Non-Coherent Demapping for Stationary Rayleigh Fading Channels using Semidefinite Programming

Steven Corroy*, Meik Dörpinghaus** and Rudolf Mathar*
*Institute for Theoretical Information Technology, RWTH Aachen University, Aachen, Germany
**Vodafone Chair Mobile Communications Systems, Technische Universität Dresden, Dresden, Germany
{corroy, mathar}@ti.rwth-aachen.de, meik.doerpinghaus@tu-dresden.de

Abstract—**In this paper, we consider the problem of detecting a QPSK signal transmitted over a flat-fading channel without using pilot symbols. Instead, our aim is to approximate ML sequence detection, which is a nonconvex and discrete problem. To solve this problem, we use semidefinite relaxation and randomization to find a solution close to optimal. We propose a new solver based on the alternative direction method of multipliers to solve this problem efficiently. To verify this approach, we have developed and implemented a complete communication framework. We thoroughly simulated the conventional scheme using pilot-based channel estimation, as well as our newly proposed method and compared the achieved bit error rates. For high channel dynamics, our approach outperforms the pilot based detection, while saving bandwidth and transmit power.**

I. Introduction

Practical mobile communication systems face the problem that communication takes place over a time varying channel, whose realization is a priori unknown to the receiver. To allow for a coherent detection, which is a prerequisite for a small receiver complexity, the channel is estimated based on pilot symbols, periodically introduced into the transmit symbol sequence. The amount of required pilot symbols increases with the channel dynamics. For high channel dynamics the pilot overhead becomes significant, as these pilot symbols consume transmit power and channel bandwidth. Non-coherent detection schemes, which do not require a channel estimate, might be an alternative.

In the present work, the use of non-coherent detection is studied. One of the main obstacles of non-coherent ML-detection is that it is quite complex and cannot be solved in a reasonable amount of time. We tackle this problem by solving a relaxed version of the detection/demapping problem. To find a solution close to optimal, we propose to perform a semidefinite relaxation and use a randomization method to generate a list of transmit sequences with a high likelihood and perform demapping based on this list.

There exist various approaches to non-coherent detection. They mainly try to decrease the complexity of the problem by using approximations. In [1] and [2], the authors consider a block based demodulation method. In [3], the authors propose a multiple differential detector receiver structure which exploits the statistical characteristics of the fading process. In [4], the authors reduce the receiver complexity by discretizing the phase space. The main difference with our work is that we search a solution close to optimal for the original problem using convex optimization instead of solving an approximated problem exactly. A similar approach, based on convex optimization, is considered in [5]. However, contrary to [5], our demapper generates soft information as input for the decoder.

This work was partly supported by the UMIC research cluster of the RWTH Aachen University and by the DFG under grant DO 1568/1-1.

The remainder of the present paper is organized as follows. In Section II, we present the fading channel model we are using and describe different demapping methods, including the typical pilot-aided method and our newly proposed method based on convex optimization. In Section III, we propose a new convex solver for performing the demapping, compare the complexity of the approach with existing ones and discuss parallelization issues. In Section IV, we describe our simulation setup and in Section V we present simulation results. Finally, Section VI concludes the present paper.

II. System Model

We consider a discrete-time flat-fading channel, with output at time k given by

$$y_k = h_k \cdot x_k + n_k, \tag{1}$$

where x_k is the transmitted symbol, h_k the channel fading coefficient, n_k the additive noise component and y_k is the channel output symbol with $y_k, h_k, x_k, n_k \in \mathbb{C}$. We consider n time instances and denote $\mathbf{y} = [y_1, \ldots, y_n]^T$ the vector containing the channel output symbols in temporal order. Analogously, we define $\mathbf{x}, \mathbf{n}$ and $\mathbf{h}$, the sorted vectors containing all x_k, n_k, h_k's respectively. Equation (1) can be expressed in vector form as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} = \mathbf{X}\mathbf{h} + \mathbf{n}, \tag{2}$$

where $\mathbf{H} = \mathrm{diag}(\mathbf{h})$ and $\mathbf{X} = \mathrm{diag}(\mathbf{x})$.

The components of the noise vector $\mathbf{n}$ are i.i.d. zero-mean jointly proper Gaussian with variance σ_n^2.

