

Convex programming for detection in structured communication problems

T. Morsy¹, J. Götze¹, and H. Nassar²

¹Information Processing Lab., TU Dortmund, 44221 Dortmund, Germany

²Faculty of Computers and Informatics, Suez Canal University, 41522 Ismailia, Egypt

Abstract. The generalized Minimum Mean Squared Error (GMMSE) detector has a bit error rate performance, which is similar to the MMSE detector. The advantage of the GMMSE detector is that it does not require the knowledge of the noise power. However, the computational complexity of the GMMSE detector is significantly higher than the computational complexity of the MMSE detector. In this paper, the complexity of the GMMSE detector is reduced by taking into account the structure of the system matrix (Toeplitz). Furthermore, by using circular approximation of the structured system matrix an approximate GMMSE detector is presented, whose computational complexity is only slightly higher than MMSE, i.e. only an iterative gradient descent algorithm based on the inversion of diagonal matrices is additionally required.

1 Introduction

The maximum likelihood (ML) detection problem can be written as a quadratic optimization problem with integer constraints (Verdu, 1998). Unfortunately this problem is in general non-deterministic polynomial hard (NP-hard) (Verdu, 1989). This observation resulted in the development of many receivers that have reasonable complexity (Tan and Rasmussen, 2001; Chang and Han, 2008), e.g. the well-known least squares (LS) and minimum mean squared error (MMSE) detectors (Lupas and Verdu, 1989; Madhow and Honig, 1994) as the most simple cases.

Recently convex programming has been successfully employed to sub-optimally solve such detection problems. Using this kind of relaxation converts the discrete optimization problem into a continuous one which can be solved itera-

tively (Boyd and Vandenberghe, 2004). Generalized minimum mean squared error detector is one important detector that uses convex programming to solve the detection problem using unconstrained gradient descent algorithm (Yener, Yates and Ulukus, 2002). The advantages of this detector are that it has a BER performance which is similar to the MMSE detector, and does not require the knowledge of the noise power. Because of these advantages it can be used in scenarios where the noise power is changing rapidly or the noise power is unknown. Associated with these advantages of GMMSE detector there is the disadvantage, that it has a significantly higher computational complexity compared to MMSE detector.

In order to decrease this computational complexity the structure of the system matrix is used in this paper. First the Toeplitz structure of the channel convolution matrix is taken into consideration. In this case, computing the solution of the GMMSE detector requires the EVD of Toeplitz matrix but reduces the effort for the iterative gradient descent algorithm significantly. Nevertheless, computing the EVD of the Toeplitz matrix (using, e.g., Lanczos algorithm) is still computationally demanding. Therefore, we approximate the banded Toeplitz matrix by a circular matrix (Vollmer et al., 2001). In this case the MMSE/GMMSE solution is obtained by computing the EVD of the circular matrix using FFT/IFFT, such that the required EVD implies no additional effort. Therefore, the additional effort of approximate GMMSE is only determined by the iteration steps of the gradient descent algorithm based on the diagonal matrix containing the eigenvalues.

This paper is organized as follows: in Sect. 2 a system model for the detection problem and the convex relaxations of the problem are introduced. LS and MMSE detectors are described from the convex programming point of view in Sect. 3. GMMSE detector is described in Sect. 4. In Sect. 5 we introduce our new detector which is derived from the GMMSE detector taking into account the structure of the



Correspondence to: T. Morsy
(tharwat.morsy@tu-dortmund.de)

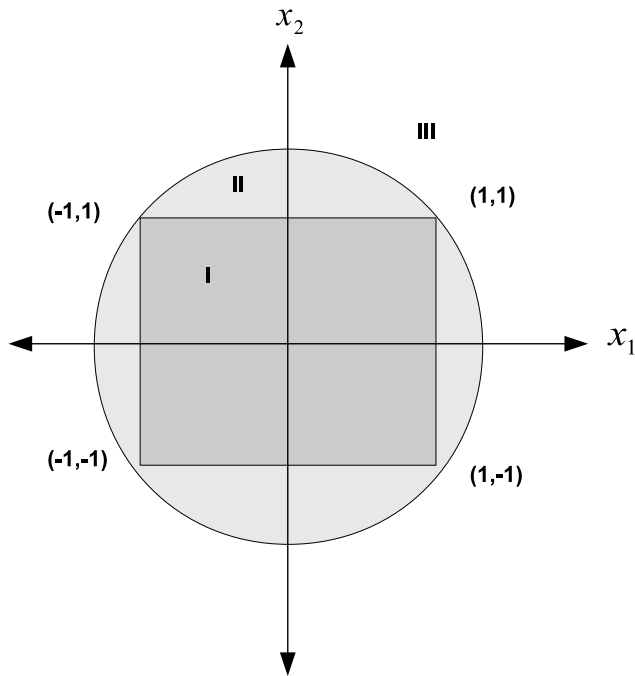


Fig. 1. Convex relaxations.

channel matrix. Simulation results are used to compare bit error rate (BER) of the different detectors in Sect. 6 and the computational complexity is discussed in Sect. 7. Conclusions are drawn in Sect. 8.

