximation with GAMP

GAMP is a very-low-complexity Bayesian iterative technique. The EM-GAMP algorithm is implemented in a dual loop technique where the outer loop (EM) calculates the $Q$-function using $q(x)$ and then maximizes the same function to update the model parameters $\alpha_1$, $\alpha_2$ and $k$. The inner loop (GAMP) makes use of the newly estimated parameters to obtain a new approximation of $q(x)$ [7]. Instead of computing $q(x)$ in the variational EM framework, GAMP is used to calculate the approximate likelihood

function $p(y|x;\beta)$ which does not involve any of the model parameters as shown in Fig 2. Also instead of implementing the inner loop in an iterative way, only a single iteration is used to approximate the likelihood function [22].

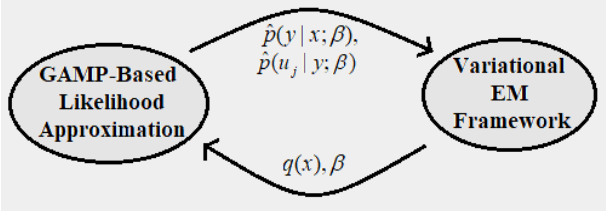



**Fig. 2:** Variational EM-GAMP Framework.

The GAMP algorithm is used to compute the approximation of the joint likelihood function $p(y|x;\beta)$ as a product of approximate marginal likelihoods as given by (22) where $\mathrm{N}(x_i|\hat{r}_i,\tau_i^r)$ is the approximate marginal likelihood [6]. To compute $\hat{r}_i$ and $\tau_i^r$, an estimate of the posterior $q(x)$ and $\beta$ are required. GAMP uses these to approximate the likelihood function $p(y|x;\beta)$. With this approximation, the variational EM gives a new estimate of the posterior distribution $q(x)$, the deterministic variables $\beta$ and $v$ as well as the posterior distributions of the other hidden variables $q(\alpha_1)$, $q(\alpha_2)$ and $q(k)$ [23].

$$p(y|x;\beta) \approx \hat{p}(y|x;\beta) \propto \prod_{i=1}^{I} \mathrm{N}(x_i|\hat{r}_i,\tau_i^r) \tag{22}$$

**Algorithm 1: GAMP-Based Likelihood Approximation**

Input:

- mean and variance of $\hat{x}_i$, where $\hat{x}_i = \langle x_i \rangle_{q(x_i)}$
- mean and variance of $\tau_i^x$, where $\tau_i^x = \langle x_i \rangle_{q(x_i)}^V$
- initialize $\hat{s}_j = 0$

where, $i = 1,.....,I$, $j = 1,.....,J$, $\hat{x}_i$ & $\tau_i^x$ are the posteriors of $q(x_i)$ and $\langle . \rangle_{q(.)}^V$ denotes the variance with respect to $q(.)$ and inverse noise variance $\beta$.

Output:

- approximate likelihoods $\aleph(x_i|\hat{r}_i;\tau_i^r)$ and $\aleph(u_j|\hat{u}_j;\tau_j^u)$, where $i = 1,.....,I$ and $j = 1,.....,J$

Step-1: For each value of $j$:

- $\tau_j^p = \sum_i A_{ji}^2 \tau_i^2$
- $\hat{p}_j = \sum_i A_{ji}^2 \hat{x}_i + \tau_j^p \hat{s}_j$

Step-2: For each value of $j$:

- $\hat{u}_j = \langle u_j \rangle_{p(u_j|y_j;\hat{p}_j;\tau_j^p)}$
- $\tau_j^u = \langle u_j \rangle_{p(u_j|y_j;\hat{p}_j;\tau_j^p)}^V$
- $\hat{s}_j = \dfrac{\hat{u}_j - \hat{p}_j}{\tau_j^p}$
- $\tau_j^s = \dfrac{1}{\tau_j^p}\left(1 - \dfrac{\tau_j^u}{\tau_j^p}\right)$

Step-3: For each value of $i$:

- $\tau_i^r = \left(\sum_j A_{ji}^2 \tau_j^s\right)^{-1}$
- $\hat{r}_i = \hat{x}_i + \tau_i^r \sum_j A_{ji} \hat{s}_j$

# 8. Variational EM-GAMP framework

## 8.1. Expectation (E) step:

This involves updations of the hidden variables. The posterior of $z$ is approximated as (23).