The channel fading vector $\mathbf{h}$ is a zero-mean jointly proper Gaussian process. The entries of $\mathbf{h}$ are unknown to the receiver a priori but the statistic of this process is known. Its temporal correlation is given by $r_h(l) = \mathbb{E}[h_{k+l}h_k^*]$ and its variance by $r_h(0) = \sigma_h^2$. In matrix form, we denote $\mathbf{R}_h = \mathbb{E}[\mathbf{h}\mathbf{h}^{\mathrm{H}}]$. We define the power spectral density (PSD) of the channel fading process as $S_h(f) = \sum_{m=-\infty}^{\infty} r_h(m)e^{-j2mf}$ for $|f| \leq 0.5$. We denote f_d as the maximum Doppler frequency, it holds $0 < f_d < 0.5$. The PSD is assumed to be supported within the interval $[-f_d, f_d]$, which implies that $S_h(f) = 0$ for $f \notin [-f_d, f_d]$. This simply models the fact that the velocity of the

ISBN 978-3-8007-3529-7 © VDE VERLAG GMBH · Berlin · Offenbach, Germany

transmitter, the receiver, and other objects is limited. To ensure ergodicity, we exclude the case $f_d = 0$.

The transmit symbol sequence consists of data symbols $x_k \in \mathcal{S}$, where $\mathcal{S}$ is a constellation set, e.g., QPSK. The domain of the complete sequence $\mathbf{x}$ is the set $\mathcal{C} = \mathcal{S}^n$. Typically the actual transmitted sequences are in a more limited set because of channel coding and we have $\mathbf{x} \in \mathcal{C}_b$, with $\mathcal{C}_b \in \mathcal{C}$. The transmit symbols are independent and have variance σ_x^2. The mean SNR is given by $\sigma_x^2\sigma_h^2/\sigma_n^2$.

A. Optimal Demapping

The task of non-coherent demapping is to minimize the conditional probability density function $p(\mathbf{y}|\mathbf{x})$ for a given received signal $\mathbf{y}$ over all possible transmitted $\mathbf{x}$. The function $p(\mathbf{y}|\mathbf{x})$ is given by

$$p(\mathbf{y}|\mathbf{x}) = \frac{\exp(-\mathbf{y}^{\mathrm{H}}(\mathbf{X}\mathbf{R}_h\mathbf{X}^{\mathrm{H}} + \sigma_n^2\mathbf{I})^{-1}\mathbf{y})}{\pi^n \det(\mathbf{X}\mathbf{R}_h\mathbf{X}^{\mathrm{H}} + \sigma_n^2\mathbf{I})}, \tag{3}$$

where $\mathbf{I}$ is the identity matrix of appropriate size. The optimal demapping consists in solving the following problem

$$\begin{aligned} &\underset{\mathbf{x}}{\text{maximize}} \quad p(\mathbf{y}|\mathbf{x}) \\ &\text{subject to} \quad \mathbf{x} \in \mathcal{C}_b. \end{aligned} \tag{4}$$

This problem cannot be solved exactly efficiently.

B. Pilot-Aided Demapping

In practice, the channel is estimated based on pilot symbols periodically inserted into the transmit sequence. In other words, the transmitted signal $\mathbf{x}$ contains data symbols and pilot symbols that are known to the receiver in advance. We call x_p a pilot symbol, T_p the pilot spacing and $\mathbf{y}_p$ the vector of received symbols consisting of all pilot symbols, i.e., wlog.,

$$y_p(i) = y((i-1)T_p + 1), \quad i = 1, \ldots, \lceil n/T_p \rceil. \tag{5}$$

Denote $\mathbf{h}_p$, defined analogously to $\mathbf{y}_p$, as the vector of fading channels coefficients corresponding to the pilot symbols. Further we define two matrices $\mathbf{R}_{h_p} = \mathbb{E}[\mathbf{h}_p\mathbf{h}_p^{\mathrm{H}}]$ and $\mathbf{R}_{hh_p} = \mathbb{E}[\mathbf{h}\mathbf{h}_p^{\mathrm{H}}]$. Finally $\mathbf{X}_p$ denotes a diagonal matrix of size $\lceil n/T_p \rceil$, containing only pilot symbols defined as $\mathbf{X}_p = x_p\mathbf{I}$. The MMSE estimate $\hat{\mathbf{h}}$ of $\mathbf{h}$ based on the pilot symbols is calculated as follows

$$\hat{\mathbf{h}} = \mathbf{R}_{hh_p}\mathbf{X}_p^{\mathrm{H}}(\mathbf{X}_p\mathbf{R}_{h_p}\mathbf{X}_p^{\mathrm{H}} + \sigma_n^2\mathbf{I})^{-1}\mathbf{y}_p. \tag{6}$$

