

RESEARCH

Open Access

Performance evaluation of time-multiplexed and data-dependent superimposed training based transmission with practical power amplifier model

Toni Levanen^{*}, Jukka Talvitie and Markku Renfors

Abstract

The increase in the peak-to-average power ratio (PAPR) is a well known but not sufficiently addressed problem with data-dependent superimposed training (DDST) based approaches for channel estimation and synchronization in digital communication links. In this article, we concentrate on the PAPR analysis with DDST and on the spectral regrowth with a nonlinear amplifier. In addition, a novel Gaussian distribution model based on the multinomial distribution for the cyclic mean component is presented. We propose the use of a symbol level amplitude limiter in the transmitter together with a modified channel estimator and iterative data bit estimator in the receiver. We show that this setup efficiently reduces the regrowth with the DDST. In the end, spectral efficiency comparison between time domain multiplexed training and DDST with or without symbol level limiter is provided. The results indicate improved performance for DDST based approaches with relaxed transmitter power amplifier requirements.

Keywords: channel estimation, data-dependent superimposed pilots, iterative receiver, nonlinear power amplifier, peak-to-average power ratio, spectral efficiency.

1 Introduction

Channel estimation and equalization are crucial parts of modern digital transmission links. As we aim for higher spectral efficiencies, the number of time instances allocated for training in the traditional time-domain multiplexed training (TDMT) systems should be minimized. At the moment, the superimposed (SI) scheme is a serious candidate for circumventing this issue, see for example [1-3] and references therein. SI pilots are added directly on top of the user data, and thus all time instances over the whole allocated spectral region contain user information. The downside is that the user information interferes greatly with the pilot sequence, increasing the mean squared error (MSE) of the initial channel estimates. Furthermore, the peak-to-average power ratio (PAPR) is considerably increased and the user-data-symbol-to-interference power ratio is decreased in detection.

To overcome this problem of self-interference (interference from the user data symbols in channel estimation), a data-dependent superimposed training (DDST)

scheme was presented in [4,5]. The basic idea is very simple. Because the cyclic pilot sequence has its energy concentrated on certain frequency bins, we set the user data frequency response to zero on these frequency bins. This is equivalent to removing the cyclic mean of the user data symbol sequence in the time domain. Therefore, there is no interference from the user data to the pilot symbols. Because the interference from the user data symbols is removed, DDST requires clearly lower pilot powers than traditional SI training to obtain the desired channel estimation MSE levels. This can also be seen as frequency-domain multiplexed (FDM) pilot based training, but the difference to the traditional approach is that the signal spectrum is not widened because of the used SI training symbols. With multicarrier systems, spectral nulling means that we lose some subcarriers for pilot symbols. Recently, a solution to circumvent this problem in multicarrier communications by the so called symbol blanking method was proposed in [6].

The DDST is suitable especially for wide-band single-carrier (SC) systems. The problem to be addressed in this article regarding the addition of DDST sequences is the increased peak power (PP) and PAPR, which violates one

* Correspondence: toni.levanen@tut.fi

Department of Communications Engineering, Tampere University of Technology, P.O. Box 553, FIN-33101, Finland

of the main benefits of using SC transmission. With increased PAPR we can expect increased spectral regrowth with nonlinear amplifiers, which are preferred in the mobile devices because of their higher efficiency. Based on the authors best knowledge, the effects of increased PP or PAPR on the spectral regrowth have not been taken into account in the recent literature in the performance comparisons between DDST and TDMT systems. More traditional SI-based training was studied in [7], where the frequency bins were in some cases nulled for improved channel estimation performance. The PAPR problem was discussed without any solutions to decrease the PAPR created by the SI pilots. We will address this problem by simply limiting the peak amplitudes at the symbol level before transmission. From now on, this symbol level amplitude limited DDST is denoted as LDDST.

In the receiver side, we have a simple feedback loop based on soft symbol estimates, which we use to estimate the missing cyclic mean and the limited amplitudes. In [8], we studied the symbol level PAPR and used an iterative receiver structure without any knowledge of the error generated by the symbol level amplitude limiter in the transmitter. In this article we will utilize the scaling information available based on Gaussian modeling of the data-dependent pilot sequence (cyclic mean) in the channel estimator.

This article is structured as follows. First we present the system model in Section 2. Then, in Section 3 we model the error caused by the symbol level limiter in the transmitted signal. Next, in Section 4 we briefly discuss the modifications used in the channel estimation algorithms because of the symbol level limiter. In Section 5, we concentrate on the symbol level PP and PAPR, on the PP and PAPR after the transmit pulse shape filtering, and show that the symbol level limiter can remove the PP increase and effectively reduce the PAPR. In addition, we discuss the spectral re-growth related to different training methods. In the Section 6, we provide improved iterative receiver algorithms taking into consideration the amplitude limiter in the transmitter and the removal of the data dependent pilots. Next, in Section 7, the throughput performance comparison of DDST and TDMT training based systems is provided. Finally, in Section 8, conclusions are provided.

Notation: Superscripts T and H denote the transpose and Hermitian transpose operators, \otimes refers to the Kronecker product and \circ defines a continuous-time convolution. For complex numbers $|z|$ defines the absolute value of z and $\angle \cdot$ gives the argument of a complex number. In addition, $\text{Re}(z)$ takes the real value of a complex number and $\text{Im}(z)$ takes the imaginary value. Exponential function is noted by $\exp(\cdot)$ and $\|\mathbf{z}\|$ defines the Euclidean vector norm. The trace and statistical expectations are denoted by $\text{tr}[\cdot]$ and $E[\cdot]$. Rounding to the largest integer

not greater than x is given by the floor function $\lfloor x \rfloor$. The $(N \times N)$ identity matrix is denoted by \mathbf{I}_N and the $(N \times M)$ matrix of all ones by $\mathbf{1}_{N \times M}$. For oversampling, we define a column vector \mathbf{r} with first element equal to one and $i - 1$ zeros after the first element, e.g., $\mathbf{r} = [1, 0, \dots, 0]^T$. We denote the length of this vector with r , which will represent the oversampling rate used in the receiver. Matrices are denoted by boldface uppercase letters and vectors by boldface lowercase letters. Finally, $\text{diag}(\mathbf{a}) = \text{diag}(a_1, \dots, a_n)$ is an $(N \times N)$ diagonal matrix whose n th entry is a_n and $\text{diag}(\mathbf{A})$ is a $(N \times 1)$ vector with values from the main diagonal of \mathbf{A} , which is a $(N \times N)$ square matrix.

2 System model

Our system design originates from the uplink assumption. Thus, the complexity of the transmitting end is kept as small as possible and most of the complexity is positioned to the receiving end. The block level design of the transmitter is given in Figure 1. The transmitter contains a bit source, channel encoder, interleaver (represented by π function), symbol mapper, pilot insertion, symbol level amplitude limiter, $L(\cdot)$, the transmitter pulse shape filter and nonlinear amplifier, $G(\cdot)$.

Let us assume that our symbol mapper produces a vector of data symbols \mathbf{d} from some finite alphabet \mathcal{A}^N , where N is the frame (vector) length. We will use a pilot sequence, \mathbf{p} , which has length N_p . The pilot sequence is an optimal channel independent (OCI) sequence that was defined in [2], and rewritten here as

$$p(k) = \sigma_p e^{j \frac{\pi}{N_p} |k(k+\nu)|}, \quad (1)$$

where $k = 0, \dots, N_p - 1$, $\nu = 1$ if N_p is odd and $\nu = 2$ if N_p is even number. In addition, we assume that our frame length is an integer multiple of N_p , given as $N = N_c N_p$, where N_c is the number of cyclic copies per frame. With the DDST, we first remove the cyclic mean of the data vector. As shown in [4], this can be expressed as

$$\mathbf{z} = (\mathbf{I} - \mathbf{J}_{Tx})\mathbf{d}, \quad (2)$$

where $\mathbf{J}_{Tx} = (1/N_c)\mathbf{1}_{N_c \times N_c} \otimes \mathbf{I}_{N_p}$. Now the data dependent pilot sequence is given as $\mathbf{p}_d = -\mathbf{J}_{Tx}\mathbf{d}$. The data dependent pilot sequence is added on top of the data sequence in order to remove the cyclic mean of the data sequence, thus removing the interference caused by data sequence on the known pilot sequence. The symbol sequence including user data symbols, data dependent pilot sequence and the cyclic pilot sequence is given as $\mathbf{s} = \mathbf{d} + \mathbf{p}_d + \mathbf{p}_c = \mathbf{z} + \mathbf{p}_c$, where the cyclic pilot sequence is defined as $\mathbf{p}_c = \mathbf{1}_{N_c \times 1} \otimes \mathbf{p}$. For a more detailed explanation on DDST, see for example [9] and references therein. The symbol sequence, \mathbf{s} , is then inserted to the peak amplitude

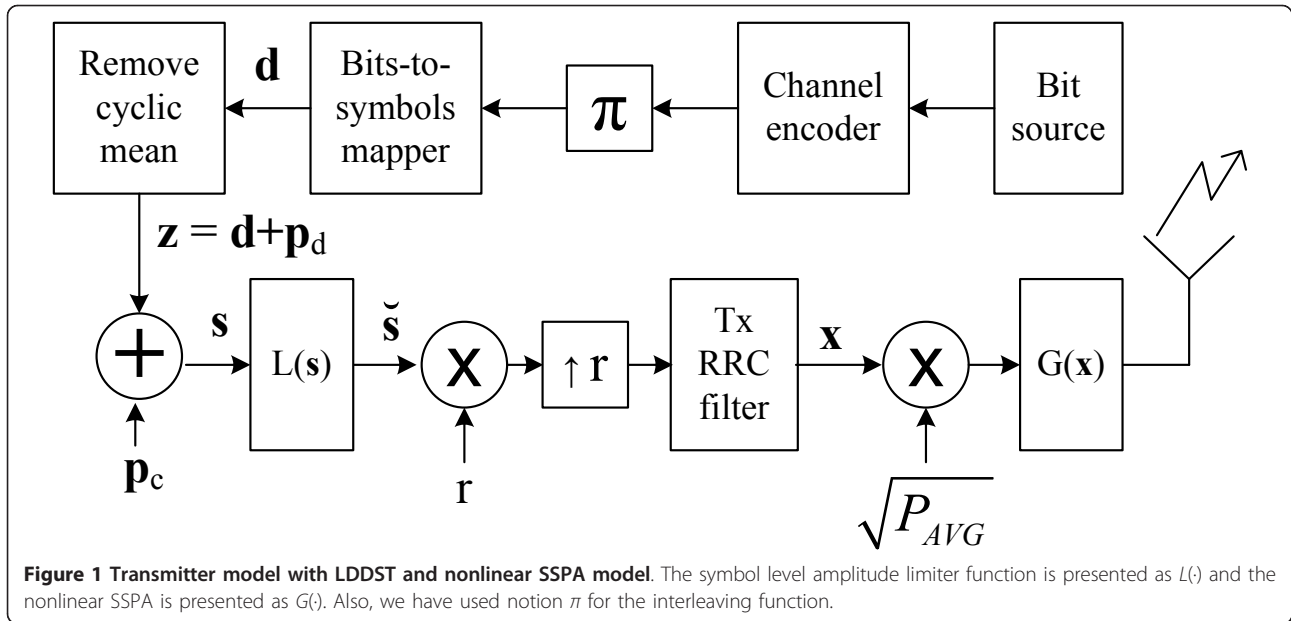


Figure 1 Transmitter model with LDDST and nonlinear SSPA model. The symbol level amplitude limiter function is presented as $L(\cdot)$ and the nonlinear SSPA is presented as $G(\cdot)$. Also, we have used notion π for the interleaving function.

limiter from which the limited signal \tilde{s} is then obtained. This sequence is then oversampled with rate r , given as $\tilde{s}_r = r\tilde{s} \otimes \mathbf{r}$, and inserted to the transmit pulse shape filter to obtain transmitted sequence \mathbf{x} . We define the power of the data sequence to be $\sigma_d^2 = 1 - \gamma$ and the power of the known pilot sequence to be $\sigma_{pc}^2 = \gamma$, where γ is the pilot power allocation factor.

The peak amplitude limiter is presented by a function $L(\cdot)$, which takes as the maximum allowed amplitude value, a_{\max} , the maximum amplitude value of the used constellation \mathcal{A} , defined as $\{a_{\max} = \max(|d|), d \in \mathcal{A}, \sigma_d^2 = 1\}$. We use this value because we wanted to achieve similar type of PAPR behavior as with TDMT and that the limiter affects mainly pilot sequences added on top of the user data. The limited symbol sequence can be defined as

$$\tilde{s}(k) = L(s(k)) = \begin{cases} s(k), & \text{if } |s(k)| \leq a_{\max}, \\ a_{\max} \cdot \exp(j\angle s(k)), & \text{if } |s(k)| > a_{\max}. \end{cases} \quad (3)$$

Now we have an amplitude limited symbol sequence whose PP is limited to the same value as the original data symbol sequence \mathbf{d} . The average power decrease, and the remaining PAPR increase, depends on the constellation. This kind of amplitude limiter, which keeps the argument difference between input and output as a constant, realizes so-called amplitude-modulation to amplitude-modulation (AM-AM) conversion [10], meaning that $|L(s(k))|$ depends only on $|s(k)|$.

We have chosen to study the hard limiting of the transmitted symbols, but of course other limiters with different input-output mappings require more studies. Furthermore, we have chosen to study symbol level

limiting instead of limiting the output of the Tx pulse shape filter, which is a more common approach for controlling the PAPR in SC transmission. From the literature concerning studies on PAPR with OFDM modulation, one can find several possible topics of study in order to reduce PAPR in DDST with a modified data-dependent pilot sequence, and these are left for future studies.