$$p(x,\alpha_1,\alpha_2,k|y;\beta,v) = q(x)q(\alpha_1)q(\alpha_2)q(k) \tag{23}$$

The updations of $q(x)$, $q(\alpha_1)$, $q(\alpha_2)$ and $q(k)$ are represented by (24), (25), (26) and (27) respectively [24].

$$\ln q(x) = \langle \ln p(y,x,\alpha_1,\alpha_2,k;\beta,v) \rangle_{q(\alpha_1)q(\alpha_2)q(k)} \tag{24}$$

$$\ln q(\alpha_1) = \langle \ln p(y,x,\alpha_1,\alpha_2,k;\beta,v) \rangle_{q(x)q(\alpha_2)q(k)} \tag{25}$$

$$\ln q(\alpha_2) = \langle \ln p(y,x,\alpha_1,\alpha_2,k;\beta,v) \rangle_{q(x)q(\alpha_1)q(k)} \tag{26}$$

$$\ln q(k) = \langle \ln p(y,x,\alpha_1,\alpha_2,k;\beta,v) \rangle_{q(x)q(\alpha_1)q(\alpha_2)} \tag{27}$$

When $x_i$ is close to the boundary point $v$, the posterior mean of the precision parameter $\alpha_{i1}$ increases which results in the prior pushing the entry $x_i$ closer to the boundary point $v$. This feed-back mechanism ensures that most of the entries are eventually located on the boundary points.

## 8.2. Maximization (M) step:

This step updates the deterministic parameters. The deterministic parameters are estimated by maximizing the $Q$-function as denoted by (28).

$$\theta^{NEW} = \max_{\theta} Q(\theta,\theta^{OLD}) = \langle \ln p(y,z;\theta) \rangle_{q(z)} \tag{28}$$

The inverse of the noise variance is updated by maximizing the $Q$-function with respect to $\beta$ as given by (29).

$$\beta^{t+1} = \frac{J}{\sum_{j=1}^{J} \langle (y_j - u_j)^2 \rangle} \tag{29}$$

The boundary parameter is updated by maximizing the $Q$-function with respect to $v$ as shown in (30) and (31), where $\hat{x}^{(t)}$ is the estimate of the signal after $t$ iterations and $\gamma = [\gamma_1.....\gamma_I]^T$.

$$v^{t+1} = v^t + \Delta v \tag{30}$$

$$\Delta v = \frac{(y - A\hat{x}^{(t)})^T A\gamma}{\|A\gamma\|_2^2}, \text{ where } \gamma_i = \begin{cases} 1, & \text{if } \hat{x}_i^{(t)} \geq 0 \\ -1, & \text{if } \hat{x}_i^{(t)} < 0 \end{cases} \tag{31}$$

However, this optimization is a complex process. So, a heuristic approach is adopted to compute an approximate value of $v$ such that the mismatch $\|y - A\hat{x}\|_2^2$ can be minimized. Here, $\hat{x}$ represents the estimated signal which is computed as the mean of the posterior distribution $q(x)$. The fundamental principle is based on the fact that since most of the entries are concentrated on the boundary points, increasing the boundary $v$ by a very small value $\Delta v$ will also cause the signal to expand accordingly leading to a lower mismatch [25].

**Algorithm 2: EM-TGM-GAMP**

Initialization:

- $\beta^{(0)} = 10^3$
- $v^{(0)} = \dfrac{\|y\|_\infty}{\|A\|_\infty}$
- $\overline{q(x)} = 0$
- $\overline{q(\alpha_1)} = 1$
- $\overline{q(\alpha_2)} = 1$

- $\overline{q(k)} = \frac{1}{2}$
- $\text{var}(q(x)) = 1$
- $t = 0$

Repeat the following steps until $t \le t_{\max}$ :

- Calculate the approximate distributions $\hat{p}(y \mid x; \beta^{(t)})$ and $\hat{p}(u_j \mid y; \beta^{(t)})$ on the basis of $\overline{q(x)}$, $\overline{\beta^{(t)}}$, $\text{var}(q(x))$ and $\text{var}(\beta^{(t)})$ with the aid of Algorithm 1.
- Using the approximate likelihood $\hat{p}(y \mid x; \beta^{(t)})$, update the posteriors of the hidden variables $q(x)$, $q(\alpha_1)$, $q(\alpha_2)$ and $q(k)$.
- Compute the new estimate of $\beta^{t+1}$ and $v^{t+1}$ using $\beta^{t+1} = \frac{J}{\sum_{j=1}^{J} \langle (y_j - u_j)^2 \rangle}$ and $v^{t+1} = v^t + \frac{(y - A\hat{x}^{(t)})^T A\gamma}{\|A\gamma\|_2^2}$, where $\gamma_i = \begin{cases} 1, & \text{if } \hat{x}_i^{(t)} \ge 0 \\ -1, & \text{if } \hat{x}_i^{(t)} < 0 \end{cases}$
- Increase the value of $t$ as $t = t + 1$