2 System model and its relaxations

Consider the system model in matrix form as

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n}. \quad (1)$$

The vector $\mathbf{r} \in \mathbb{R}^m$ is the received signal vector, the matrix $\mathbf{H} \in \mathbb{R}^{m \times n}$ is the channel matrix, and the vector $\mathbf{n} \in \mathbb{R}^m$ is additive white Gaussian noise with noise power σ^2 . The transmitted symbols $\mathbf{x} \in \mathbb{R}^n$ are drawn from Binary Phase Shift Keying (BPSK) constellation, i.e. $\mathbf{x} \in \{-1, +1\}^n$.

Under the white Gaussian noise assumption the ML detector of \mathbf{x} is given by

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \{-1, +1\}^n} \|\mathbf{r} - \mathbf{H}\mathbf{x}\|_2^2. \quad (2)$$

The ML problem in Eq. (2) can be equivalently written as

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \{-1, +1\}^n} \mathbf{x}^T \mathbf{H}^T \mathbf{H} \mathbf{x} - 2\mathbf{r}^T \mathbf{H} \mathbf{x}. \quad (3)$$

Substituting the value of the matched filter output

$$\mathbf{y} = \mathbf{H}^T \mathbf{r} \quad (4)$$

into Eq. (3), we get

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \{-1, +1\}^n} \mathbf{x}^T \mathbf{H}^T \mathbf{H} \mathbf{x} - 2\mathbf{y}^T \mathbf{x}. \quad (5)$$

This problem is NP hard and solving Eq. (5) by exhaustive search has a complexity which grows with 2^n (Verdu, 1989). This makes computationally less complex solutions of Eq. (5) interesting.

We use the benefits of convex programming as an important mathematical tool to solve problem Eq. (5) by relaxing the constraint set in Eq. (5). Our constraint set $\mathbf{x} \in \{-1, +1\}^n$ which contains only the corners of the unit hypercube is not a convex set. Therefore we relax this constraint set using three relaxations which yield a convex set.

Figure 1 shows the relaxed constraint sets for $n = 2$ taking into account that the original problem contains only the corners of the unit hypercube. Three relaxations are considered: relaxation of the constraint set to the whole unit hypercube (region I), relaxation of the constraint set to the sphere which covers the unit hypercube (region I+II), and the relaxation to the whole space (region I+II+III). The solution in each case can be mapped to the feasible set of the original problem by taking the sign of each component of the relaxed solution vector.

3 Least squares and MMSE detectors

We first discuss the LS and MMSE solution from the convex programming point of view. Relaxing the constraint set to be the whole space, i.e. (region I+II+III), problem Eq. (5) will take the form

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x}^T \mathbf{H}^T \mathbf{H} \mathbf{x} - 2\mathbf{y}^T \mathbf{x}. \quad (6)$$

The following theorem that was stated in (Boyd and Vandenberghe, 2004) describes LS and MMSE solutions from the convex programming point of view.

Theorem 1 Suppose that the objective function f in an unconstrained convex optimization problem is differentiable, so the well known necessary and sufficient optimality condition is

$$\nabla f = 0. \quad (7)$$

Applying condition Eq. (7) to problem Eq. (6), which has an objective function

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{H}^T \mathbf{H} \mathbf{x} - 2\mathbf{y}^T \mathbf{x}, \quad (8)$$

the necessary and sufficient optimality condition gives the solution

$$\mathbf{x}^* = \left(\mathbf{H}^T \mathbf{H} \right)^{-1} \mathbf{y}, \quad (9)$$

which is the well known least squares solution.

When the noise power σ^2 is known, using the same relaxation (region I+II+III) we get the minimum mean square error solution

$$\mathbf{x}^* = \left(\mathbf{H}^T \mathbf{H} + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{y}. \quad (10)$$

4 Generalized MMSE detector

If we relax the constraint set in problem Eq. (5) to be the sphere which contains the unit hypercube, i.e. (region I+II), then our detection problem takes the form

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}^T \mathbf{x} \leq n} \mathbf{x}^T \mathbf{H}^T \mathbf{H} \mathbf{x} - 2\mathbf{y}^T \mathbf{x}. \quad (11)$$

Since problem Eq. (11) has a convex objective function over a convex constraint set, it is a convex optimization problem and it has a unique minimum (Boyd and Vandenberghe, 2004). The convex duality theorem guarantees that no duality gap exists and one can solve for the dual problem instead (Nash and Sofer, 1996). Problem Eq. (11) has a single constraint such that there is only one dual variable and a simple iterative algorithm can be employed to solve this dual problem.