Let us define an error vector $\mathbf{e}_{\text{limiter}} = \tilde{s} - \mathbf{s}$, which contains the information removed by the limiter from the sequence \mathbf{s} . It represents an additive error sequence generated by the limiter. This model is used when we present the receiver feedback structure in Section 7.

The signal after the symbol level limiter, \tilde{s} , is then fed to the transmit pulse shape filter after over-sampling. We have used traditional root-raised-cosine (RRC) filtering with rolloff factor $\rho = 0.1$ and filter order $N_{\text{RRC}} = 64$. We have chosen two different scenarios for simulations. For the PAPR and spectral leakage simulations we have used four times oversampling, $r = 4$, and for the performance evaluations we have used two times oversampling, $r = 2$. We have chosen this setup for better understanding of the spectral spreading and because the used filter bank (FB) based equalizer is designed to work with two times oversampled sequences.

The nonlinear power amplifier model is a widely-used basic model, based on solid-state power amplifier (SSPA) model by Rapp [11]. The AM-to-AM conversion function for an input amplitude A is given as

$$G(A) = v \frac{A}{\left(1 + \left[\frac{vA}{A_0}\right]^{2p}\right)^{-2p}}, \quad (4)$$

where ν is the small signal amplification, A_0 is the saturation amplitude of the amplifier and p defines the smoothness of the transition from linear region to the limiter region. The actual values chosen for the simulations are discussed in more detail in Section 7.

Based on Bussgang's theorem [12], we model the output of the power amplifier as $G(\mathbf{x}) = \alpha\sqrt{P_{\text{AVG}}}\mathbf{x} + \mathbf{n}_G$, where α is a scaling factor for the input signal, P_{AVG} is the average power of the transmitted frame, and \mathbf{n}_G is uncorrelated Gaussian noise vector caused by the nonlinear power amplifier $G(\cdot)$. P_{AVG} is used to scale the average power of the transmitted frame in order to stay inside the spectral mask to be defined in Section 5. The Bussgang's theorem is based on Gaussian variables, but it's results are widely used, e.g., in PAPR modeling for orthogonal frequency domain multiplexing (OFDM) systems. Also in our case, the signals are not purely Gaussian, but after the pulse shape filter they are Gaussian like and we can apply Bussgang's theorem to model the non-linear limiting caused by the power amplifier model.

We have assumed a discontinuous block wise transmission where the channel is assumed to be time invariant during the transmission time of one frame. The used channel model is a modified ITU-R Vehicular A channel [13].

In Figure 2, we have presented a block diagram of our multiantenna receiver. We have extended the model provided in [4] to our SC model with FB-based frequency-domain equalizer structure, presented in [14]. The analysis FB converts the time domain signal to the frequency domain (similar to the well known DFT operation) and the synthesis FB converts the frequency domain presentation back to time domain (similar to the IDFT operation). The channel estimates are obtained in time domain after which the sub-channel wise equalization (SCE) is performed in the frequency domain with 3-tap complex FIR filter for each sub-channel. The equalizers for each diversity branch are designed based on the maximum ratio combining (MRC) criteria, presented in [15]. The channel estimates could also be obtained in the frequency domain and after suitable interpolation with DDST they could be directly used for defining the SCE equalizer tap values for each sub-channel. The FB-based receiver structure is used because it does not require a cyclic prefix (improved throughput), provides close to ideal linear equalizer performance, has good spectral containment properties (adjacent channel suppression is clearly better than with DFT based solutions) and is equally applicable also to SC-FDMA (DFT-S-OFDMA) as used in 3GPP-LTE uplink.

We assume perfect synchronization in frequency and time domain and ideal down conversion of the received signal in the Rx block. Several studies on DDST suitability for time and frequency synchronization have been

performed, e.g., [16,17], where it has been shown that DDST is also a viable solution for low SNR synchronization. We can present the channel between transmitter and receiver as an r times oversampled discrete-time equivalent channel, $h_{\text{eq}}(n) = |h_{\text{RRC}}(t) \circ h_{\text{channel}}(t) \circ h_{\text{RRC}}(t)|_{t=nT/r} = |h_{\text{RRC}} \circ h_{\text{channel+RRC}}|_{t=nT/r}$. The n th received sample $y_i(n)$ from the i th antenna can be given as

$$y_i(n) = \alpha\sqrt{P_{\text{AVG}}} \sum_{m=0}^{M-1} h_{\text{eq},i}(m)\tilde{s}_r(n-m) + \sum_{k=0}^{K-1} h_{\text{channel+RRC},1}(k)n_G(n-k) + \sum_{l=0}^{L-1} h_{\text{RRC}}(l)w_i(n-l), \quad (5)$$

where M is the channel length in samples, n is the time index for r times oversampled symbol sequence, $n_G(n)$ is a noise term caused by the nonlinear amplifier, and $\tilde{s}_r(n)$ is a possibly limited, oversampled transmitted symbol, which is zero if $n < 0$ or $n > rN - 1$. The noise term $w_i(n)$ is complex additive white Gaussian noise (AWGN). Because of the r times oversampling, in our case $s(k) = d(k) = p_d(k) = p_c(k) = 0$ when k modulus $r \neq 0$. The channel estimation procedures are simply repeated for each diversity branch. For this reason and for the sake of clarity, we drop out the antenna index i .

We can now rewrite the received discrete-time signal in the matrix notation as

$$\mathbf{y} = \alpha\sqrt{P_{\text{AVG}}}\tilde{\mathbf{S}}_r\mathbf{h}_{\text{eq}} + \mathbf{N}_G\mathbf{h}_{\text{channel+RRC}} + \mathbf{W}\mathbf{h}_{\text{RRC}}, \quad (6)$$

where the matrix $\tilde{\mathbf{S}}_r = \mathbf{D}_r + \mathbf{P}_{d,r} + \mathbf{P}_{c,r} + \mathbf{E}_{\text{limiter},r}$ is built from the oversampled user data symbols, data dependent pilot sequence, known cyclic pilot sequence and the additional error generated by the symbol level limiter (only with LDDST), respectively. Here \mathbf{N}_G and \mathbf{W} are the matrix presentations of the amplifier induced and channel induced noise terms, respectively.

Because we assume a discontinuous block-wise transmission, all matrices \mathbf{D}_r , $\mathbf{P}_{d,r}$, $\mathbf{P}_{c,r}$ and $\mathbf{E}_{\text{limiter},r}$ have the form

$$\mathbf{B} = \begin{bmatrix} b_0 & 0 & \cdots & 0 & 0 \\ b_1 & b_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{rN_p-1} & b_{rN_p-2} & \cdots & b_1 & b_0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{rN-1} & b_{rN-2} & \cdots & b_{rN-rN_p+1} & b_{rN-rN_p} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & b_{rN-1} \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (7)$$

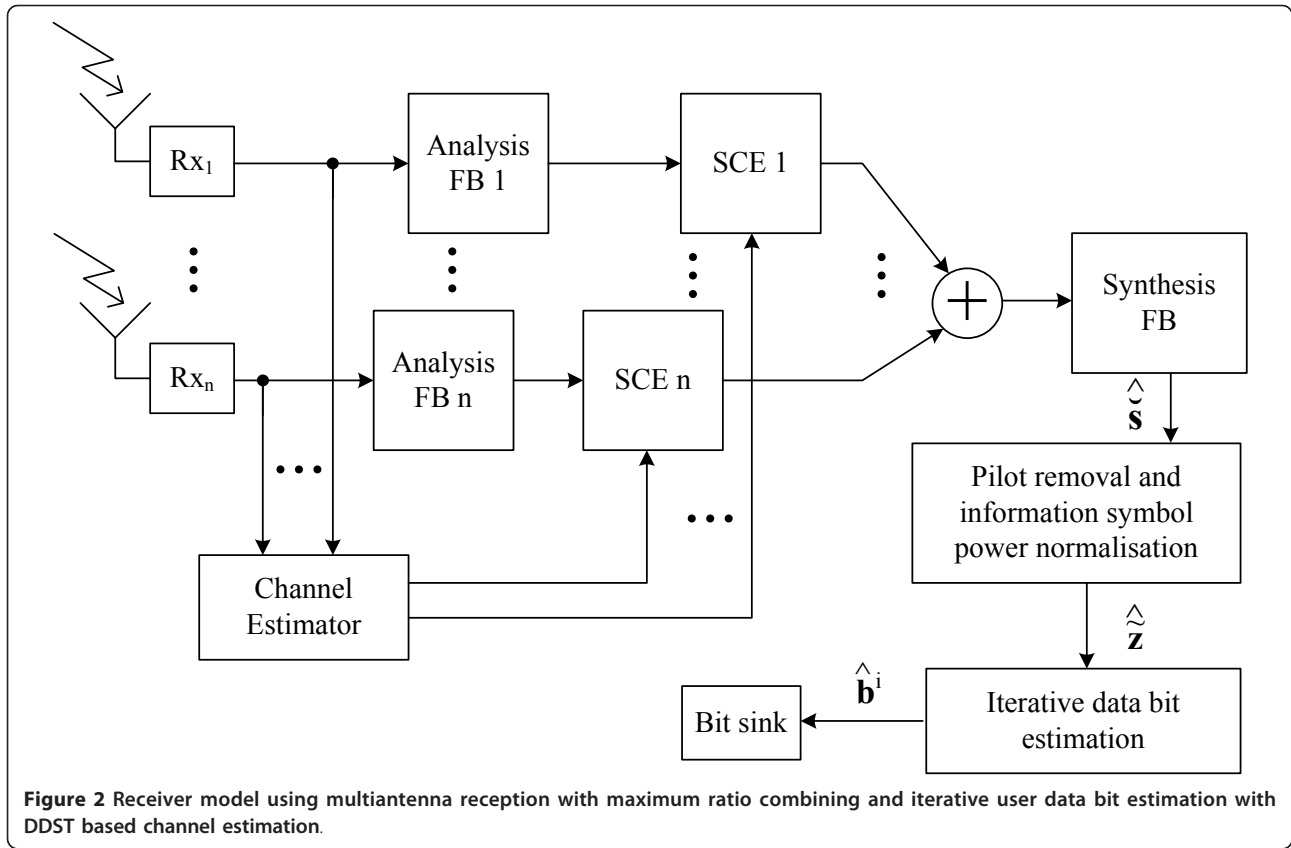


Figure 2 Receiver model using multiantenna reception with maximum ratio combining and iterative user data bit estimation with DDST based channel estimation.

including the zeros before and after the transmitted frame. Note that the oversampled matrices \mathbf{D}_r , $\mathbf{P}_{d,r}$, $\mathbf{P}_{c,r}$, $\mathbf{E}_{\text{limiter},r}$ are now of dimension $(rN + rN_p \times rN_p)$ and that we have assumed that $M = rN_p$. This means that in the receiver we have to do the cyclic mean calculation over $N_c + 1$ copies. Thus, the cyclic mean of the received sequence is given as

$$\begin{aligned} \hat{\mathbf{m}}_y &= \mathbf{J}_{R_x} \mathbf{y} \\ &= \alpha \sqrt{P_{\text{AVG}}} [\mathbf{P}_r + \hat{\mathbf{M}}_{e_{\text{limiter},r}}] \mathbf{h}_{\text{eq}} \\ &\quad + \hat{\mathbf{M}}_{n_c} \mathbf{h}_{\text{channel}+\text{RRC}} + \hat{\mathbf{M}}_w \mathbf{h}_{\text{RRC}}, \end{aligned} \quad (8)$$

where $\mathbf{J}_{R_x} = (1/N_c) \mathbf{1}_{1 \times N_c+1} \otimes \mathbf{I}_{rN_p}$. In our notation, for any vector \mathbf{b} , the cyclic mean vector is defined as $\hat{\mathbf{m}}_b = \mathbf{J}_{R_x} \mathbf{b} = [\hat{m}_b(0) \hat{m}_b(1) \dots \hat{m}_b(rN_p - 1)]^T$, and for any matrix \mathbf{B} , the cyclic mean matrix is defined as

$$\hat{\mathbf{M}}_b = \mathbf{J}_{R_x} \mathbf{B} = \begin{bmatrix} \hat{m}_b(0) & \hat{m}_b(rN_p - 1) & \dots & \hat{m}_b(2) & \hat{m}_b(1) \\ \hat{m}_b(1) & \hat{m}_b(0) & \dots & \hat{m}_b(3) & \hat{m}_b(2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{m}_b(rN_p - 1) & \hat{m}_b(rN_p - 2) & \dots & \hat{m}_b(1) & \hat{m}_b(0) \end{bmatrix}. \quad (9)$$

For example, if you set $\mathbf{b} = \mathbf{e}_{\text{limiter},r}$ then $\hat{\mathbf{M}}_{e_{\text{limiter},r}}$ is a cyclic matrix having $\hat{\mathbf{m}}_{e_{\text{limiter},r}}$ as the first column. The pilot matrix \mathbf{P}_r is a cyclic matrix, having the r times

oversampled OCI pilot sequence $\mathbf{p}_r = r\mathbf{p} \otimes \mathbf{r}$ as its first column.

From the receiver frontend, the oversampled signal is provided for the channel estimator and for the analysis FB. After obtaining a channel estimate, SCE is performed in the frequency domain. More details on the equalizer structure can be found from [14,18], and references therein. After the SCE, different antenna branches are added together sub-channel wise according to the MRC principle. The composite sub-channels are then recombined in the synthesis FB, which also efficiently realizes the sampling rate reduction by 2.