## 9. Results

The performance metrics include a measure of the MUI the complementary cumulative distribution function (CCDF), which is used to evaluate the PAPR reduction performance. The CCDF computes the probability that the PAPR of the estimated signal exceeds a given threshold and is given as (32).

$$CCDF(PAPR_0) = \Pr(PAPR > PAPR_0) \tag{32}$$

The MUI is computed as denoted by (33).

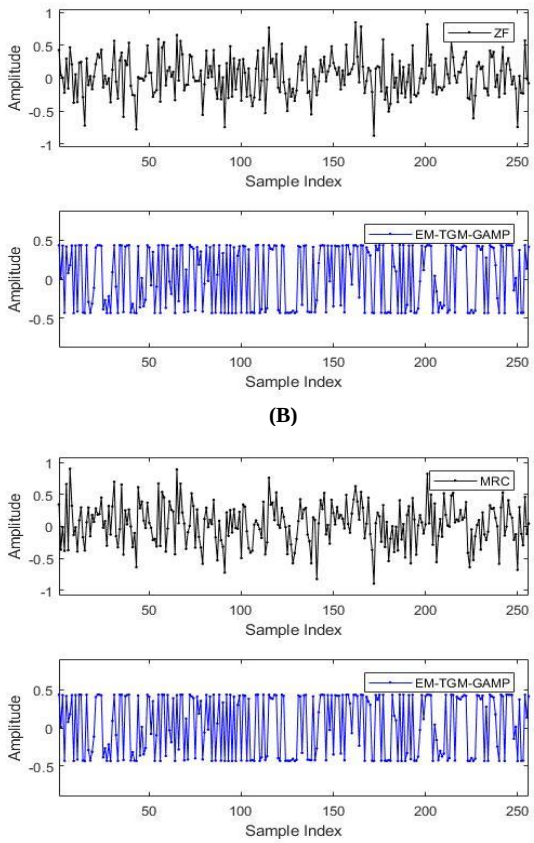
$$MUI = \frac{\sum_{n \in \tau} \| v_n - P_n s_n \|^2}{\sum_{n \in \tau} \| v_n \|^2} \tag{33}$$

MATLAB simulations have been carried out to compare the performance of the Bayesian PAPR reduction and MUI cancellation approach with the ZF and MRC techniques. The method has been referred to as the EM-TGM-GAMP algorithm. Our simulations have been carried out for a massive MIMO system which has 256 antennas at the BS and serves 64 single-antenna users. A BPSK signal constellation has been considered.

The real part of the time-domain signal as estimated by the EM-TGM-GAMP scheme (the complex part shows similar response) has been depicted in Fig. 3. It can be clearly observed that the solution has most of its entries on the boundary points, which is the basic principle behind using the TGM mixture prior model. Such a solution has a low PAPR as it almost behaves like a constant modulus signal. As far as the ZF and MRC schemes are concerned, the solutions have several high peaks and show a large deviation.

The number of trials has been restricted to 100 in our simulations. To choose the number of iterations, the convergence rate of the algorithm is computed with respect to the PAPR reduction and MUI cancellation. It can be observed in Fig. 4 that the algorithm converges at around 300 iterations, the value which has been used for all the simulations.