The covariance matrix of the estimation error $\mathbf{R}_e$ is given by

$$\mathbf{R}_e = \mathbf{R}_h - \mathbf{R}_{hh_p}\mathbf{X}_p(\mathbf{X}_p\mathbf{R}_{h_p}\mathbf{X}_p^{\mathrm{H}} + \sigma_n^2\mathbf{I})^{-1}\mathbf{X}_p^{\mathrm{H}}\mathbf{R}_{hh_p}^{\mathrm{H}}. \tag{7}$$

First, the pilot symbols rate has to be at least as large as the Nyquist rate. Second, there is a trade-off regarding the number of pilot to be used. A larger number of pilots enables a better channel estimation. On the other side, it requires more bandwidth and transmit power. For more details on this trade-off refer to [6]. Finally we fed the signal $\mathrm{diag}(\hat{\mathbf{h}})^{-1}\mathbf{y}$ to a demodulator [7], where the total noise variance V_k for symbol k is given by

$$V_k = (\mathbf{R}_e(k,k)\sigma_x^2 + \sigma_n^2)/|\hat{h}_k|^2. \tag{8}$$

The main advantage of pilot-based demapping is its low computational complexity. The main drawback is that, for high channel dynamics, its achievable rate falls significantly below the channel capacity [8, Fig. 4].

C. Semidefinite Programming Demapping

Most communication systems perform demapping and decoding in two steps. In the present work, we focus on demapping, i.e., we search for sequences $\mathbf{x}$ in $\mathcal{C}$ and not in $\mathcal{C}_b$. The main idea of the present work is to relax the problem (4), with the constraint $\mathbf{x} \in \mathcal{C}$, to find an upper bound on the maximal value of $p(\mathbf{y}|\mathbf{x})$ and to obtain a feasible solution using the randomization method. For simplicity, we consider a QPSK modulation scheme, i.e.,

$$\mathcal{S} = \{1+i, 1-i, -1+i, -1-i\}. \tag{9}$$

In this case, $\det(\mathbf{X}\mathbf{R}_h\mathbf{X}^{\mathrm{H}} + \sigma_n^2\mathbf{I}) = \det(2\mathbf{R}_h + \sigma_n^2\mathbf{I})$ is constant and maximizing $p(\mathbf{y}|\mathbf{x})$ is equivalent to minimizing $\mathbf{y}^{\mathrm{H}}(\mathbf{X}\mathbf{R}_h\mathbf{X}^{\mathrm{H}} + \sigma_n^2\mathbf{I})^{-1}\mathbf{y}$ since the exponential function is strictly increasing. Finally it holds that

$$\begin{aligned} \mathbf{y}^{\mathrm{H}}(\mathbf{X}\mathbf{R}_h\mathbf{X}^{\mathrm{H}}+\sigma_n^2\mathbf{I})^{-1}\mathbf{y} &= \mathbf{y}^{\mathrm{H}}\mathbf{X}(\mathbf{R}_h+(\sigma_n^2/2)\mathbf{I})^{-1}\mathbf{X}^{\mathrm{H}}\mathbf{y} \\ &= \mathbf{x}^{\mathrm{H}}(\mathbf{Y}^{\mathrm{H}}(\mathbf{R}_h+(\sigma_n^2/2)\mathbf{I})^{-1}\mathbf{Y})^*\mathbf{x}. \end{aligned}$$

The optimal non-coherent demapping problem is equivalent to

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}} \quad \mathbf{x}^{\mathrm{H}}(\mathbf{Y}^{\mathrm{H}}(\mathbf{R}_h + (\sigma_n^2/2)\mathbf{I})^{-1}\mathbf{Y})^*\mathbf{x} \\ &\text{subject to} \quad x_k \in \mathcal{S}, \quad k = 1, \ldots, n.. \end{aligned} \tag{10}$$

The elements of $\mathcal{S}$ are complex and for simplification we use the equivalent real formulation of (10). Using the Cholesky decomposition we factorize the matrix in the objective function

$$(\mathbf{Y}^{\mathrm{H}}(\mathbf{R}_h + (\sigma_n^2/2)\mathbf{I})^{-1}\mathbf{Y})^* = \mathbf{L}\mathbf{L}^{\mathrm{H}}. \tag{11}$$

We define the matrices

$$\mathbf{A} = \mathbf{B}\mathbf{B}^{\mathrm{T}}, \quad \mathbf{B} = \begin{bmatrix} \mathrm{Re}(\mathbf{L}) & -\mathrm{Im}(\mathbf{L}) \\ \mathrm{Im}(\mathbf{L}) & \mathrm{Re}(\mathbf{L}) \end{bmatrix}. \tag{12}$$

Finally, we denote $\mathbf{z} = [\mathrm{Re}(\mathbf{x})^{\mathrm{T}}\ \mathrm{Im}(\mathbf{x})^{\mathrm{T}}]^{\mathrm{T}}$ and the problem (10) is equivalent to the problem

$$\begin{aligned} &\underset{\mathbf{z}}{\text{minimize}} \quad \mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z} \\ &\text{subject to} \quad z_k \in \{-1, 1\}, \quad k = 1, \ldots, 2n. \end{aligned} \tag{13}$$

This problem is nonconvex and cannot be solved efficiently. In the following, we propose a method to find a solution close to optimal.