After the synthesis FB, we have the Pilot removal and information symbol power normalization block. Inside this block, the received sequence power is normalized to $\sigma_s^2 = 1 + \sigma_w^2 \|\mathbf{h}_{\text{RRC}}\|^2$, which corresponds to the total received power. We have assumed that we exactly know the noise variance in the receiver. Next, we scale the power based on the pilot power allocation and remove the cyclic mean of the received sequence. If we use LDDST, we normalize the sequence based on our estimate on the average transmit power σ_s^2 , to be defined in (18), to obtain an estimate for the distorted data sequence,

$$\hat{\mathbf{z}} = \sigma_{\tilde{\mathbf{s}}} (\mathbf{I} - \mathbf{J}) \sqrt{\frac{1}{1 - \gamma}} \sqrt{\frac{1 + \sigma_w^2 \|\mathbf{h}_{\text{RRC}}\|^2}{\sigma_{\tilde{\mathbf{s}}}^2}} \hat{\mathbf{s}}. \quad (10)$$

Here $\hat{\mathbf{z}}$ is an estimate for \mathbf{z} with cyclic mean set to zero and including the limiter error. Note that the cyclic mean of the limiter error is also zero.

Next, we have the Iterative data bit estimation block, where we iteratively obtain the data bit estimates. The procedures performed inside this block are described in detail in Section 6. Finally, the bit estimates are collected for bit error rate (BER) and block error rate (BLER) evaluations. The concept of (data) block in our system will be described in more detail in Section 7.

3 Symbol level limiter error modeling

Even though the earlier discussion assumed that the error caused by the symbol level limiter is purely additive, we will adopt another model for the channel estimator modifications. In this Section, we will assume that symbol level amplitude limiter will only affect the data dependent pilot sequence, \mathbf{p}_d , and cyclic pilot sequence, \mathbf{p}_c . We model the effects by a common scaling factor and added noise. We refer to this model as the double-scaling model. We start by rewriting the limited symbol sequence as

$$\tilde{\mathbf{s}} = L(\mathbf{s}) = \mathbf{d} + \beta(\mathbf{p}_d + \mathbf{p}_c) + \mathbf{n}_L. \quad (11)$$

Here the additive noise component caused by the limiter, \mathbf{n}_L , is assumed to be uncorrelated with \mathbf{p}_d and \mathbf{p}_c , and it is assumed to have complex Gaussian distribution. This model is a rough approximation of the phenomena that take place in the symbol level limiter, but based on our experience it provides sufficient accuracy for the channel estimator. The main difficulty in the modeling is to incorporate the effect of the limiter on the random data-dependent pilot sequence. We have tried several models, but they all have similar or worse accuracy than the Gaussian model we are going to present here, so we chose it because of its simplicity.

We can rewrite the purely additive limiter error given in the previous Section as $\mathbf{e}_{\text{limiter}} = \tilde{\mathbf{s}} - \mathbf{s} = (\beta - 1)(\mathbf{p}_d + \mathbf{p}_c) + \mathbf{n}_L$. The cyclic mean of the received sequence can now be rewritten as

$$\begin{aligned} \hat{\mathbf{m}}_y &= \mathbf{J}_{\text{Rx}} \mathbf{y} \\ &= \mathbf{J}_{\text{Rx}} \alpha \sqrt{P_{\text{AVG}}} (D_r + \beta(\mathbf{P}_{d,r} + \mathbf{P}_{c,r}) + \mathbf{N}_{L,r}) \mathbf{h}_{\text{eq}} \\ &\quad + \mathbf{N}_G \mathbf{h}_{\text{channel} + \text{RRC}} + \mathbf{W} \mathbf{h}_{\text{RRC}} \\ &= \alpha \sqrt{P_{\text{AVG}}} (\beta \mathbf{P}_r + (\beta - 1) \hat{\mathbf{M}}_{d,r} + \hat{\mathbf{M}}_{n_{L,r}}) \mathbf{h}_{\text{eq}} \\ &\quad + \hat{\mathbf{M}}_{n_G} \mathbf{h}_{\text{channel} + \text{RRC}} + \hat{\mathbf{M}}_w \mathbf{h}_{\text{RRC}}. \end{aligned} \quad (12)$$

Because we have assumed that the limiter would affect only the pilot sequences, we have to define new methods for approximating these scaling parameters. We approximate β by generating a symbol vector consisting of all possible data symbol and pilot symbol combinations, defined as $\mathbf{s}_{\text{comb},1} = \sqrt{(1 - \gamma)} \mathbf{d}_l + \sqrt{\gamma} \mathbf{p}_l = \mathbf{1}_{N_p \times 1} \otimes \mathbf{d} + \mathbf{p} \otimes \mathbf{1}_{2^Q \times 1}$, where \mathbf{d} is a vector containing all possible symbols, \mathbf{p} is the OCI pilot sequence and Q is the number of bits per symbol. Next, we run this test sequence through the limiter and approximate the scaling factor as

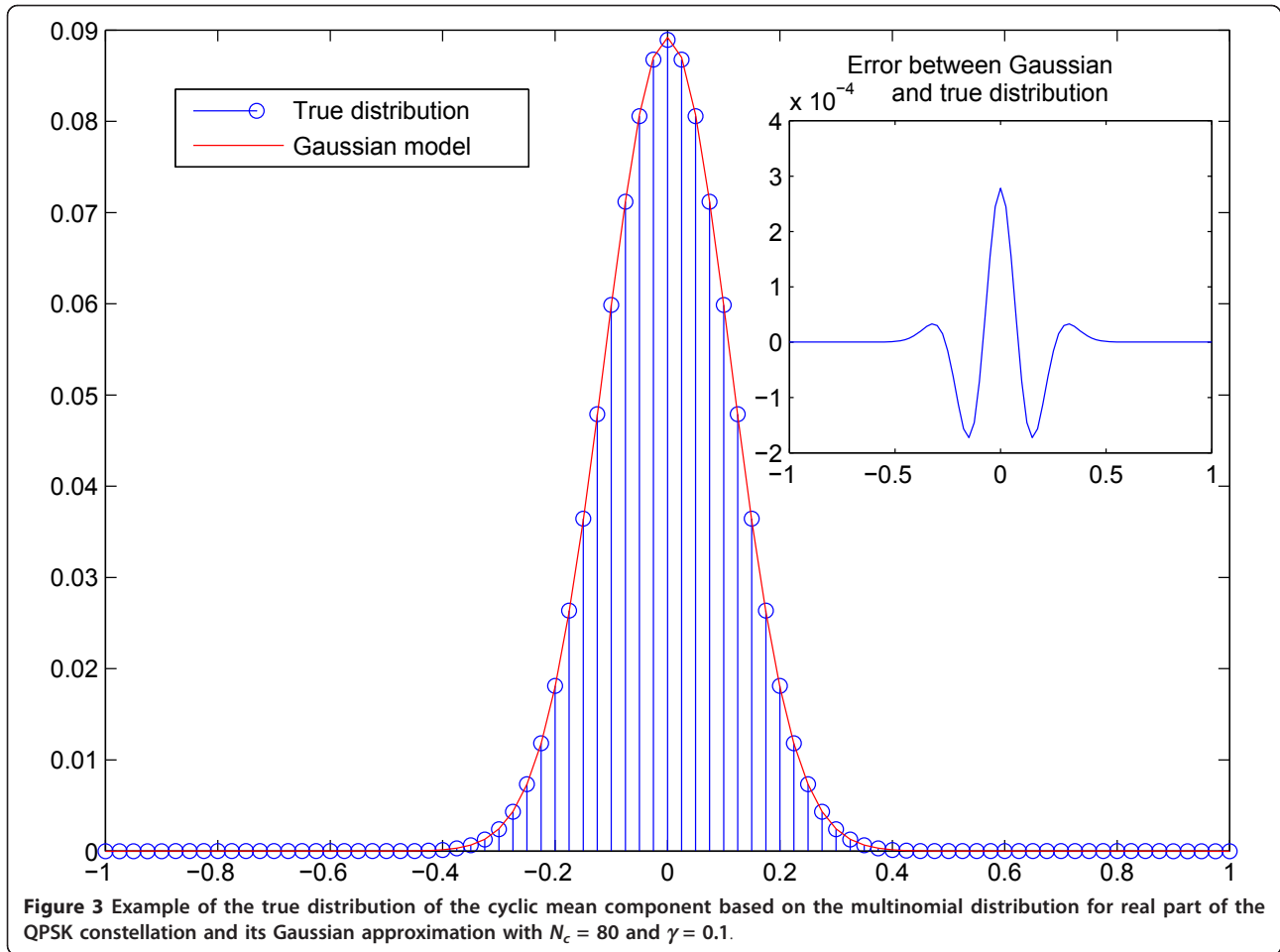
$$\beta = \frac{|\mathbf{p}_l^H L(\mathbf{s}_{\text{comb},1})|}{|\mathbf{p}_l^H \mathbf{p}_l|}, \quad (13)$$

where we basically calculate a correlation based weighting factor for the extended pilot sequence, \mathbf{p}_l . We use this same weighting factor for data dependent pilot sequence because it undergoes similar effects in the symbol level amplitude limiter.

Now the difficult question is, how can we approximate $\sigma_{\mathbf{e}_{\text{limiter}}}^2 = E[|\tilde{\mathbf{s}} - \mathbf{s}|^2]$. First we have to somehow model the distribution of the cyclic mean of the transmitted sequence. The probability of a certain combination of N_c symbols follows the multinomial distribution

$$\begin{aligned} &p(x_1, x_2, \dots, x_k; n, p_1, p_2, \dots, p_k) \\ &= \begin{cases} \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}, & \text{when } \sum_{i=1}^k x_i = n \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (14)$$

where x_i is the number of observations of a certain constellation point on a real or imaginary axis, p_i is the probability of that constellation point and in our case $n = N_c$ is the number of realizations in total per cyclic mean value. Here k is the number of constellation points per real or imaginary axis and takes the value of 2, 4 or 6 for QPSK, 16-QAM and 64-QAM, respectively. In this case, because all symbols are equally probable, $p_i = 1/k$ for all i . To get the true probability of a certain cyclic mean value, one has to add together all the probabilities of different combinations leading to that specific cyclic mean value. With high number of cyclic copies, the distribution of the cyclic mean value tends toward the Gaussian distribution, as expected based on the central limit theorem. For this reason, we have chosen to model the data dependent pilot sequence \mathbf{p}_d with a continuous complex Gaussian distribution $n_{pd} \in \mathcal{N}(0, \sigma_{pd}^2)$, where $\sigma_{pd}^2 = E[|p_d|^2] = \sigma_d^2 / N_c$, is the expected power of the data-dependent pilot sequence. In Figure 3, we have shown the true distribution of the real part of the cyclic mean component of QPSK constellation based on the multinomial distribution (which in this case



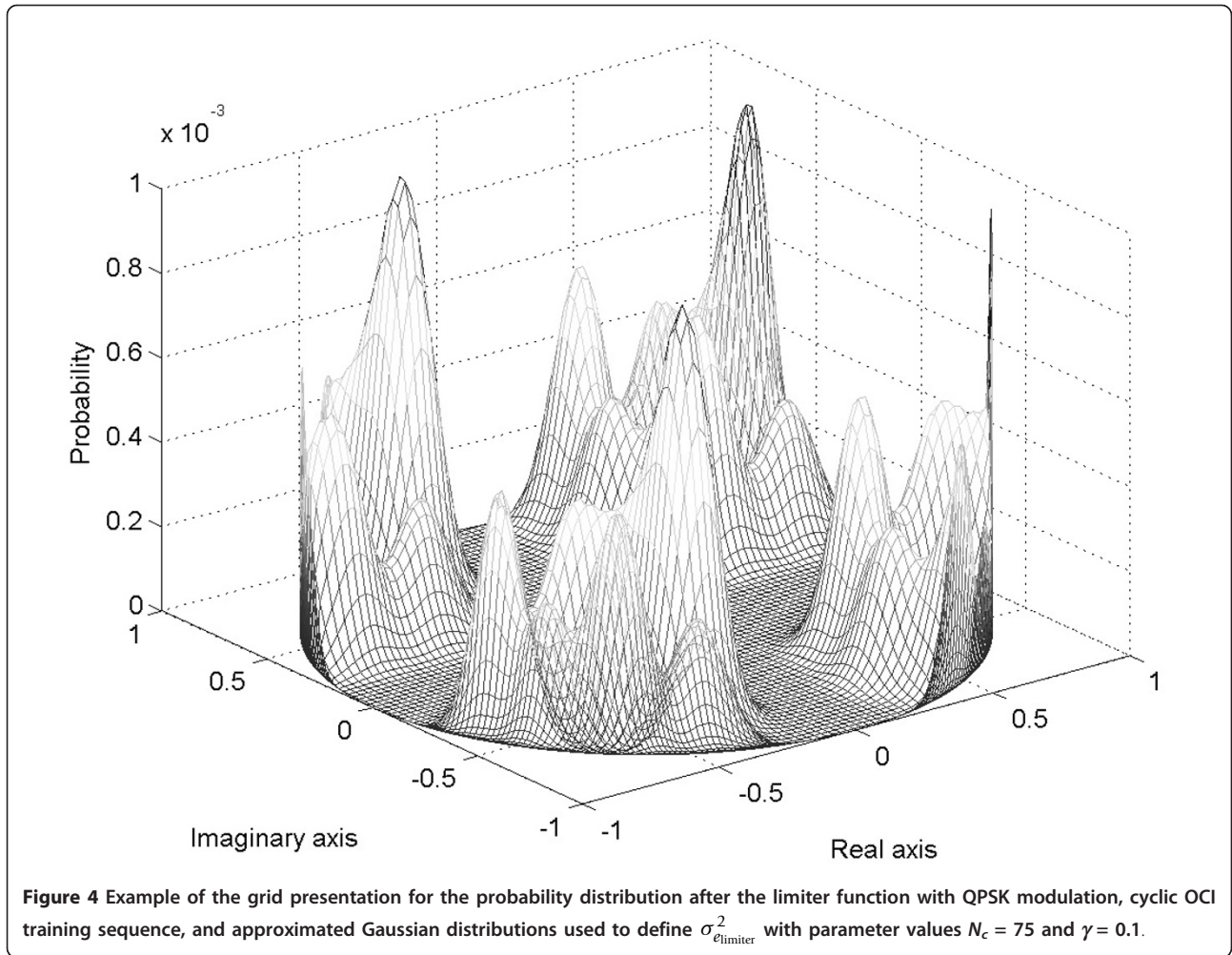
is actually binomial), its Gaussian approximation and the error between these two models. The Gaussian approximation is a good compromise for modeling purposes.