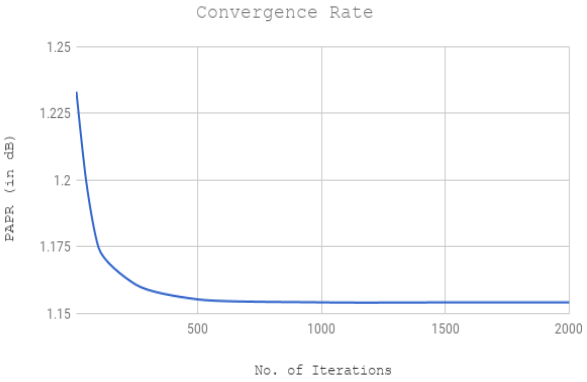
**(A)**




**(B)**

**Fig. 3:** Time Domain OFDM Signal: (A) ZF vs. EM-TGM-GAMP (B) MRC vs. EM-TGM-GAMP

**(A)**




**(B)**

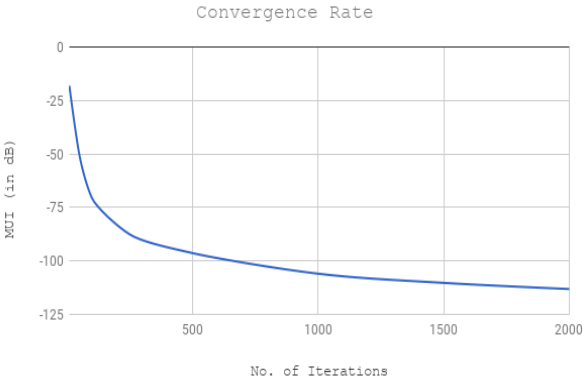
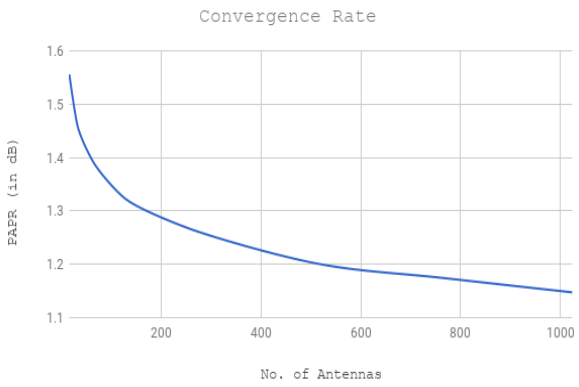



**Fig. 4:** Convergence Rate For Number of Iterations: (A) PAPR (in Db) (B) MUI (in Db) [Suitable Value = 300].

The convergence rate was also calculated for the number of transmit antennas in order to determine a suitable value of $N$ by fixing the number of single antenna users as $K = 64$. It is seen from Fig. 5 that the algorithm converges when around 256 antennas are used which satisfies our model assumption $K \ll M$.

**(A)**




**(B)**

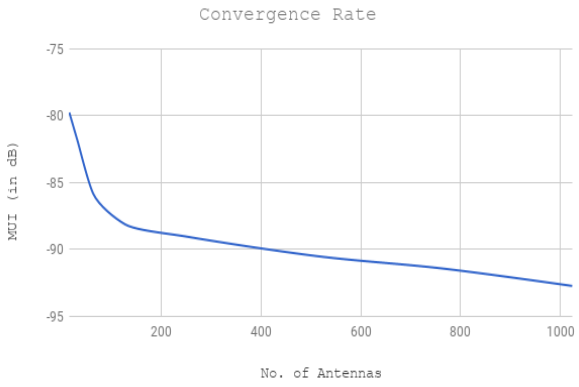
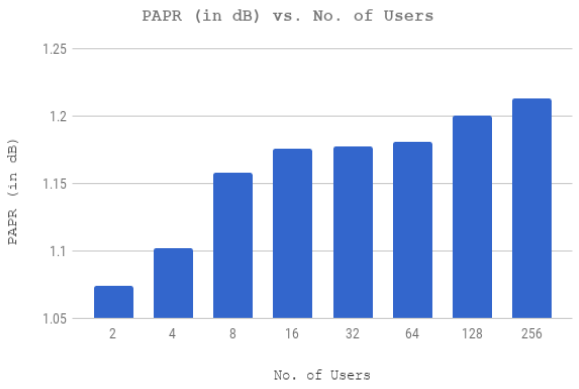



**Fig. 5:** Convergence Rate for Number of Antennas: (A) PAPR (in Db) (B) MUI (in Db) [Suitable Value = 256].

The PAPR reduction and MUI cancellation are also investigated when the number of users is increased. It can be clearly seen from Fig. 6 that an increase in the number of users hinders the performance of the system.