1) Lower Bound through Semidefinite Relaxation: We take the semidefinite relaxation of problem (13). See [9] and [10] for an in-depth presentation. For the sake of completeness we briefly sum-up this relaxation method:

1) We perform a variable change $\mathbf{Z} = \mathbf{z}\mathbf{z}^{\mathrm{T}}$ and introduce two constraints, $\mathbf{Z} \succcurlyeq 0$ and $\mathrm{rank}(\mathbf{Z}) = 1$.
2) Since $z_k \in \{-1, 1\} \Leftrightarrow z_k^2 = 1$, this constraint is equivalent to $\mathrm{diag}(\mathbf{Z}) = \mathbf{1}$, where $\mathbf{1}$ is a vector containing only 1's.
3) The actual semidefinite relaxation is performed by dropping the rank constraint.

© VDE VERLAG GMBH · Berlin · Offenbach, Germany

The relaxed problem can be therefore expressed as

$$\begin{aligned} \underset{\mathbf{Z}}{\text{minimize}} \quad & \operatorname{tr}(\mathbf{AZ}) \\ \text{subject to} \quad & \mathbf{Z} \succcurlyeq 0,\ \operatorname{diag}(\mathbf{Z}) = \mathbf{1}. \end{aligned} \tag{14}$$

This problem is a convex optimization problem and can be solved efficiently using standard tools (e.g., interior point methods). The optimal value of problem (14) is a lower bound on the optimal value of problem (13). We will use this lower bound to test the quality of diverse heuristics delivering a feasible solution.

2) Feasible Solution using the Randomization Method: In order to generate a feasible solution we will use the solution of problem (14) in the randomization method [10]. This method works as follows

1) Solve (14) and get $\mathbf{Z}^\star$.
2) Draw $\hat{\mathbf{z}} \sim \mathcal{N}(\mathbf{0}, \mathbf{Z}^\star)$.
3) Take $\mathbf{z} = \operatorname{sgn}(\hat{\mathbf{z}})$.
4) Repeat from 2. and keep the best $\mathbf{z}$.

After only few iterations, the randomization method delivers a good feasible solution.

III. Solving the Relaxed Problem

The most computing power intensive step in this new detection method is solving problem (14). This optimization problem is a semidefinite problem, which can be solved in polynomial time. In order to have a clear view on the complexity of such a solver and the possibilities to parallelize it, we propose in this section, a new convex solver based on the alternating direction method of multipliers (ADMM) [11]. We recast problem (14) to adapt it to the ADMM framework as follows

$$\begin{aligned} \underset{\mathbf{Z},\mathbf{W}}{\text{minimize}} \quad & \operatorname{tr}(\mathbf{AZ}) + \mathrm{i}_{\succcurlyeq 0}(\mathbf{W}) \\ \text{subject to} \quad & \mathbf{Z} = \mathbf{W} \\ & \operatorname{diag}(\mathbf{Z}) = \mathbf{1}, \end{aligned} \tag{15}$$

where $\mathrm{i}_{\succcurlyeq 0}(\mathbf{W})$ is the indicator function of the set of positive semidefinite matrices and is equal to zero if $\mathbf{W} \succcurlyeq 0$ and $+\infty$ otherwise. Note that a typical ADMM problem would only have the equality constraint $\mathbf{Z} = \mathbf{W}$. In the following, we modify the ADMM algorithm to account for a second equality constraint. The augmented Lagrangian for problem (15) is given by

$$\begin{aligned} L_\rho(\mathbf{Z}, \mathbf{W}, \mathbf{U}) \;=\; & \operatorname{tr}(\mathbf{AZ}) + \mathrm{i}_{\succcurlyeq 0}(\mathbf{W}) \\ & +(\rho/2) \left\| \begin{bmatrix} \mathbf{Z} - \mathbf{W} & \operatorname{diag}(\mathbf{Z}) - \mathbf{1} \end{bmatrix} + \mathbf{U} \right\|_F^2 \end{aligned}$$

where $\|.\|_F$ denotes the Frobenius norm, ρ is a solver intern parameter and the matrix $\mathbf{U}$ is the scaled dual variable and is defined as the concatenation

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{u}_2 \end{bmatrix},$$

where $\mathbf{U}_1$ and $\mathbf{u}_2$ are the scaled dual variables corresponding to the first and second equality constraints of (15) respectively.