In order to approximate $\sigma_{\text{limiter}}^2$, let us first define another symbol vector consisting of all possible data symbol and pilot symbol combinations, defined as $\mathbf{s}_{\text{comb},2} = \sqrt{(1 - 1/N_c)(1 - \gamma)}\mathbf{d}_l + \sqrt{\gamma}\mathbf{p}_l$, where the power scaling factor $\sqrt{1 - 1/N_c}$ is used to ensure that the total probability over the grid model, after adding Gaussian noise modeling the cyclic mean, equals to unity. Next, we add together probability grids, in which the different grids are based on the Gaussian distribution of n_{pd} centered on a certain point of vector $\mathbf{s}_{\text{comb},2}$. The overall distribution can be given as

$$\begin{aligned}
 &P(\text{probability of symbols } \mathbf{s}_{\text{comb}} \text{ at point } x, y) \\
 &= P(\mathbf{s}_{\text{comb}}, x, y) = \frac{\text{step}^2}{2^{2Q}N_p} \sum_{k=1}^{2^{2Q}N_p} \frac{1}{\sqrt{\pi\sigma_{pd}^2}} \quad (15) \\
 &\exp\{1/\sigma_{pd}^2[(\text{Re}(s_{\text{comb},2}(k)) - x)^2 + (\text{Im}(s_{\text{comb},2}(k)) - y)^2]\},
 \end{aligned}$$

where x and y present the real and imaginary axes, respectively, in a grid with values from -2 to 2. The step size used for real and imaginary axis for calculating the probabilities of cyclic mean values from the Gaussian distribution is determined by the constellation, power normalization, pilot power allocation factor and the number of cycles used in the cyclic mean calculation. For example, if we are using 16-QAM constellation with $\gamma = 0.05$ and have $N_c = 80$ cycles, the step size used is $\text{step} = 2\sqrt{1 - 0.05}/(80\sqrt{10})$, where $\sqrt{10}$ is the power normalization factor to set 16-QAM constellation average power to unity. This step now corresponds to the smallest change in the cyclic mean over possible symbols in real or imaginary axis and directly provides us a model for the discrete distribution of the cyclic mean with the defined parameters.

In Figure 4, we show as an example the generated grid model for QPSK constellation with pilot power allocation factor $\gamma = 0.1$ and number of cyclic means $N_c = 80$ after the limiter function. With QPSK the constellation power normalization factor is one, thus the step size is $\text{step} = 2\sqrt{0.9}/80$.



If we define $g(x, \gamma) = \sqrt{x^2 + \gamma^2}$ as a vectorized function of the distances of grid points (x, γ) from the origo, we can approximate $\sigma_{e_{\text{limiter}}}^2$, given as

$$\sigma_{e_{\text{limiter}}}^2 = \sum_{x, \gamma} |g(x, \gamma) - L(g(x, \gamma))|^2 P(s_{\text{comb}}, x, \gamma). \quad (16)$$

We will use the $\sigma_{e_{\text{limiter}}}^2$ value in the ML-LMMSE channel estimator to incorporate a priori knowledge of the symbol limiter based error term.

If we now assume that \mathbf{p}_c , \mathbf{p}_d , and $\mathbf{n}_{\text{limiter}}$ are uncorrelated, we can obtain the power of the limiter error with double-scaling model to be

$$\begin{aligned} \sigma_{n_L}^2 &= \sigma_{e_{\text{limiter}}}^2 - (\beta - 1)^2 (\sigma_{p_d}^2 - \sigma_p^2) \\ &= \sigma_{e_{\text{limiter}}}^2 - (\beta - 1)^2 (\sigma_d^2 / N_c - \sigma_p^2). \end{aligned} \quad (17)$$

By using the same grid model, we can obtain our estimate of the average power of the limited symbol sequence $\sigma_s^2 = E[|\mathbf{s}|^2]$, as

$$\sigma_s^2 = \sum_{x, \gamma} |L(g(x, \gamma))|^2 P(s_{\text{comb}}, x, \gamma). \quad (18)$$

Here, the average power of the amplitude limited signal and the limiter error power could also be estimated by Bussgang's method [12]. However, based on our simulations, the developed model gives similar estimates and is simpler because it does not require averaging simulations for the framewise correlation calculations. Thus, it provides an alternative approach to define these parameters.

4 Channel estimation with LDDST

In this Section, we will provide the used channel estimator for LDDST. When defining the LMMSE channel estimator, we want to minimize the expected value of the squared error, $E\{|\hat{\mathbf{h}} - \mathbf{h}|^2\}$. If we now make the assumptions that the noise and the total interference experienced by the pilot sequence is AWGN, channel taps are i.i.d. and have zero mean, i.e., $E\{\mathbf{h}\} = \mathbf{0}$, the LMMSE estimator can be simplified to [19]

$$\hat{\mathbf{h}} = \left(\sigma^2 \mathbf{C}_{\hat{\mathbf{h}}_{\text{apriori}}}^{-1} + \mathbf{P}_{c,r}^H \mathbf{P}_{c,r} \right)^{-1} \mathbf{P}_{c,r}^H \mathbf{y}_r \quad (19)$$

where $\sigma^2 = \|\mathbf{h}_{\text{RRC}}\|^2 \sigma_w^2 + E[\|\mathbf{h}_{\text{channel+RRC}}\|^2] \sigma_{n_c}^2 + E[\|\mathbf{h}_{\text{eq}}\|^2] \sigma_{n_l}^2$ models the total interference power based on the Gaussian channel noise, nonlinear power amplifier caused interference and the limiter error. The channel covariance matrix, $\mathbf{C}_{\hat{\mathbf{h}}_{\text{apriori}}}$, contains the a priori information of the channel tap values. The a priori information of the channel taps is obtained through a least squares (LS) channel estimator. From (12), the LS channel estimator can be defined as

$$\begin{aligned} \hat{\mathbf{h}}_{LS} &= \frac{\mathbf{P}_r^H}{\beta r^2 N_p \sigma_p^2} \hat{\mathbf{m}}_y = \left(\alpha \sqrt{P_{\text{AVG}}} - 1 \right) \mathbf{h}_{\text{eq}} \\ &+ \frac{\alpha \sqrt{P_{\text{AVG}}} \mathbf{P}_r^H}{\beta r^2 N_p \sigma_p^2} \left[(1 - \beta) \hat{\mathbf{M}}_{d,r} + \hat{\mathbf{M}}_{n_l} \right] \mathbf{h}_{\text{eq}} \\ &+ \frac{\mathbf{P}_r^H}{\beta r^2 N_p \sigma_p^2} \left(\hat{\mathbf{M}}_{n_c} \mathbf{h}_{\text{channel+RRC}} + \hat{\mathbf{M}}_w \mathbf{h}_{\text{RRC}} \right). \end{aligned} \quad (20)$$

We have assumed independent tap coefficients, which allows us to model the a priori channel correlation matrix $\mathbf{C}_{\hat{\mathbf{h}}_{\text{apriori}}}$ as a diagonal matrix. Because of the receiver pulse shape filtering, this assumption is not exactly true, but it is used to provide us simpler diagonalized LMMSE estimator model, which reduces the channel estimation complexity. We shall refer to this LMMSE estimator, that uses LS based channel estimates as a priori information, as LS-LMMSE channel estimator. The performance of the receiver could be improved with more advanced methods taking the correlation into account, like the universal basis based decomposition of the receiver pulse shape filter correlation, as was discussed in [20]. In a sense, the idea of using only the most significant components of the decomposition is similar to our idea of truncating the time window of the channel estimator to take into account only the most significant channel taps. Both methods gain in noise power reduction in the channel estimation but lose in the asymptotic accuracy.

In the channel estimator, we approximate the diagonal correlation matrix \mathbf{C} by the instantaneous tap power obtained from the LS channel estimator, i.e.,

$$\mathbf{C}_{\hat{\mathbf{h}}_{LS}} = \text{diag} \left\{ \left| \hat{h}_{LS}(0) \right|^2, \left| \hat{h}_{LS}(1) \right|^2, \dots, \left| \hat{h}_{LS}(rN_p - 1) \right|^2 \right\}. \quad (21)$$

By assuming the cyclic OCI training sequence, the LS-LMMSE estimator can be reduced to

$$\hat{\mathbf{h}}_{LS\text{-LMMSE}} = \frac{\mathbf{P}_r^H}{\beta \left(\sigma_{\text{est}}^2 \mathbf{C}_{\hat{\mathbf{h}}_{LS}}^{-1} + r^2 N_p \sigma_p^2 \mathbf{I}_{rN_p \times rN_p} \right)} \hat{\mathbf{m}}_y. \quad (22)$$

The variable σ_{est}^2 corresponds to the total interference power on top of each received pilot symbol and is estimated as

$$\sigma_{\text{est}}^2 = \frac{1}{\beta^2 N_c} \left[\left\| \hat{\mathbf{h}}_{LS} \right\|^2 \sigma_{n_l}^2 + (1 + 1/N_c) \sigma_w^2 \|\mathbf{h}_{\text{RRC}}\|^2 \right], \quad (23)$$

where we do not have a term related to $\sigma_{n_c}^2$ because this value is unknown to the receiver. Similar channel estimator structure with traditional SI pilots and iterative interference canceling feedback was studied in [21].

5 PAPR analysis and spectral leakage comparison

One drawback with DDST in SC transmission is the increased PP and PAPR in the transmitted signal and spectral leakage caused by the non-linear amplifier due to the increased PAPR. These problems are well known but have received relatively little attention in the recent literature.

In a SC transmission, the PAPR of the transmitted sequence is defined after the Tx pulse-shape filter. The PP we see in the filter output depends on the maximum amplitude of the input symbols and on a portion of the absolute values of the filter coefficients, depending on the oversampling. Because we have fixed the Tx pulse-shape filter, only the maximum amplitudes of the input symbols effect the observed PAPR.

There are two main reasons for increased symbol level amplitude in DDST. First of all, we increase the amplitude range related to a certain constellation by adding a power scaled pilot sequence on top of a power scaled symbol sequence. The second main reason for increased amplitude is the possibility of a cyclic mean (data dependent pilot) component with relatively high amplitude. When this component is added on top of data and known pilot symbols, and if the angles of these complex variables happen to align, then the total symbol amplitude is significantly increased.

In this Section, we will first discuss the worst case PP and PAPR effects in more detail and after that we will describe the reference spectral power mask and related simulations and results.

5.1 PAPR analysis and simulated results

For the analysis and results in this section, we have used oversampling ratio equal to four, $r = 4$. The worst case evaluations are based on the filter taps with separation of r samples that have the highest sum-power. This is because the transmitted symbol sequence is oversampled by factor r , so then for each output only every r th filter tap value participates in the corresponding power value. In other words, the filter model used in the following derivations is defined as $h_{\text{RRC}}(\mathbf{i})$, where the set of indices

\mathbf{i} is chosen based on criteria

$$\left\{ \mathbf{i} = [k, k+r, \dots, k+nr] \mid \max_k \left[\left(\sum_{i \in \mathbf{i}} |h_{\text{RRC}, \text{Tx}}(i)| \right)^2 \right] \right\}, \quad (24)$$

where $k \in [0, 1, \dots, r-1]$ and $k+nr \leq N_{\text{RRC}}$. With RRC transmit pulse shape filter of degree 64 and $r=4$, the starting index which maximizes the sum-power is $k=2$. Because the RRC filter acts also as a oversampling filter, the taps of the filter are multiplied by the oversampling factor r in order to keep the average transmitted power equal to unity.

First, we define the worst case symbol level PP. Assume now that $d(k) = ae^{j\phi}$ is some corner symbol with amplitude a and all the other symbols present in the cyclic mean calculation, $d(k+iN_p) = ae^{j(\phi-\pi)}$ with $i=1, 2, \dots, N_c-1$, are opposite corner symbols with amplitude a . Then the data dependent pilot added on top of $d(k)$ is equal to

$$\begin{aligned} p_d(k) &= -\frac{1}{N_c} \sum_{i=0}^{N_c-1} d(k+iN_p) \\ &= -\frac{1}{N_c} [(N_c-1)(ae^{j(\phi-\pi)}) + ae^{j\phi}] \\ &= \frac{(N_c-2)}{N_c} ae^{j\phi} \\ &= \frac{N_c-2}{N_c} \sqrt{1-\gamma} a_{\text{max}} e^{j\phi}, \end{aligned} \quad (25)$$

which corresponds to the worst case peak amplitude with the data dependent pilot sequence and its value depends on the used constellation and the pilot power allocation factor γ . The worst case symbol level PP is defined for an aligned pilot $p_c(k)$ which has amplitude $\sqrt{\gamma}$. By aligned, we mean that the arguments of data and the pilot are equal, $\angle d(k) = \angle p_c(k) = \phi$. Now we can write the worst case symbol level PP as

$$\begin{aligned} \text{WPP}_s &= |d(k) + p_d(k) + p_c(k)|^2 \\ &= \left[\left(1 + \frac{N_c-2}{N_c} \right) \sqrt{1-\gamma} a_{\text{max}} + \sqrt{\gamma} \right]^2. \end{aligned} \quad (26)$$

By using (26), we can define then the worst case PP after the transmit pulse shape filtering to be

$$\begin{aligned} \text{WPP}_{\text{Tx}, \text{DDST}} &= \left(\sum_{i \in \mathbf{i}} |h_{\text{RRC}}(i)| \right)^2 \\ &= \left[\left(1 + \frac{N_c-2}{N_c} \right) \sqrt{1-\gamma} a_{\text{max}} + \sqrt{\gamma} \right]^2, \end{aligned} \quad (27)$$

For TDMT, the worst case PP after the transmit pulse shape filtering is

$$\text{WPP}_{\text{Tx}, \text{TDMT}} = a_{\text{max}}^2 \left(\sum_{i \in \mathbf{i}} |h_{\text{RRC}}(i)| \right)^2. \quad (28)$$

If we use the presented hard symbol level limiter in the transmitter, then the worst case symbol level PP can be given as

$$\text{WPP}_{s, \text{limited}} = |L(d(k) + p_d(k) + p_c(k))|^2 = a_{\text{max}}^2, \quad (29)$$

which is the same as with TDMT. Then the worst case PP after the RRC filtering is

$$\text{WPP}_{\text{Tx}, \text{DDST}, \text{limited}} = a_{\text{max}}^2 \left(\sum_{i \in \mathbf{i}} |h_{\text{RRC}}(i)| \right)^2. \quad (30)$$

which is equal to TDMT case.