**(A)**




**(B)**

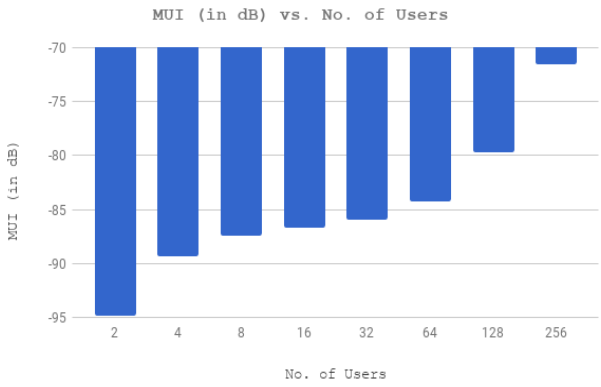
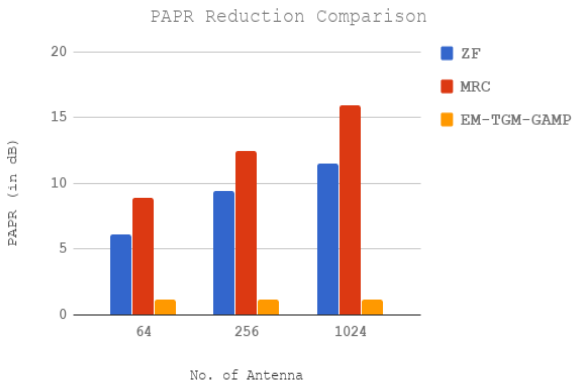



**Fig. 6:** (A) PAPR (in Db) (B) MUI (in Db) For Variable Number of Users (K) Served by K*4 Antennas.

Fig. 7 compares the performance of the Bayesian methodology with the ZF and MRC schemes for varying number of transmit antennas $N$ serving $N/4$ users. The PAPR values (in dB) have been tabulated in Table 1 and the MUI values (in dB) have been enlisted in Table 2. It can be clearly seen that the Bayesian algorithm gives the lowest PAPR. As far as the MUI is concerned, ZF achieves complete MUI cancellation, but MRC fails in that aspect. Thus, the EM-TGM-GAMP scheme shows a substantial reduction in MUI, thereby ensuring a joint MUI cancellation and PAPR reduction attribute.

**(A)**




**(B)**

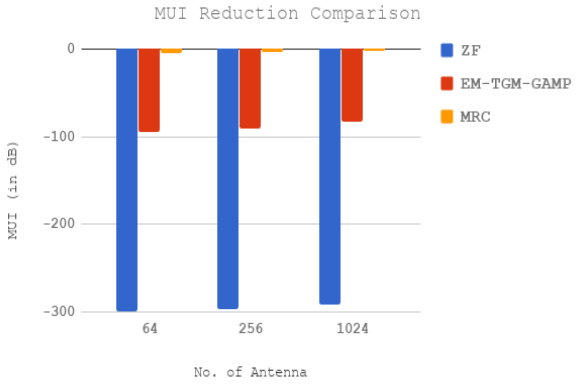



**Fig. 7:** Comparison of (A) PAPR (in Db) (B) MUI (in Db) Reduction for Different Schemes for A Variable Number of Antennas (N) Serving N/4 Users

**Table 1:** PAPR (In Db) For Different Schemes with 300 Iterations for A Variable Number of Antennas (N) Serving N/4 Users

| No. of Antenna | 64 | 256 | 1024 |
|---|---|---|---|
| ZF | 6.1584932 | 9.4655246 | 11.547964 |

| | | | |
|---|---|---|---|
| MRC | 8.8977103 | 12.439619 | 15.988378 |
| EM-TGM-GAMP | 1.1915911 | 1.1818215 | 1.1887495 |

**Table 2:** MUI (In Db) For Different Schemes with 300 Iterations for A Variable Number of Antennas (N) Serving N/4 Users

| No. of Antenna | 64 | 256 | 1024 |
|---|---|---|---|
| ZF | -299.39531 | -296.46173 | -290.93918 |
| MRC | -4.9440519 | -3.5302091 | -2.0260346 |
| EM-TGM-GAMP | -82.275613 | -94.367416 | -90.974505 |