The ADMM algorithm is an iterative method. At each iteration q the variable updates take the form

$$\begin{aligned} \mathbf{Z}^{(q+1)} &= \underset{\mathbf{Z}}{\operatorname{argmin}}(\operatorname{tr}(\mathbf{AZ}) \\ &\quad +(\rho/2) \left\| \begin{bmatrix} \mathbf{Z} - \mathbf{W}^{(q)} & \operatorname{diag}(\mathbf{Z}) - \mathbf{1} \end{bmatrix} + \mathbf{U}^{(q)} \right\|_F^2) \end{aligned}$$

$$\begin{aligned} \mathbf{W}^{(q+1)} &= \underset{\mathbf{W}}{\operatorname{argmin}}(\mathrm{i}_{\succcurlyeq 0}(\mathbf{W}) \\ &\quad +(\rho/2) \left\| \begin{bmatrix} \mathbf{Z}^{(q+1)} - \mathbf{W} & \operatorname{diag}(\mathbf{Z}^{(q+1)}) - \mathbf{1} \end{bmatrix} + \mathbf{U}^{(q)} \right\|_F^2) \end{aligned}$$

$$\mathbf{U}^{(q+1)} = \mathbf{U}^{(q)} + \begin{bmatrix} \mathbf{Z}^{(q+1)} - \mathbf{W}^{(q+1)} & \operatorname{diag}(\mathbf{Z}^{(q+1)}) - \mathbf{1} \end{bmatrix}.$$

The key for ADMM to solve our problem quickly is to be able to solve the $\mathbf{Z}$- and $\mathbf{W}$-updates fast, at best analytically. Due to space constraints, we explain this next only briefly.

A. $\mathbf{Z}$-update

Let us define the following matrix

$$\mathbf{T} = -(1/\rho)\mathbf{A} + \mathbf{W}^{(q)} + \mathbf{U}_1^{(q)} + \operatorname{diag}(\mathbf{1} - \mathbf{u}_2^{(q)}).$$

By derivating the matrix function $\operatorname{tr}(\mathbf{AZ}) + (\rho/2) \left\| \begin{bmatrix} \mathbf{Z} - \mathbf{W}^{(q)} & \operatorname{diag}(\mathbf{Z}) - \mathbf{1} \end{bmatrix} + \mathbf{U}^{(q)} \right\|_F^2$ and setting the result to zero, it can be shown that

$$\mathbf{Z}^{(q+1)} = \mathbf{T} \circ (\mathbf{1}\mathbf{1}^{\mathrm{T}} - (1/2)\mathbf{I}), \tag{16}$$

where $\circ$ denotes the Hadamard product.

B. $\mathbf{W}$-update

For this update, the matrix $\mathbf{W}$ has to be positive semidefinite and close to $\mathbf{Z}^{(q+1)} + \mathbf{U}_1^{(q)}$ in the Frobenius norm sense. It is shown in [11] that the $\mathbf{W}$-update is in this case given by

$$\mathbf{W}^{(q+1)} = \prod_{\succcurlyeq 0} (\mathbf{Z}^{(q+1)} + \mathbf{U}_1^{(q)}), \tag{17}$$

where $\prod_{\succcurlyeq 0}(\mathbf{Z}^{(q+1)} + \mathbf{U}_1^{(q)})$ is the projection of $\mathbf{Z}^{(q+1)} + \mathbf{U}_1^{(q)}$ on the set of positive semidefinite matrices. As explained in [12], we can find $\mathbf{W}^{(q+1)}$ by the eigenvalue decomposition $\mathbf{Z}^{(q+1)} + \mathbf{U}_1^{(q)} = \mathbf{V\Lambda V}^{\mathrm{T}}$, where $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\mathbf{Z}^{(q+1)} + \mathbf{U}_1^{(q)}$. We have

$$\mathbf{W}^{(q+1)} = \mathbf{V}(\mathbf{\Lambda})^{+}\mathbf{V}^{\mathrm{T}},$$

where $\lambda_{ii}^+ = \lambda_{ii}$ if $\lambda_{ii} > 0$ and $\lambda_{ii}^+ = 0$ otherwise.