With the PPs defined, we can define the PAPRs for different cases. While reading the results for PAPR from Table 1, one should note the difference in the average powers used to define these PAPR results. The average power of a TDMT signal is given as $E[|s_{\text{TDMT}}|^2] = 1$. For DDST based system, the average power of the signal is $E[|s|^2] = (1 - 1/N_c)\sigma_d^2 + \sigma_p^2$. The weighting factor $(1 - 1/N_c)$ is caused by the removal of the cyclic mean from the data sequence. Now the worst case PAPR for DDST without limiter before and after the transmitter pulse shape filter can be given as

$$\begin{aligned} \text{WPAPR}_s &= \frac{\text{WPP}_s}{E[|s|^2]} \\ &= \frac{\left[\left(1 + \frac{N_c-2}{N_c} \right) \sqrt{1-\gamma} a_{\text{max}} + \sqrt{\gamma} \right]^2}{(1 - 1/N_c)\sigma_d^2 + \sigma_p^2}, \end{aligned} \quad (31)$$

and

$$\begin{aligned} \text{WPAPR}_{\text{Tx}, \text{DDST}} &= \frac{\text{WPP}_{\text{Tx}, \text{DDST}}}{E[|s|^2]} \\ &= \frac{\left(\sum_{i \in \mathbf{i}} |h_{\text{RRC}}(i)| \right)^2 \left[\left(1 + \frac{N_c-2}{N_c} \right) \sqrt{1-\gamma} a_{\text{max}} + \sqrt{\gamma} \right]^2}{(1 - 1/N_c)\sigma_d^2 + \sigma_p^2}. \end{aligned} \quad (32)$$

The average power for LDDST is given as $E[|\tilde{s}|^2] = \sigma_{\tilde{s}}^2$ and is defined based on the Gaussian grid model in (18) in Section 3. The PAPRs for the limited case can be written as

$$\text{WPAPR}_{s, \text{limited}} = \frac{\text{WPP}_{s, \text{limited}}}{E[|\tilde{s}|^2]} = \frac{a_{\text{max}}^2}{\sigma_{\tilde{s}}^2}, \quad (33)$$

and

$$\text{WPAPR}_{\text{Tx}, \text{DDST}, \text{limited}} = \frac{a_{\text{max}}^2 \left(\sum_{i \in \mathbf{i}} |h_{\text{RRC}}(i)| \right)^2}{\sigma_{\tilde{s}}^2}. \quad (34)$$

Finally, the PAPR for the TDMT case equals

$$\begin{aligned} \text{WPAPR}_{T_x, \text{TDMT}} &= \frac{\text{WPP}_{T_x, \text{TDMT}}}{E[|s_{\text{TDMT}}|^2]} \\ &= a_{\max}^2 \left(\sum_{i \in \mathcal{I}} |h_{\text{RRC}}(i)| \right)^2. \end{aligned} \quad (35)$$

In Table 1, we have calculated different symbol level and transmitted signal related worst case PPs and PAPRs for different constellations with pilot power allocation factor $\gamma = 0.1$. As we can see, the hard limiter significantly decreases the worst case PPs and PAPRs and the limited worst case PAPRs are close to the TDMT cases, as was desired.

If we assume that with DDST we want to set the PP at the transmit pulse shape filter output to be at a similar level as with TDMT, based on Table 1, a significant back-off is required. With symbol level amplitude limiter we can remove this backoff requirement. As a downside, the amplitude limiter causes additional interference in the transmitted symbols, which might be significant especially with higher order modulations.

In Table 2, the different simulated PPs and PAPRs are given for each constellation. The simulated values were obtained by finding the maximum PAPR over 100,000 random frame realizations. These results provide more insight on the average PAPR performance of the given system with different training methods, and show that the defined analytic worst case PPs and PAPRs are reliable upper bounds.

As expected, the PP and PAPR results with DDST are not as bad as the worst case studies suggested. The main benefit of using symbol level limiter seems to be with QPSK and 16-QAM constellations, where significant reduction in PAPR can be achieved. 64-QAM has quite similar performance with and without symbol level limiter. In Figure 5, an example of the complementary cumulative distribution functions (CCDF) for PP and PAPR distributions with QPSK constellation are shown.

Table 1 WPP and WPAPR for the used constellations with parameter values $N_c = 75$, $N_p = 60$, and $\gamma = 0.1$

	QPSK	16-QAM	64-QAM
WPP _s (26)	4.8	8.0	10
WPP _{s,limited} (29)	1	1.8	2.3
WPP _{T_x,DDST} (27)	25.6	42.7	53.8
WPP _{T_x,LDDST} (30)	5.3	9.6	12.5
WPP _{T_x,TDMT} (28)	5.3	9.6	12.5
WPAPR _{T_x,DDST} (32)	25.9	43.3	54.6
WPAPR _{T_x,LDDST} (34)	5.4	10.2	12.7
WPAPR _{T_x,TDMT} (35)	5.3	9.6	12.5

All values are given in linear scale

Table 2 Simulated PPs and PAPRs for the used constellations with parameter values $N_c = 75$, $N_p = 60$, and $\gamma = 0.1$

	QPSK	16-QAM	64-QAM
PP _s	2.8	3.9	4.6
PP _{s,limited}	1	1.8	2.3
PP _{T_x,DDST}	6.6	8.7	9.3
PP _{T_x,LDDST}	4.7	7.6	8.9
PP _{T_x,TDMT}	5.3	7.7	9.1
PAPR _{T_x,DDST}	6.8	9.0	9.5
PAPR _{T_x,LDDST}	5.9	8.2	9.2
PAPR _{T_x,TDMT}	5.3	7.8	9.2

All values are given in linear scale

Here we can see that the PAPR distributions are similar but the PP distributions are quite different.

5.2 Spectral leakage with SSPA amplifier model

In this section we will study the spectral re-growth with different training methods and with QPSK, 16-QAM, and 64-QAM constellations. The power amplifier model was given in Section 2. We have chosen to use values $\nu = 1$ and $p = 3$ for the simulations. Because we have assumed that the power amplifier is matched to work with TDMT transmission, we have set the 1 dB compression point of the power amplifier based on the 64-QAM constellation PP distribution. The chosen amplitude limit is related to the PP which gives us 1% probability in the CCDF. Thus, from the results obtained in the previous section, we can look for the PP with 64-QAM that $P(\text{PP}_{64\text{-QAM}} \leq P_{1\text{dB}}) = 0.01$. Based on our simulations, this value is equal to $P_{1\text{dB}} = 4.8$ dB. Now, we use this power value to solve the power amplifier saturation amplitude. The amplitude corresponding to the 1 dB compression point is $A = 10^{4.8/20}$ and the saturation amplitude can be solved to be

$$A_0 = \nu A \left(10^{p/10} - 1 \right)^{\frac{-10}{2p}}, \quad (36)$$

which gives us $A_0 \approx 1.739$.

The used spectral mask is based on 3GPP technical specification for E-UTRA user equipment [22]. The used required attenuation levels are based on 23 dBm transmission power in the used 20 MHz bandwidth and Table 6.6.2.2.2-1 in page 44 of [22]. We chose the values of this Table because it provides the most strict attenuation mask. The obtained attenuation levels are given in Table 3 with respect to the distance from the channel band edge. This distance is defined as an out-of-band frequency distance, Δf_{OOB} . The required attenuation levels are defined for a measurement bandwidth of 1 MHz.

For the simulations, we have assumed to use 20 MHz channel bandwidth, 18 MHz symbol frequency and a roll-off factor 0.1 in the RRC filter. We wanted to keep

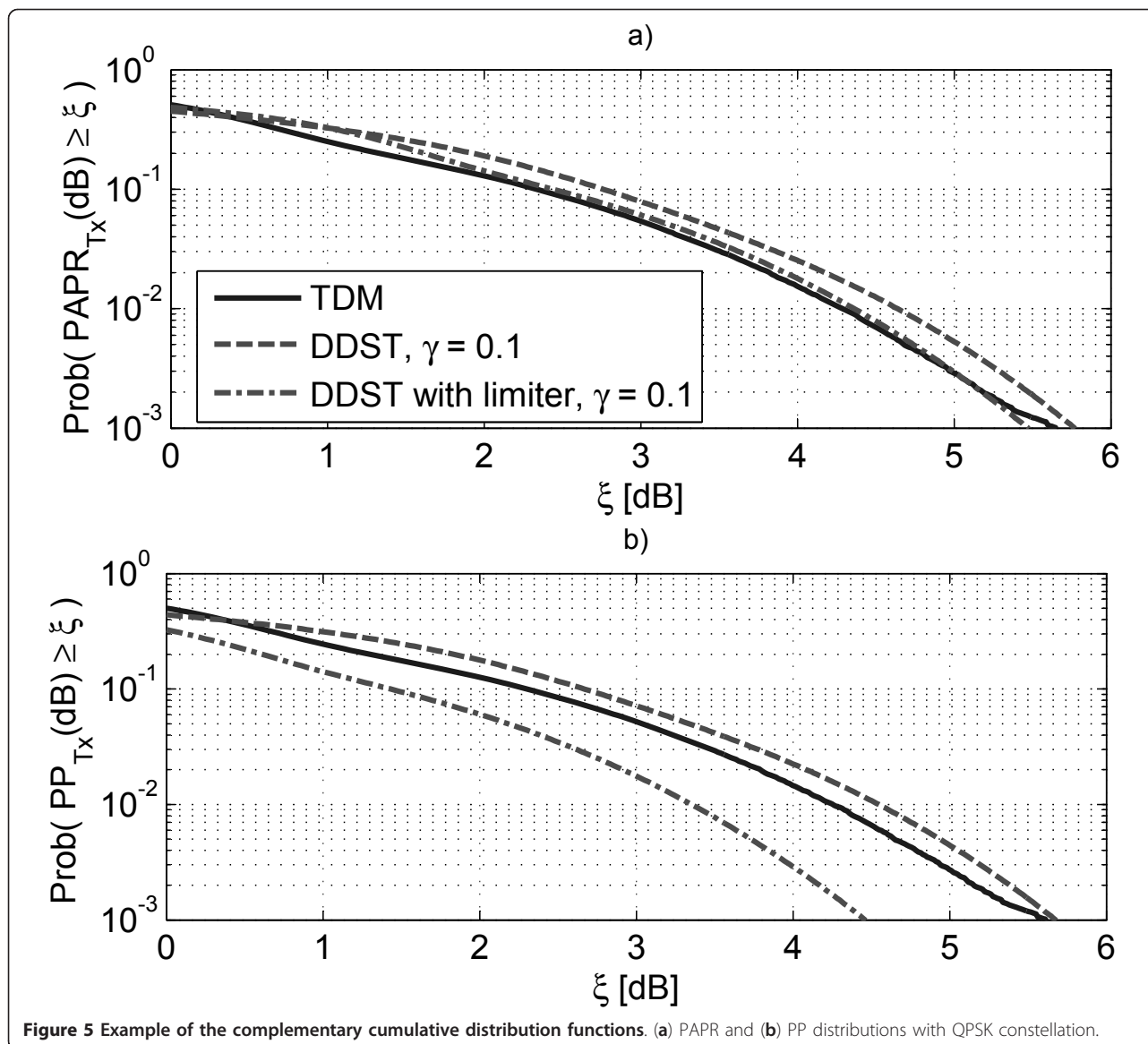


Figure 5 Example of the complementary cumulative distribution functions. (a) PAPR and (b) PP distributions with QPSK constellation.

the roll-off factor small because we are aiming toward very high spectral efficiency. For different training methods and constellations, we ran the simulations looking for smallest IBO with 0.5 dB step in the average transmitted power, P_{AVG} . We have defined the input backoff (IBO) as $IBO = 10\log_{10}(A_0^2/P_{AVG})$. Based on the results, we chose the smallest IBO for each training method and constellation which leads to spectral leakage that stays below the given spectral mask. The obtained IBO and output backoff (OBO) results are provided in the Table 4. The OBO is defined as the maximum output power to the average output power ratio, given as $OBO = 10\log_{10}(A_0^2/E[G(\mathbf{x})^2])$.

As expected, based on the PP and PAPR analysis, we can reach significantly lower OBO when using limited

DDST with QPSK constellation. With 16-QAM constellation we can decrease the OBO somewhat with symbol level limiter. With 64-QAM, meaningful gains were not achieved with symbol level amplitude limiter. These IBO values are used in Section 7 when we compare the throughput performance of different training methods.