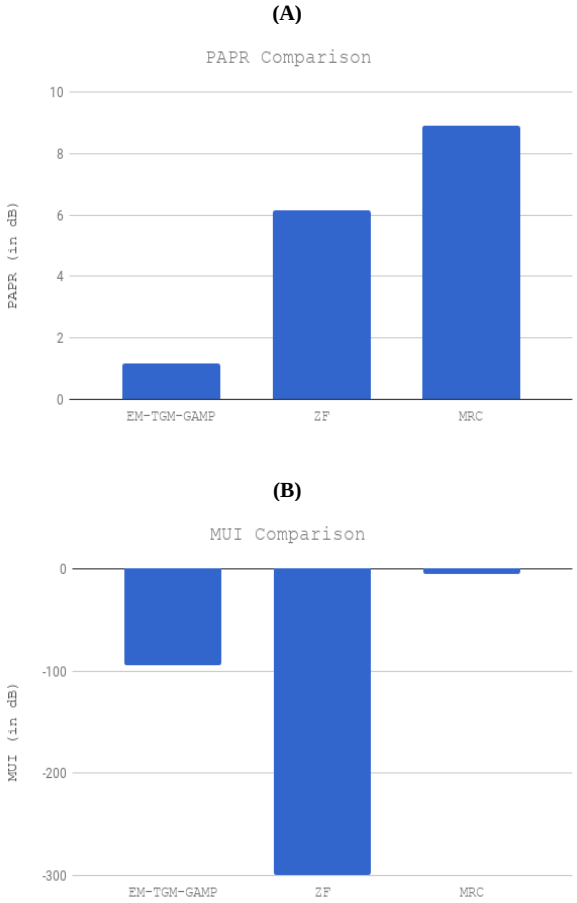



**Fig. 8:** Comparison of (A) PAPR (In Db) (B) MUI (In Db) Reduction for Different Schemes with 300 Iterations for 256 Antennas Serving 64 Users.

The PAPR and MUI values (in dB) obtained by using the Bayesian method for a simulation involving 256 transmit antennas serving 64 single-antenna users with 300 iterations have been tabulated in Table 3. The same has also been plotted in Fig. 8. Fig. 9 presents a CCDF plot of the three schemes under consideration. We can clearly see that the Bayesian approach reduces the PAPR by around 5 dB as compared to ZF and MRC.

**Table 3:** PAPR & MUI for Different Schemes with 300 Iterations for 256 Antennas Serving 64 Users

| Scheme | PAPR (in dB) | MUI (in dB) |
|---|---|---|
| EM-TGM-GAMP | 1.1818215 | -94.3674157 |
| ZF | 6.1584932 | -299.395307 |
| MRC | 8.8977103 | -4.94405194 |

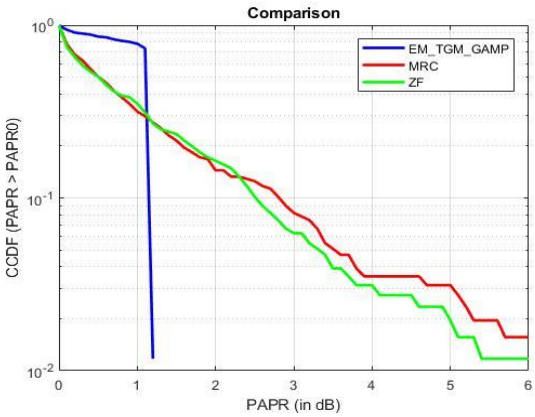



**Fig. 9:** CCDF Plot for Comparison of Reduction of PAPR by EM-TGM-GAMP, ZF and MRC Schemes

## 10. Conclusion

This paper presents a solution to the problem of high PAPR in an OFDM-based downlink massive MIMO system with a Bayesian inference-based joint PAPR reduction and MUI cancelation technique. The system is modelled as a hierarchical TGM prior model to facilitate the choice of a low PAPR signal. A variational EM algorithm is embedded to calculate the estimates of the hyperparameters of the prior. Finally, the GAMP technique is combined with the EM framework to improve the convergence rate. A comparative study has been performed with the popular existing schemes such as ZF and MRC. Simulation results show that the Bayesian scheme achieves a significant reduction in PAPR and even ends up performing improved MUI cancelation. It also shows a fast convergence rate, which makes it a suitable candidate for real-time systems.

## References

[1] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," IEEE Signal Process. Mag., vol. 30, no. 1, pp. 40–60, Jan. 2013. https://doi.org/10.1109/MSP.2011.2178495.

[2] 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. https://doi.org/10.1109/MSP.2012.2218138.

[3] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for OFDM signals," IEEE Trans. Broadcasting, vol. 54, no. 2, pp. 257–268, Jun. 2008. https://doi.org/10.1109/TBC.2008.915770.