C. Discussion

1) Complexity: The $\mathbf{Z}$-update and $\mathbf{U}$-update steps have a linear complexity in terms of the number of entries N of $\mathbf{Z}$ (here $N = 4n^2$). The only nonlinear step is the $\mathbf{W}$-update, which requires an eigenvalue decomposition, which has complexity $\mathcal{O}(N^3)$. It is proven that ADMM converges [11] and we assume that it converges in a number of steps bounded by ζ. So we have to solve at most ζ problems with complexity $\mathcal{O}(N^3)$. The exact value of ζ depends on ρ, is difficult to determine and is out of the scope of this paper. Nevertheless we could experimentally verify that the cost of the eigenvalue decomposition in the $\mathbf{W}$-update absolutely dominates the overall run time.

ISBN 978-3-8007-3529-7
© VDE VERLAG GMBH · Berlin · Offenbach, Germany

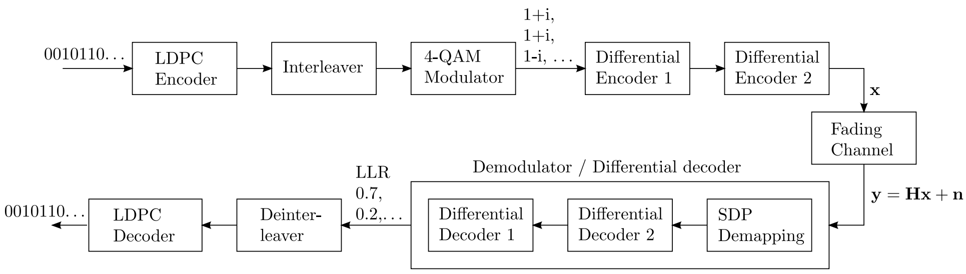

Fig. 1. Simulation framework.

2) Parallelization: As the eigenvalue decomposition in the $\mathbf{W}$-update, is the most expensive operation performed by the solver, this is where we can gain a lot of time using parallelization. A common and simple way to perform an eigenvalue decomposition in an iterative and parallel manner is the well-known parallel Jacobi algorithm [13], which can significantly reduce the overall run time of the solver. The $\mathbf{Z}$-update and $\mathbf{U}$-update are trivially parallelized.

3) Comparison with Optimal and Pilot-Based Detection: The pilot-based detection has a linear complexity and is much more computation efficient than the convex optimization based detection. However the proposed method has a much lower complexity than optimal demapping.

IV. Simulation Framework

In order to evaluate numerically the performance of our proposed convex optimization-based demapping and to compare it to the coherent demapping using a pilot-based channel estimate, we describe our system setup illustrated in Figure 1.

In our simulations we transmit bit sequences with a length of 504 bits. We use a low density parity check (LDPC) code of rate $r = 1/2$, with a parity check matrix of size 504×1008 [14]. A random bit interleaver is used. After that we map two bits to a QPSK symbol. At his point we have a sequence of $n = 504$ symbols. Next we proceed to two sequential rounds of differential encoder. Denoting an input symbol as $x(k) = \sqrt{2}\mathrm{e}^{i\alpha(k)}$ and an output symbol as $x_{\text{enc}}(k) = \sqrt{2}\mathrm{e}^{i\alpha_{\text{enc}}(k)}$, one round is described as follows,

$$\alpha_{\text{enc}}(k) = \alpha(k) - \alpha_{\text{enc}}(k-1) + \pi/4, \tag{18}$$

with $\alpha_{\text{enc}}(1) = \alpha(1)$. The reason why to use differential encoding and especially two rounds of it is subtle. When looking at the objective function of problem (10), we can see that given a specific feasible point $\mathbf{x}$, then each other points $\mathbf{x}_r$ that are a constant rotation of $\mathbf{x}$, i.e., $\mathbf{x}_r = \mathbf{x}\mathrm{e}^{i\theta}$, achieve the same objective value. In other words, problem (10) has at least 4 different solutions assuming a QPSK modulation scheme. In terms of convex optimization, none of these solutions is better than the other. However for a communication system, we are only interested in the transmitted signal, all other solutions are actually completely wrong. Most communication systems tackle this problem by only using distinguishable sequences. However we are not solving the problem exactly and we might have a large range of feasible points, with similar objective values, but with a much different number of demapping errors. This is a huge problem, since we cannot say that a demapped signal with a better objective value is closer to the transmitted signal. The two rounds of differential decoders are present for this reason. We get rid of a main rotation using the first differential decoder and some internal local rotations with the second one. We could add more encoder/decode pairs, but each new pair starts to introduce random errors in the demapped signal.

The fading channel realization is generated using the sum of sinusoids model as described by Jakes [15]. After the signal is received, the demapping is performed as follows.