Next, we will return to the actual implementation of the iterative receiver used with limited DDST before we study the throughput performance with different training methods.

6 Iterative receiver algorithms

The receiver operations before the iterative data bit estimation were already described in Section 2. In this section we discuss in more detail the operations performed

Table 3 Attenuation at distance Δf_{OoB} from the channel band edge

Δf_{OoB} [MHz]	Attenuation requirement [dB]
$\pm 0-1$	-15.76
$\pm 1-5.5$	-22.99
$\pm 5.5-25$	-34.99

inside the iterative data bit estimation block, shown in more detail in Figure 6.

We have used notation $\hat{\mathbf{z}}$ to represent our estimates of the data symbol sequence, including the limiter error, with cyclic mean set to zero, obtained from the pilot removal and information symbol power normalization block, as shown in Figure 2. We use $\hat{\mathbf{z}}$ as a initial data symbol estimates to generate hard symbol based cyclic mean estimate in the hard symbol based \mathbf{p}_d estimation and compensation block. Inside this block, we generate hard symbol estimates based on $\hat{\mathbf{z}}$, calculate their cyclic mean and add it to $\hat{\mathbf{z}}$, to obtain initial symbol estimates $\hat{\mathbf{d}}^0$. Here superscript 0 points out that these symbol estimates are obtained before coded feedback. This idea was presented in [4], and we use it before the first soft symbols to bits mapping.

We start the iterative reception process by using $\hat{\mathbf{d}}^0$ to generate soft coded bit estimates $\hat{\mathbf{b}}$ in the *soft symbols-to-bits* block. These are then provided to the soft-input soft-output (SISO) decoder from which we obtain our first soft decoded bit estimates to be provided for the \mathbf{p}_d and $\mathbf{e}_{\text{limiter}}$ estimation and compensation block and for bit error evaluation. This block is presented in more detail in Figure 7, where superscript i refers to the iteration number. These procedures, before we obtain the first feedback data symbol estimates, $\hat{\mathbf{d}}^1$, are considered to happen in the zeroth feedback iteration ($i = 0$). In our notation, after first pass through channel decoder, symbol estimation and compensation processes, we obtain our first feedback data symbol estimates $\hat{\mathbf{d}}^1$, to be used for soft bit estimation.

Table 4 Simulation based IBO and OBO results for different training methods and constellations

Training method/constellation	QPSK	16-QAM	164-QAM
Required IBO [dB]			
TDMT	5.3	5.8	5.8
DDST	5.3	5.8	5.8
DDST with limiter	3.8	5.3	5.8
Corresponding OBO [dB]			
TDMT	5.5	6.0	6.0
DDST	5.6	6.1	6.1
DDST with limiter	5.0	5.8	6.1

All values are given in decibels [dB]

The operations inside the \mathbf{p}_d and $\mathbf{e}_{\text{limiter}}$ estimation and compensation block, shown in Figure 7, are performed as follows. First we generate soft symbol estimates based on the latest soft bit estimates $\hat{\mathbf{b}}^i$, which are equal to the log-likelihood presentation of the a posteriori probabilities obtained from the soft decoder. The soft symbols are given by equation

$$\hat{d}_v^i = \sum_{a=1}^{|A|} d_a p(d_a | \hat{\mathbf{b}}_v^i), \quad 0 \leq v \leq N-1, \quad (37)$$

where $|A|$ gives the number of symbols in alphabet A , v is a symbol index, $\hat{\mathbf{b}}_v^i$ are the soft bit estimates related to the v th symbol, and $p(d_a | \hat{\mathbf{b}}_v^i)$ is the probability of a symbol d_a , given the latest soft bit estimates $\hat{\mathbf{b}}_v^i$. The probability of a symbol d_a is defined as

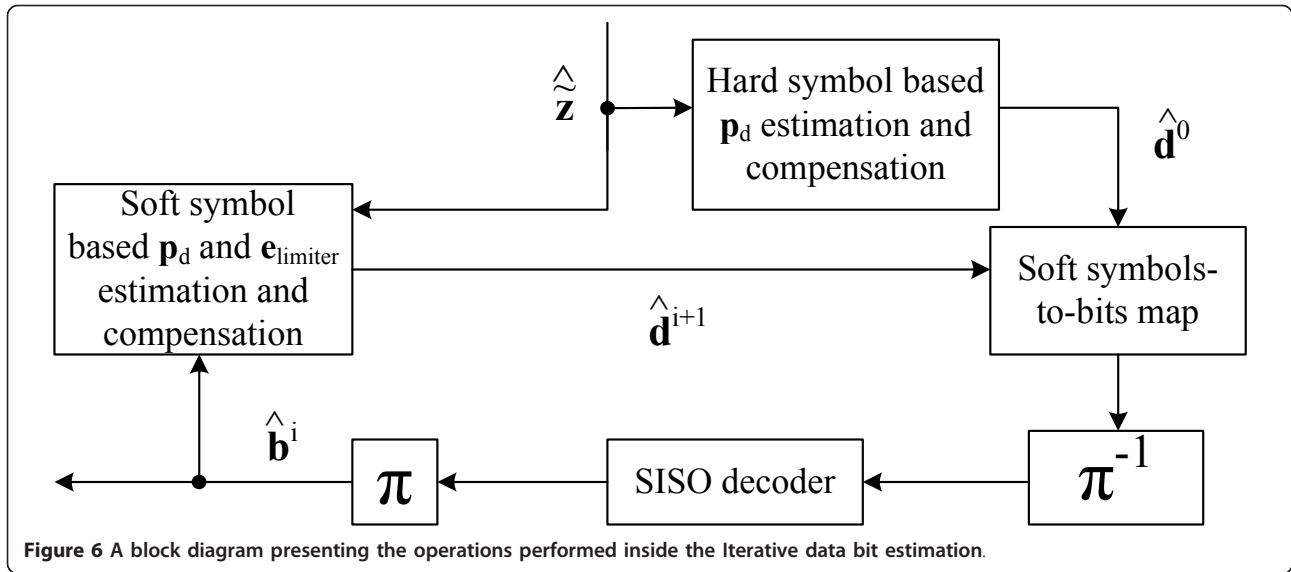
$$p(d_a | \hat{\mathbf{b}}_v^i) = 2^{-Q} \prod_{q=1}^Q \left[1 + \bar{b}_{d_a}(q) \tanh\left(\frac{\hat{b}_v^i(q)}{2}\right) \right], \quad (38)$$

where Q is the number of bits per symbol, $\bar{b}_{d_a}(q) \in [-1, +1]$ is the q th bit of the hypothesis d_a , and $\hat{b}_v^i(q)$ is the log-likelihood presentation of the a posteriori probability related to the q th bit of the v th symbol in the i th iteration, given as

$$\hat{b}_v^i(q) = \log \left(\frac{P_{\text{app}}(b_v^i(q) = 1)}{P_{\text{app}}(b_v^i(q) = 0)} \right). \quad (39)$$

We have also normalized the variance of the soft symbol vector, $\hat{\mathbf{d}}^i$, to be equal to unity. This improves the feedback performance when the soft bit estimates have very low reliability. In our simulations, using soft symbol feedback for the limiter error estimation provided better results than using hard symbol feedback.

Then, we calculate the symbol wise cyclic mean and remove it from the symbol sequence to obtain $\hat{\mathbf{z}}^i$. Now $-\hat{\mathbf{p}}_d^i$ is an improved estimate of the cyclic mean, assuming that the SISO decoder has been able to reduce the number of bit errors in the detected bit sequence. Next, we add the known pilot sequence on top of the sequence $\hat{\mathbf{z}}^i$ to get $\hat{\mathbf{s}}^i$ and provide this sequence to the amplitude limiter. Then we calculate the limiter error estimate based on the input and the output of the limiter function and an improved estimate of the average power, $\sigma_{\hat{\mathbf{s}}^i}^2$. At this point, when $i > 0$, we obtain our first estimate of the limiter error. Based on our results, it is better to estimate the limiter error after the channel



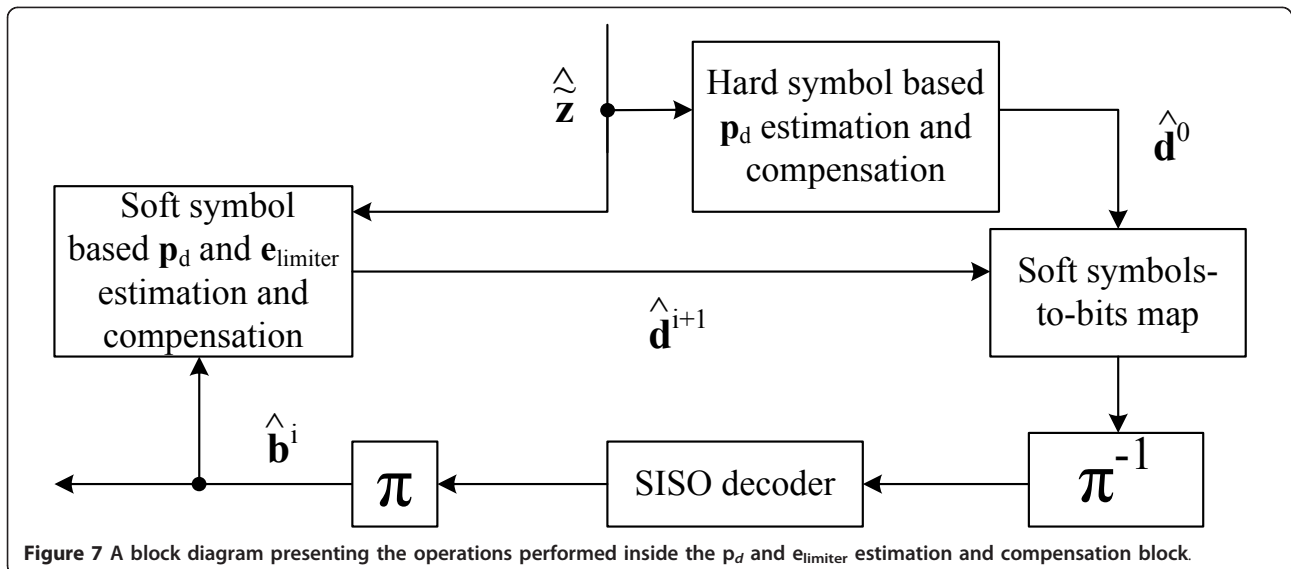
decoder and not based on the uncoded hard symbol estimates $\hat{\mathbf{d}}^0$. With low code rates (low E_b/N_0 region) the uncoded limiter error estimation leads to worse performance in all iterations. Then again, with high code rates (high E_b/N_0 region) uncoded limiter error estimation improves the BLER performance at the 0th iteration, but the iterative gain decreases, leading to worse performance at the fifth iteration.

Based on this improved average amplitude estimate, we can obtain improved symbol estimates by rescaling the average power of the received sequence, remembering that we have already scaled the incoming sequence by $\sigma_{\hat{s}}$ in (10). Finally, we can generate new symbol

estimates by adding to the received symbol estimates $\hat{\mathbf{z}}$ the latest cyclic mean and limiter error estimates, given as

$$\begin{aligned} \hat{\mathbf{d}}^{i+1} &= \frac{\sigma_{\hat{s}}}{\sigma_{\hat{s}}} \hat{\mathbf{z}} - \hat{\mathbf{e}}_{\text{limiter}}^i - \hat{\mathbf{p}}_d^i \\ &= \frac{\sigma_{\hat{s}}}{\sigma_{\hat{s}}} \hat{\mathbf{z}} - (\mathbf{I} - \mathbf{J}_{Tx}) \hat{\mathbf{e}}_{\text{limiter}}^i + \mathbf{J}_{Tx} \hat{\mathbf{d}}^i. \end{aligned} \quad (40)$$

We remove the cyclic mean of the estimated limiter error $\hat{\mathbf{e}}_{\text{limiter}}^i$, because we have completely removed the cyclic mean from $\hat{\mathbf{z}}$, including the limiter error.



Based on our results, it is better not to use the extrinsic information obtained from the channel decoder as a priori information in the soft symbols-to-bits mapping, if this information is already used to improve the cyclic mean estimate. This is probably because we are using the same information twice inside the same loop, thus losing the independence of the a priori information. We can use it as a priori information if we do not improve the cyclic mean, but based on our studies this does not provide as good iterative gain in the receiver. This could be because of the error averaging nature of the cyclic mean computation.

Here we remind the reader, that even without symbol level amplitude limiter, we have to use iterative detection algorithm for the cyclic mean estimation. Of course, the limiter error estimation is not required. Therefore, in the simulation results presented in Section 7, the throughput results obtained with DDST also include five feedback iterations.

For a reader interested in a pure SI training with iterative reception, a good starting point is, for example, [23]. In this article a computationally efficient, iterative frequency-domain equalization and channel estimation is presented. In this article, we have not considered of including the channel estimation process in the iterative loop because with DDST there is no interference from the data symbols to the known pilot symbols. Nonetheless, when there is symbol level limiter involved, we could feedback the cyclic mean of the limiter error estimate in order to improve the channel estimates with LDDST. In addition, in SISO case or in spatially multiplexed MIMO case, the feedback filtering used also in [23], is of great interest and provides interesting topics for future research.

7 Performance comparisons

In this section, we will first provide some results demonstrating the performance of our iterative receiver algorithm. In the end, spectral efficiency comparisons between TDMT and DDST based training are provided. This is, after all, the most important topic of this article. We will investigate whether the end user spectral efficiency is really improved with DDST and do we gain something by using a symbol level amplitude limiter.