[4] 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. https://doi.org/10.1109/MWC.2005.1421929.

[5] Hengyao Bao, Jun Fang, Zhi Chen, Hongbin Li, Senior Member, IEEE, and Shaoqian Li" An Efficient Bayesian PAPR Reduction Method for OFDM-Based Massive MIMO Systems, IEEE Transactions on Wireless Communications, 15(6), 4183-4195. https://doi.org/10.1109/TWC.2016.2536662.

[6] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," IEEE International Symposium, Information Theory Proceedings (ISIT), pp. 2168- 2172, Aug. 2011. https://doi.org/10.1109/ISIT.2011.6033942.

[7] J. Tellado, "Peak to average power reduction for multi-carrier modulation", PhD thesis, Stanford University, 2000.

[8] C. L. Wang, S. S. Wang, & H. L. Chang, "A low-complexity SLM based PAPR reduction scheme for SFBC MIMO-OFDM systems", Wireless Communications and Networking Conference (WCNC), IEEE, pp. 1449-1453, 2011. https://doi.org/10.1109/WCNC.2011.5779373.

[9] Y. L. Lee, Y. H. You, W. G. Jeon, J. H. Paik, & H. K. Song, "Peak-to-average power ratio in MIMO-OFDM systems using selective mapping", IEEE Communications letters, 7(12), 575-577, 2003. https://doi.org/10.1109/LCOMM.2003.821329.

[10] S. J. Ku, C. L. Wang, & C. H. Chen, "A reduced-complexity PTS-based PAPR reduction scheme for OFDM systems", IEEE Transactions on Wireless Communications, 9(8), 2455-2460, 2010. https://doi.org/10.1109/TWC.2010.062310.100191.

[11] S. A. Adegbite, S. G. McMeekin, & B. G. Stewart, "A PAPR reduction and data decoding for SLM based OFDM systems without SI", Vehicular Technology Conference (VTC Spring), IEEE, pp. 0001-0005, May 2015. https://doi.org/10.1109/VTCSpring.2015.7145638.

[12] A. Thakur & N. Dhillon, "Hybrid Approach using SLM and PTS Techniques to Reduce PAPR", International Journal of Science and Research (IJSR), Vol. 4, Issue 5, pp. 173-177, May 2015.

[13] K. Padarti & N. Venkateswara, "A novel method for joint PAPR mitigation in OFDM based massive MIMO downlink systems", International Journal of Engineering & Technology (IJET), 7 (3), pp. 1185- 1188, 2018. https://doi.org/10.14419/ijet.v7i3.13009.

[14] 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. https://doi.org/10.1109/MWC.2005.1421929.

[15] K. Prasanna K, "Reducing the PAPR of OFDM Based MIMO Systems Using Bayesian Approach", International Journal of Advanced Technology and Innovative Research, Vol.09, Issue 06, pp. 0958-0963, May 2017.

[16] C. Studer and E. G. Larsson, "PAR-aware large-scale multi-user MIMO OFDM downlink," IEEE J. Sel. Areas Commun., vol. 31, no. 2, pp. 303–313, Feb. 2013. https://doi.org/10.1109/JSAC.2013.130217.

[17] 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.

[18] 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. https://doi.org/10.1049/el:19961384.

[19] C. Tellambura, "Computation of the continuous-time PAR of an OFDM signal with BPSK subcarriers," IEEE Commun. Lett., vol. 5, no. 5, pp. 185–187, May 2001. https://doi.org/10.1109/4234.922754.

[20] M. Tipping, "Sparse Bayesian learning and the relevance vector machine," Journal of Machine Learning Research, vol. 1, pp. 211–244, 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. https://doi.org/10.1109/MSP.2008.929620.

[22] 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, Jun. 2013. https://doi.org/10.1109/LSP.2013.2256783.

[23] 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. https://doi.org/10.1109/TSP.2013.2272287.

[24] M. Abramowitz and I. A. Stegun, Handbook of mathematical functions: with formulas, graphs, and mathematical tables. New York: Dover Publication, 1965. https://doi.org/10.1115/1.3625776.

[25] J. Vila, P. Schniter, S. Rangant, F. Krzakala, and L. Zdeborovd, "Adaptive damping and mean removal for the generalized approximate message passing algorithm," in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), South Brisbane, QLD, pp. 2021-2025, Apr. 19-24, 2015. https://doi.org/10.1109/ICASSP.2015.7178325.