1) Generate a feasible vector $\mathbf{z}_l$ as in Section II-C2.
2) Calculate its objective value p_l, according to (13).
3) Perform two rounds of differential decoding on the complex representation of $\mathbf{z}_l$. Denoting an input symbol as $x(k) = \sqrt{2}\mathrm{e}^{i\alpha(k)}$ and an output symbol as $x_{\text{dec}}(k) = \sqrt{2}\mathrm{e}^{i\alpha_{\text{dec}}(k)}$, one round is described as follows,

$$\alpha_{\text{dec}}(k) = \alpha(k) + \alpha(k-1) - \pi/4, \tag{19}$$

   with $\alpha_{\text{dec}}(1) = \alpha(1)$.
4) Store $\mathbf{x}_l$, the results of the differential decoding, and p_l.
5) Repeat n_{rand} times (For our results, we generated $n_{\text{rand}} = 1000$ vectors from $|\mathcal{S}|^n = 4^{504}$ possible vectors).

Using the stored vectors $\mathbf{x}_l$ with $l = 1, \ldots, n_{\text{rand}}$ (each of size n), we calculate a log-likelihood ratio (LLR) vector of size $2n$, containing the LLR for all code bits as follows

$$\text{LLR}(k) = \log\left(\frac{\sum_{\{\mathbf{x}_l | c_k(\mathbf{x}_l)=0\}} p_l}{\sum_{\{\mathbf{x}_l | c_k(\mathbf{x}_l)=1\}} p_l}\right), \tag{20}$$

where $c_k(\mathbf{x}_l)$ is k-th bit of $\mathbf{x}_l$. The LLR vector is then fed to the deinterleaver and finally to the LDPC decoder.

V. Simulation Results

In this section, we evaluate the performance of the convex optimization based demapping and the pilot-aided demapping. The main factor influencing the results is f_d. In Fig. 2, we simulate a fading channel with a relatively high dynamic with $f_d = 10^{-2}$. We plot the bit error rate (BER) with respect to the E_b/N_0, where E_b is the energy per information bit and N_0 is noise power spectral density for

- the coherent demapping using the pilot-aided channel estimate (6), denoted PA, for different pilot spacing T_p and
- the convex optimization-based demapping, denoted CO, for different block lengths. In this regard, the number of transmit symbols is 504. Solving the problem (14) might be very time consuming. To shorten it, we demodulate

ISBN 978-3-8007-3529-7 © VDE VERLAG GMBH · Berlin · Offenbach, Germany

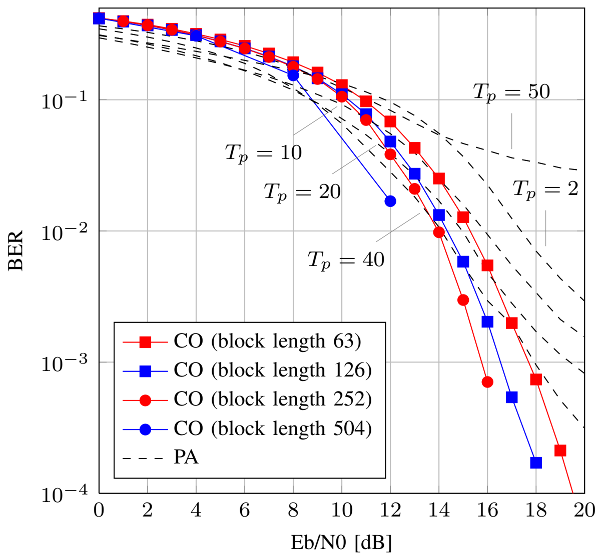

Fig. 2. BER with respect to Eb/No, $f_d = 0.01$.

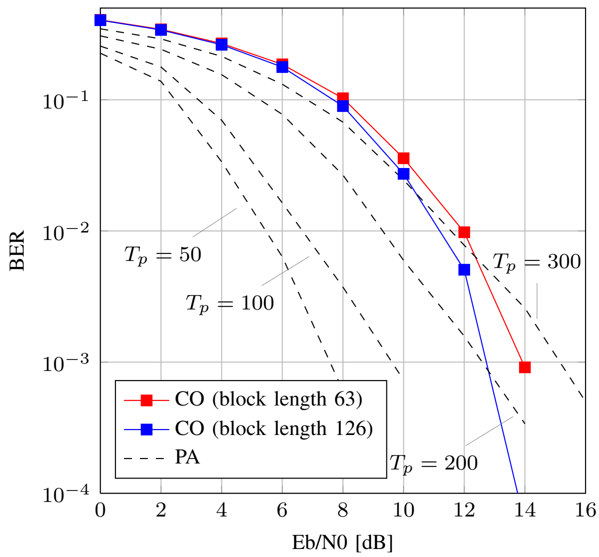

Fig. 3. BER with respect to Eb/No, $f_d = 0.001$.

smaller blocks of the received signal but we loose information by dropping the correlation between blocks.