The used channel model is a block-fading extended ITU-R Vehicular A channel with approximately 20 MHz bandwidth [13]. The maximum delay spread of the channel is 78 samples. In [13], the channel model was defined for sampling interval $t_s = 32.55$ ns where as in our system the sampling interval is $t_s = 27.78$ ns. This modification has a minor effect on the spectral correlation properties of the channel. However, the main idea is only to do some initial comparisons in the possible throughput performance between DDST and TDMT training based

systems. Therefore, the used model provides a good starting point for the simulations.

The oversampling in the receiver allows us to efficiently realize the RRC filtering in frequency domain in combination with the channel equalization process. More details can be found in [14] and references therein. In this article we have considered single-input single-output (SISO), and 1×2 and 1×4 single-input multiple-output (SIMO) antenna configurations with MRC equalizer.

In our simulations, the channel estimator length is $rN_p = 120$ while the true equivalent channel length, including the effects of transmitter and receiver RRC filters, is $N_{\text{channel}} + 1 + 2N_{\text{RRC}} = 206$ samples. This kind of short channel estimator was studied in [21,24]. The reason behind using short channel estimator is to maximize the number of cycles, N_c , with the cost of minimizing the estimator length, N_p . Because we are estimating the equivalent channel, we can ignore channel tap values close to zero, which are caused by the heavy tailing of the RRC filters. In the presented simulations we have used values $N_c = 75$ and $N_p = 60$ with DDST and LDDST. This gives us a good compromise with the estimator accuracy and achievable number of cyclic copies. Especially with QPSK modulation, when we are working in a high noise environment, it is worth to consider sacrificing the channel estimation accuracy to achieve better noise power averaging through increased number of cyclic copies. With higher order constellations, in addition to the improved noise averaging, with increased number of copies we can also decrease the variance of the data dependent training sequence, \mathbf{p}_d , and this improves the accuracy of the first symbol estimates.

The channel codec uses turbo code [25] with generator matrix $G = [1 \frac{15}{13}]$. We have used the max-log-MAP algorithm presented in [26] without any correction factor for the max-operator. The extrinsic information exchanged between the component decoders is weighted by a factor 0.75 to reduce the error propagation, as proposed in [27]. Iterations in the turbo decoder are terminated based on the hard-data-aided algorithm presented in [28]. The used interleavers are bitwise S -interleavers [29], where the distance parameter is defined as $S = \sqrt{U/2}$ where U is the length of the unit which is interleaved. In channel interleaving the unit is the whole transmitted frame $U = QN$, where Q is the number of bits per symbol and N is the number of symbols per transmitted frame. We divide each transmitted frame into Q coded blocks. Inside the turbo codec the length of the interleaved unit is equal to one uncoded data block $U = \lfloor R(N - 2m) \rfloor$, where $m = 3$ is the memory length of the component encoder and the term $2m$ is caused by the unpunctured termination bits [30].

We have run the simulations for QPSK, 16-QAM, and 64-QAM constellations with code rates $R = 0.5$, $R = 0.67$ and $R = 0.75$. With TDMT pilots, the number of transmitted data symbols in each frame is decreased by the number of pilot symbols, which is set to be 450 in our simulations (10% of the frame duration). The TDMT pilots are the first 450 binary symbols from a Gold code of length 512 symbols [31] with unity power. The channel estimator length is equal to the equivalent channel length. With DDST, we decided to provide same portion of total power for the pilots, thus $\gamma = 0.1$. This gives us a fair comparison between TDMT training and DDST based transmission, because the channel estimation MSE of basic least-squares channel estimator with DDST is the same as with TDMT, if equal amount of power is allocated for the pilots [4]. The optimization of the pilot powers with TDMT or DDST for channel estimation with transmitted average power and PP restrictions is an interesting and open problem, but is out of the scope of this article. Some additional simulation parameters related to the simulation model are given in Table 5.

In all the simulated cases we have used the maximum of five feedback iterations for $\hat{\mathbf{p}}_d$ and $\hat{\mathbf{e}}_{\text{limiter}}$ estimation. Typically, for QPSK modulation two and for 16-QAM modulation three feedback iterations already provide relatively good performance. With 64-QAM modulation we need five feedback iterations to ensure convergence in all of the cases. Example of the typical BLER behavior over iterations with LDDST using amplitude limiter with different constellations, compared to TDMT, is shown in Figure 8. We have assumed that the receiver does not know the IBO used in the transmitter and this degrades the performance results in all of the simulated cases.

One rather intriguing problem while planning the spectral efficiency comparison was the choice of the reference power. The comparison of performance with DDST and TDMT based systems is not so trivial and one has to be

careful about what to compare and how these results should be interpreted.

In the simulations, we chose to do the performance comparisons with respect to the energy per transmitted data bit over one sided noise spectral density, E_b/N_0 . We have chosen this parameter because what matters most in modern wireless communications is the used energy per data bit to transmit with certain spectral efficiency. We have defined the SNR based on E_b/N_0 as

$$\text{SNR} = \frac{E_b Q R_{\text{true}}}{N_0 r}, \quad (41)$$

where Q is the number of bits per symbol, R_{true} is the true coding rate (including the effect of possible termination bits, block length modifications with zero padding, etc.), and $r = 2$ is the oversampling rate used in the receiver.

Figures 9 and 10 present spectral efficiency results for DDST, LDDST and for TDMT training, using also a LS-LMMSE type equalizer, with QPSK modulation and with 16-QAM and 64-QAM modulations, respectively. From Figure 9 we can observe how the increased average transmit power allowed by the symbol level amplitude limiter improves the spectral efficiency in the low E_b/N_0 range with QPSK modulation. In Figure 10 we have shown the performance with higher order modulations. Here, the performance of LDDST compared to DDST is quite similar. Clearly, both DDST based systems improves the spectral efficiency over the whole E_b/N_0 range for each antenna configuration. The maximum spectral efficiency difference for each constellation is equal to 10%, which corresponds to the pilot overhead of TDMT.

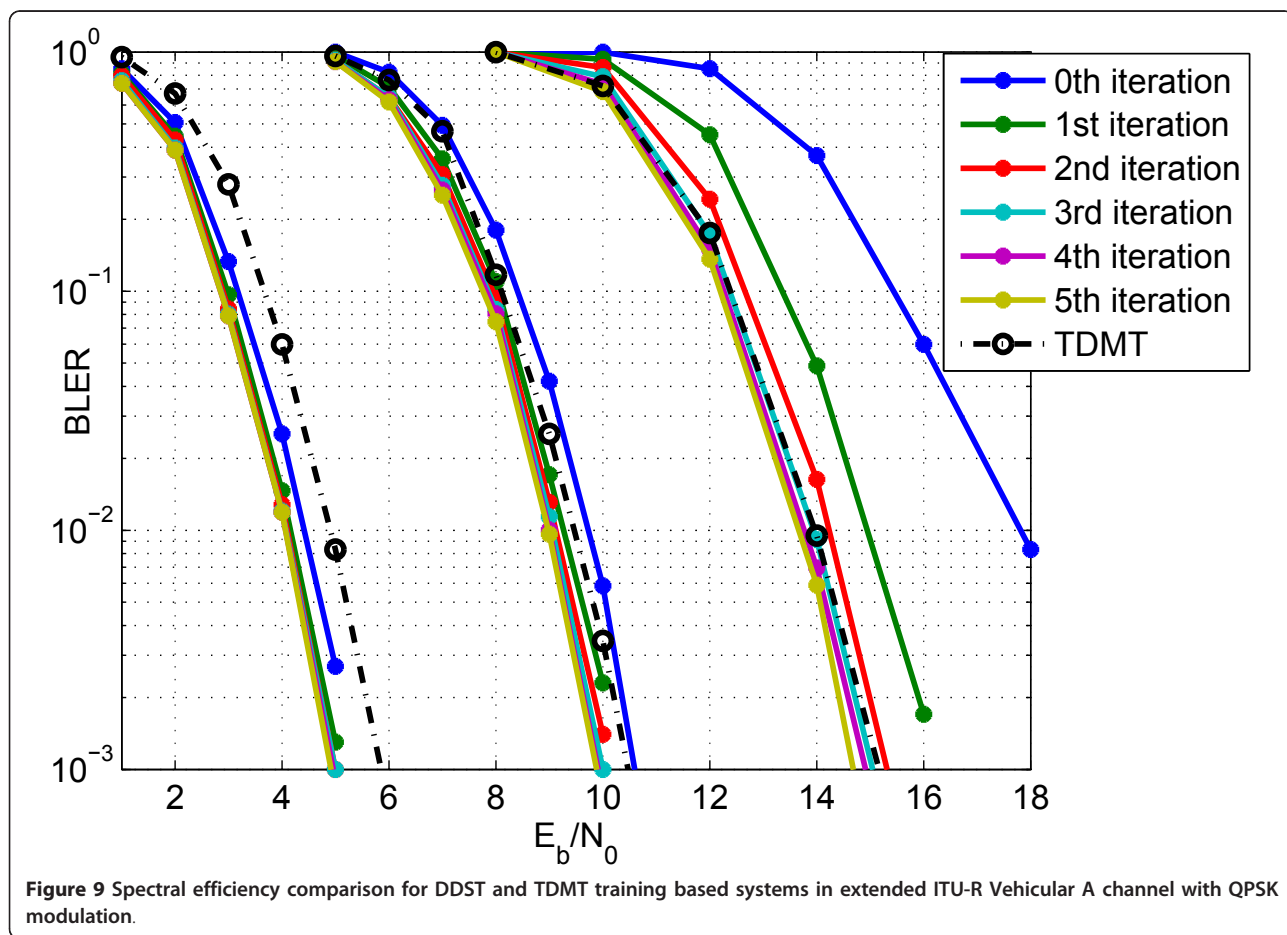
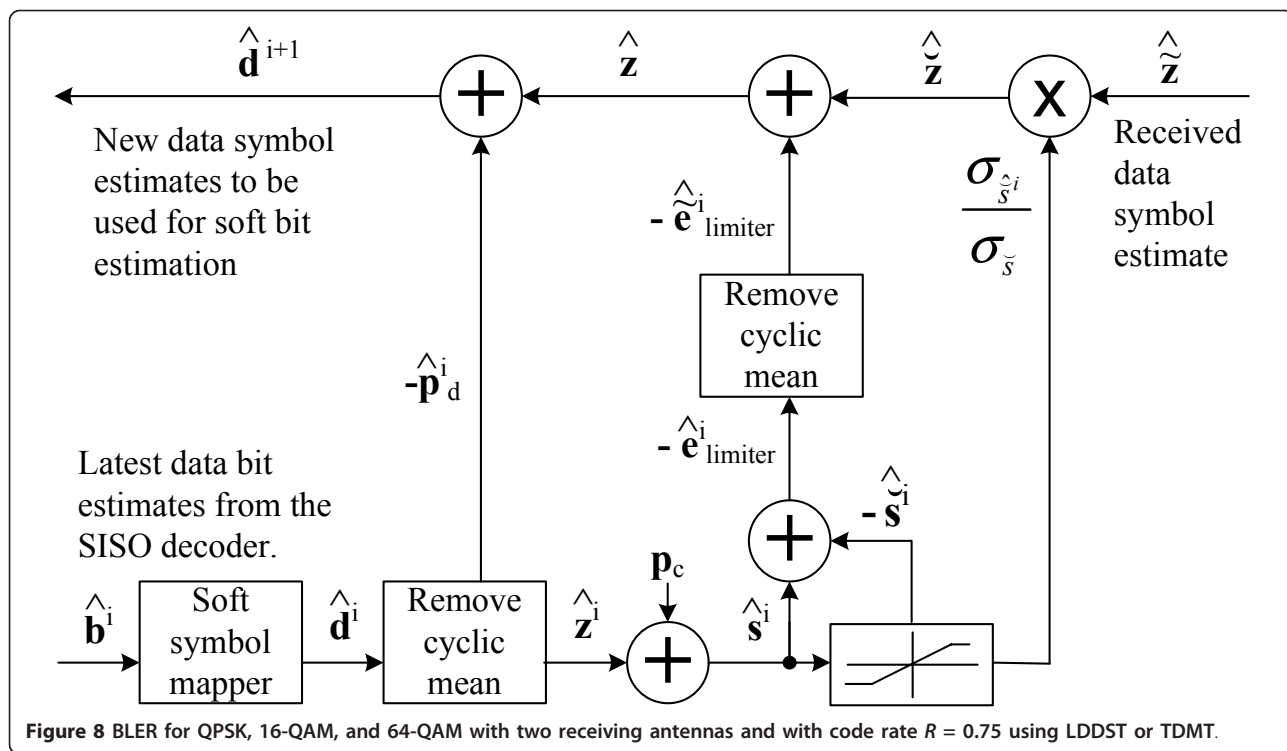
With the proposed symbol level amplitude limiter we can obtain improved spectral efficiency performance with QPSK modulation in all antenna configurations. With 16-QAM or 64-QAM modulations, LDDST and DDST have quite the same performance. Possibly, one could improve the LDDST performance with higher order modulations by tighter limiting bounds. In addition, by first performing tighter limiting and after that removing the cyclic mean, we could decrease the limiter error effect in the channel estimation and possibly improve the system performance. These topics are left for future studies.

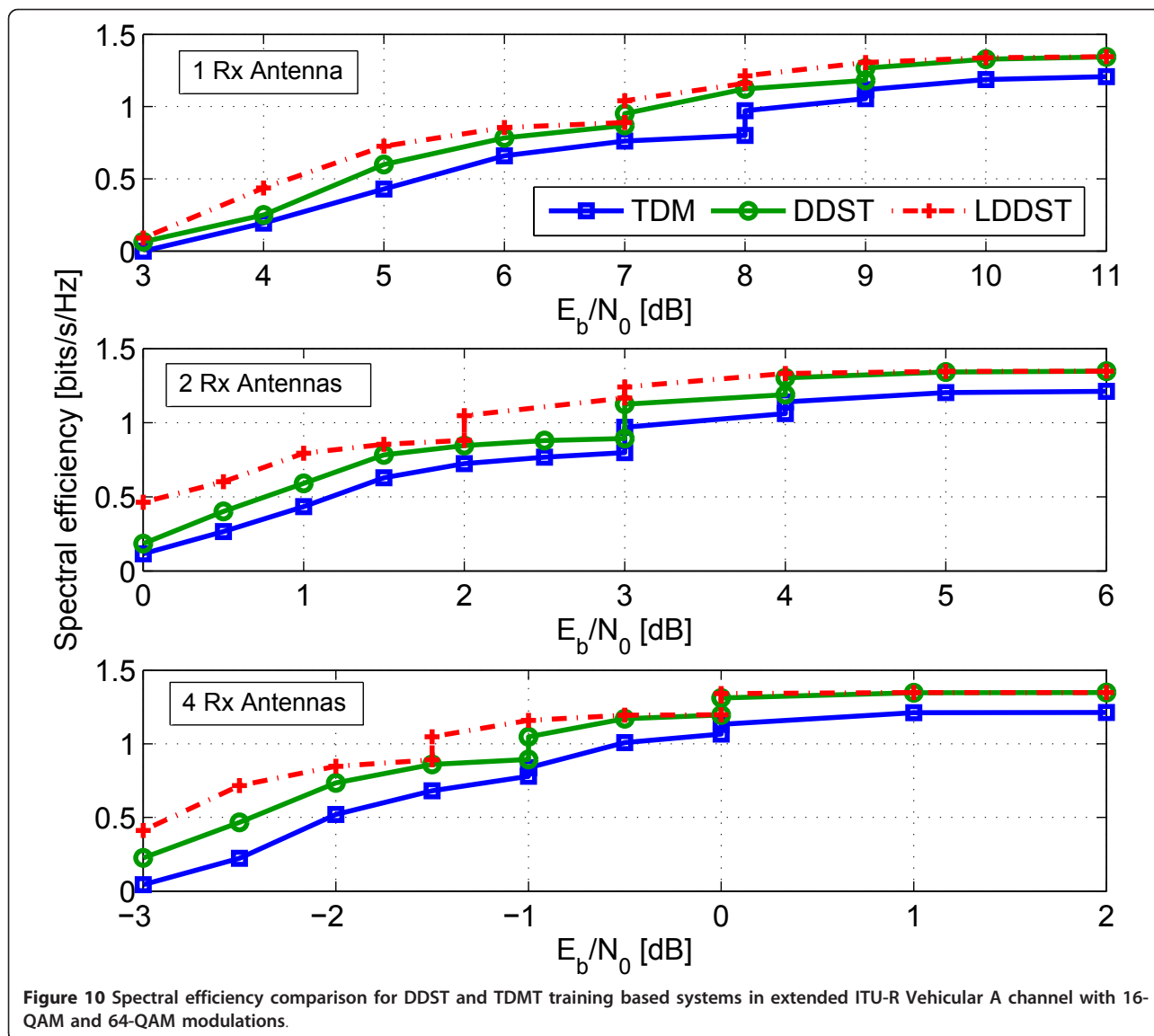
8 Conclusion

In this article, we have discussed the effects of a DDST based training on the signal PP and PAPR distributions. We demonstrated that the PP and PAPR distributions of the DDST based training have longer tails and therefore there is a higher probability for big PAPR values. Especially, with constant amplitude modulations like QPSK, the average PAPR is significantly increased. Furthermore, the effects of the increased PAPR on the spectral leakage

Table 5 Simulation parameters

Symbol rate	18 MHz
Signal bandwidth	19.8 MHz
Frame duration	250 μ s
Order of the RRC filter	64
RRC roll-off	0.1
Symbols per frame	4,500
TDMT pilot symbols per frame	450
Number of feedback iterations	5
No. of subbands in the analysis bank	1,024
No. of subbands in the synthesis bank	512
FB Overlapping factor	5
FB roll-off	1





with SSPA amplifier model were studied. It was shown, that DDST does not require higher IBO compared to TDMT, but does provide slightly worse OBO performance. The proposed symbol level limiter can decrease further the IBO and OBO requirements with QPSK and 16-QAM constellations. The reduced OBO and IBO may significantly ease the design, implementation and cost of the required power amplifier. With QPSK modulation the symbol level limiter also clearly decreases the spectral re-growth and improves the spectral efficiency performance via higher average transmitted power.

Based on our results, with QPSK and 16-QAM, one should consider using LDDST to allow higher average transmitted power (lower OBO) and to achieve improved throughput compared to DDST. With higher order constellations symbol level amplitude limiter, as

presented in this article, doesn't seem to provide significant benefit.

With DDST, with or without symbol level amplitude limiter, the complexity increase compared to traditional TDMT training can be approximated by the complexity of the SISO decoder used. In the soft feedback loop with DDST, with or without symbol level amplitude limiter, the SISO decoder is dominating the detection complexity. Thus, the average increase in the detection complexity compared to TDMT, is roughly the average number of feedback iterations times the number of blocks decoded in average in each feedback iteration times the average complexity of decoding one block in the SISO decoder. With TDMT no feedback iterations are required.

The performance comparisons between DDST and TDMT based system showed that DDST can provide

similar or better performance over the whole E_b/N_0 range with all antenna configurations. The proposed symbol level amplitude limiter improves the throughput performance of the DDST in the low E_b/N_0 range with all antenna configurations tested.

In addition to careful performance analysis and comparisons, we have provided some new ideas for PAPR control with DDST, for modeling the effects of symbol level limiter in channel estimation, and for modeling the cyclic mean distribution based on multinomial distribution or its Gaussian approximation.

Acknowledgements

The authors would like to thank Dr. Ali Shahed Hagh Ghadam for enlightening the mysteries of power amplifiers. This work was supported by the Tampere Graduate School in Information Science and Engineering (TISE), the Nokia Foundation and the Academy of Finland (under Project No. 129077, "Hybrid Analog-Digital Signal Processing for Communications Transceivers").

Competing interests

The authors declare that they have no competing interests.

Received: 11 May 2011 Accepted: 16 February 2012

Published: 16 February 2012

References

1. P Hoehner, F Tufvesson, Channel estimation with superimposed pilot sequence, in *Proc IEEE Global Telecommunications Conference 1999 GLOBECOM '99*, vol. 4. (Janeiro, Brazil, 1999), pp. 2162–2166
2. AG Orozco-Lugo, MM Lara, DC McLernon, Channel estimation using implicit training. *IEEE Trans Signal Process.* **52**(1), 240–254 (2004). doi:10.1109/TSP.2003.819993
3. SAK Jagannatham, BD Rao, Superimposed pilot vs. conventional pilots for channel estimation, in *Fortieth Asilomar Conference on Signals Systems and Computers 2006 ACSSC '06*, (Pacific Grove, California USA, 2006), pp. 767–771
4. M Ghogho, DC McLernon, E Alameda-Hernandez, A Swami, Channel estimation and symbol detection for block transmission using data-dependent superimposed training. *IEEE Signal Process Lett.* **12**(3), 226–229 (2005)
5. DC McLernon, E Alameda-Hernandez, AG Orozco-Lugo, MM Lara, Performance of data-dependent superimposed training without cyclic prefix. *Electron Lett.* **42**(10), 604–606 (2006). doi:10.1049/el:20060127
6. E Gayosso-Rios, MM Lara, AG Orozco-Lugo, DC McLernon, Symbol-blanking superimposed training for orthogonal frequency division multiplexing systems, in *7th International Symposium on Wireless Communications Systems (ISWCS)*, York, UK, pp. (19–22 Sept 2006), 204–208
7. C-T Lam, DD Falconer, F Danilo-Lemoine, R Dinis, Channel estimation for SC-FDE systems using frequency domain multiplexed pilots, in *IEEE 64th Vehicular Technology Conf 2006 VTC-2006 Fall*, (Montreal, Canada, 2006), pp. 1–5
8. T Levanen, J Talvitie, M Renfors, Performance evaluation of a DDST based SIMO SC system with PAPR reduction, in *6th International Symposium on Turbo Codes & Iterative Information Processing, ISTC 2010*, (Brest, France, 2010), pp. 186–190
9. DC McLernon, E Alameda-Hernandez, A Orozco-Lugo, MM Lara, New results for channel estimation via superimposed training, in *Proc Second International Symposium on Communications Control and Signal Processing ISCCSP 2006*, (Marrakech, Morocco, 2006). (Article ID cr1001). ISBN: 2-908849-17-8
10. R Raich, H Qian, GT Zhou, Optimization of SNDR for amplitude-limited nonlinearities. *IEEE Trans Com-mun.* **53**(11), 1964–1972 (2005)
11. C Rapp, Effects of HPA-nonlinearity on a 4-DPSK/OFDM-signal for a digital sound broadcasting system, in *Second European Conference on Satellite Communications, ECSC-2*, (Liege, Belgium, 1991), pp. 179–184
12. JJ Bussgang, Crosscorrelation functions of amplitude-distorted gaussian signals, (Technical report (Massachusetts Institute of Technology. Research Laboratory of Electronics), 1952) Report no.: **216**,
13. TB Sorensen, PE Mogensen, F Frederiksen, Extension of the ITU channel models for wideband (OFDM) systems, in *IEEE 62nd Vehicular Technology Conference 2005 (VTC-2005-Fall)*, (Dallas, Texas, USA, 2005), pp. 392–396
14. Y Yang, T Ihalainen, M Rinne, M Renfors, Frequency-domain equalization in single-carrier transmission: filter bank approach. *EURASIP J Adv Signal Process* (2007). **2007**, (Article ID 10438)
15. MV Clark, Adaptive frequency-domain equalization and diversity combining for broadband wireless communications. *IEEE J Sel Areas Commun.* **16**(8), 1385–1395 (1998). doi:10.1109/49.730448
16. E Alameda-Hemndez, DC McLernon, AG Orozco-Lugo, MM Lara, M Ghogho, Improved synchronization for superimposed training based channel estimation, in *IEEE/SP 13th Workshop on Statistical Signal Processing*, (Bordeaux, France, 2005), pp. 1324–1329
17. SMA Moosvi, DC McLernon, AG Orozco-Lugo, MM Lara, M Ghogho, Carrier frequency offset estimation using data-dependent superimposed training. *IEEE Commun Lett.* **12**(3), 179–181 (2008)
18. Y Yang, T Ihalainen, M Renfors, Filter bank based frequency domain equalizer in single carrier modulation, in *Proc 14th IST Mobile & Wireless Communications Summit*, (Dresden, Germany, 2005)
19. M Pukkila, Iterative Receivers and Multichannel Equalisation for Time Division Multiple Access Systems, Ph.D. dissertation, (Helsinki University of Technology, Espoo, Finland, 2003) ISBN 951-22-6717-9
20. R Carrasco-Alvarez, R Parra-Michel, AG Orozco-Lugo, JK Tugnait, Enhanced channel estimation using superimposed training based on universal basis expansion. *IEEE Trans Signal Process.* **57**(3), 1217–1222 (2009)
21. T Levanen, M Renfors, Improved performance bounds for iterative IC LMMSE channel estimator with SI pilots, in *21st Annual IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, (Istanbul, Turkey, 2010), pp. 9–14
22. 3GPP TS36.101 V10.1.0 (2010-12), 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (Release 10) http://www.3gpp.org/ftp/Specs/archive/36_series/36.101/36101-a10.zip (2012)
23. R Dinis, C-T Lam, D Falconer, Joint frequency-domain equalization and channel estimation using superimposed pilots, in *IEEE Wireless Communications and Networking Conference (WCNC)*, (Las Vegas, NV, USA, 2008), pp. 447–452
24. T Levanen, J Talvitie, M Renfors, Improved performance analysis for superimposed pilot based short channel estimator, in *IEEE International Workshop on Signal Processing Advances for Wireless Communications SPAWC 2010*, (Marrakech, Morocco, 2010), pp. 1–6
25. C Berrou, A Glavieux, P Thitimajshima, Near shannon limit error-correcting coding and decoding: turbo-codes, in, in *IEEE International Conference on Communications*, vol. 2. (Geneva, 1993), pp. 1064–1070
26. P Robertson, E Villebrun, P Hoehner, A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain, in *Proc IEEE International Conference on Communications ICC 95 Gateway to Globalization*, vol. 2. (Seattle, WA, USA, 1995), pp. 1009–1013
27. S Sharma, S Attri, RC Chauhan, A simplified and efficient implementation of FPGA-based turbo decoder, in *2003 IEEE International Performance Computing and Communications Conference*, (Longowal, Sangnu, 2003), pp. 207–213
28. CL Kei, WH Mow, Improved stopping criteria for iterative decoding of short-frame multi-component turbo codes, in *Proc IEEE International Conference on Communications Circuits and Systems and West Sino Expositions*, vol. 1. (Chengdu, Sichuan, China, 2002), pp. 42–45
29. D Divsalar, F Pollara, Turbo codes for PCS applications, in *Proc IEEE International Conference on Communications ICC'95 Gateway to Globalization*, vol. 1. (Seattle, WA, USA, 1995), pp. 54–59
30. 3rd Generation Partnership Project. 3GPP TS 25.212 V7.2.0 (2006-09); 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Multiplexing and channel coding (FDD) (Release 7) ftp://ftp.3gpp.org/specs/2006-09/Rel-7/25_series/25212-720.zip (2011)
31. R Gold, Optimal binary sequences for spread spectrum multiplexing. *IEEE Trans Inf Theory.* **13**(4), 619–621 (1967)

doi:10.1186/1687-1499-2012-49

Cite this article as: Levanen et al.: Performance evaluation of time-multiplexed and data-dependent superimposed training based transmission with practical power amplifier model. *EURASIP Journal on Wireless Communications and Networking* 2012 **2012**:49.