Let us first analyze the PA results. Using a too large number of pilot symbols (e.g., $T_p = 2$) is not beneficial, it increases the required E_b/N_0 to achieve a specific BER and is inefficient. On the other side, using a too low number of pilot symbols (e.g., $T_p = 50$) leads to a poor channel estimation and consequently to a high BER. When comparing the CO demapping for different block length, as expected, a larger block length provides better performance, since we simply use more available information. Finally, we compare the PA and CO demapping methods. From a block length of 126 on, the CO demapping outperforms the PA demapping for a BER of 10^{-2}. For a block length of 252, we observe a difference of over 2dB at a BER of 10^{-3}.

Now we would like to analyze a less preferential scenario for the CO demapping, i.e., when the channel dynamic is lower. In Fig. 3, we simulate a fading channel with $f_d = 10^{-3}$. In this case the PA approach outperforms the CO approach for $T_p \leq 100$ at a BER of 10^{-3}, while the CO approach still performs reasonably and achieve an BER of $\sim 10^{-4}$ at 14dB. For a lower pilot frequency, the CO approach performs slightly better and is saving bandwidth and transmit power resources by using no pilots. In general, it is expected that using the CO approach with a block length of 504 will show a more favorable comparison.

VI. Conclusion

We have proposed a new method to perform non-coherent QPSK demapping using convex optimization. We first formulate the problem of minimizing the conditional probability density function of the received signal over all possible transmit signals. We then relax this problem using semidefinite relaxation and solve the relaxed problem using a new ADMM solver. The solution of the relaxed problem is then used as input for the randomization method that finally delivers a LLR measure of the transmit signal. We have implemented a complete communication framework, evaluated our new method and compared it to coherent detection using pilot symbols for channel estimation. We have shown that the convex optimization based demapping outperforms the pilot-aided demapping for high channel dynamics and performs well at lower dynamics while saving bandwidth and transmit power.

References

[1] D. Divsalar and M. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300–308, 1990.

[2] D. Warrier and U. Madhow, "Spectrally efficient noncoherent communication," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 651–668, 2002.

[3] I. Marsland and P. Mathiopoulos, "Multiple differential detection of parallel concatenated convolutional (turbo) codes in correlated fast Rayleigh fading," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 265–275, 1998.

[4] R. Chen, R. Koetter, U. Madhow, and D. Agrawal, "Joint noncoherent demodulation and decoding for the block fading channel: A practical framework for approaching Shannon capacity," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1676–1689, 2003.

[5] Z. Ma, P. Fan, E. G. Larsson, and B. Honary, "Quasi-maximum-likelihood multiple-symbol differential detection for time-varying Rayleigh fading channel," *Electronics Letters*, vol. 45, no. 22, pp. 1127–1128, 2009.

[6] J. Baltersee, G. Fock, and H. Meyr, "An information theoretic foundation of synchronized detection," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2115–2123, 2001.

[7] B. Sklar, *Digital Communications: Fundamentals and Applications*. Prentice-Hall, Inc., 1988.

[8] M. Dörpinghaus, H. Meyr, and R. Mathar, "On the achievable rate of stationary Rayleigh flat-fading channels with Gaussian inputs," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2208–2220, Apr. 2013.

[9] Y. Nesterov, "Semidefinite relaxation and nonconvex quadratic optimization," *Optim. Method. Softw.*, vol. 9, no. 1-3, pp. 141–160, 1998.

[10] D. Bertsimas and Y. Ye, "Semidefinite relaxations, multivariate normal distributions, and order statistics," *Handbook of Combinatorial Optimization*, vol. 3, pp. 1–19, 1998.

[11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.

[12] N. J. Higham, "Computing a nearest symmetric positive semidefinite matrix," *Linear Algebra and its Applications*, vol. 103, pp. 103 – 118, 1988.

[13] G. Golub and C. Van Loan, *Matrix Computations*. JHUP, 1996.

[14] D. J. MacKay. Encyclopedia of sparse graph codes. [Online]. Available: http://www.inference.phy.cam.ac.uk/mackay/codes/data.html

[15] W. Jakes, *Microwave Mobile Communications*. Wiley & Sons, 1975.

ISBN 978-3-8007-3529-7
© VDE VERLAG GMBH · Berlin · Offenbach, Germany