We can express the Lagrange dual function as

$$L(\mathbf{x}, \lambda) = \mathbf{x}^T \mathbf{H}^T \mathbf{H} \mathbf{x} - 2\mathbf{y}^T \mathbf{x} + \lambda (\mathbf{x}^T \mathbf{x} - n), \quad (12)$$

which is minimized over \mathbf{x} and maximized over $\lambda \geq 0$. Solving for \mathbf{x} in terms of λ and substituting back, we obtain

$$\max_{\lambda \geq 0} -\mathbf{y}^T (\mathbf{H}^T \mathbf{H} + \lambda \mathbf{I})^{-1} \mathbf{y} - \lambda n. \quad (13)$$

This problem has the advantage, that it is a one dimensional optimization problem. Now we can solve this dual problem Eq. (13) instead of the primal problem Eq. (11) because there is no duality gap between these two problems. Problem Eq. (13) can be solved by different iterative algorithms (Hansen, 1979). A simple unconstrained gradient descent algorithm given by

$$\bar{\lambda}(t+1) = \bar{\lambda}(t) + \mu \left(\mathbf{y}^T (\mathbf{H}^T \mathbf{H} + \bar{\lambda}(t) \mathbf{I})^{-2} \mathbf{y} - n \right), \quad (14)$$

converges to $\bar{\lambda}$ for a reasonable choice of μ . The solution of Eq. (13) is given by

$$\lambda^* = \max(0, \bar{\lambda}). \quad (15)$$

Then, the unique minimizer of Eq. (11) is

$$\mathbf{x}^* = (\mathbf{H}^T \mathbf{H} + \lambda^* \mathbf{I})^{-1} \mathbf{y}. \quad (16)$$

This solution looks familiar because of its similarity to the MMSE detector. When $\lambda^* = \sigma^2$, the GMMSE detector reduces to the MMSE detector. Therefore this detector which depends on the value of optimum dual solution λ^* is named *Generalized MMSE detector*. The advantage of the GMMSE detector is, that it improves the BER performance (Compared to LS detector as the MMSE detector) and does not require the knowledge of the noise power σ^2 . However GMMSE detector has the disadvantage that it requires a significantly higher computational complexity than MMSE detector.

5 Structured problem

In this Sect. we consider the reduction of the computational complexity of the GMMSE detector by taking the structures of the underlying detection problems into account. In the first case the Toeplitz structure of the channel convolution matrix \mathbf{H} is used. In this case we will express the matrix $\mathbf{H}^T \mathbf{H}$ in problem Eq. (11) by its eigenvalue decomposition $\mathbf{H}^T \mathbf{H} = \mathbf{V} \Lambda \mathbf{V}^T$ where \mathbf{V} is the matrix whose columns are the eigenvectors of $\mathbf{H}^T \mathbf{H}$ and Λ is a diagonal matrix that contains the corresponding eigenvalues as its diagonal elements. Problem Eq. (11) can be rewritten as

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}^T \mathbf{x} \leq n} \mathbf{x}^T (\mathbf{V} \Lambda \mathbf{V}^T) \mathbf{x} - 2\mathbf{y}^T \mathbf{x}. \quad (17)$$

The dual problem for the problem Eq. (17) in this case takes the form

$$\max_{\lambda \geq 0} -\mathbf{y}^T ((\mathbf{V} \Lambda \mathbf{V}^T) + \lambda \mathbf{I})^{-1} \mathbf{y} - \lambda n. \quad (18)$$

The unconstrained gradient descent algorithm takes the form

$$\bar{\lambda}(t+1) = \bar{\lambda}(t) + \mu \left(\mathbf{y}^T \mathbf{V} (\Lambda + \bar{\lambda}(t) \mathbf{I})^{-2} \mathbf{V}^T \mathbf{y} - n \right) \quad (19)$$

and the GMMSE solution will be

$$\mathbf{x}^* = \mathbf{V} (\Lambda + \lambda^* \mathbf{I})^{-1} \mathbf{V}^T \mathbf{y}. \quad (20)$$

Besides computing $\mathbf{V}^T \mathbf{y}$ only diagonal matrices must be converted in Eq. (19) and Eq. (20), which simplifies the computations significantly. We can also make use of the Toeplitz structure of $\mathbf{H}^T \mathbf{H}$ when computing the EVD by using the Lanczos algorithm (Golub and Loan, 1996). Although this approach reduces the computational complexity of the GMMSE detector significantly, it is still much more complex than MMSE because of the required EVD (the iterations of Eq. (19) on the diagonal matrices are only of $O(n)$). In the following we will discuss the circular structure case, which is obtained by an approximation of the Toeplitz case. A banded Toeplitz structure gives a circular structure by adding $L - 1$ columns to the Toeplitz matrix, where L is the length of the channel impulse response. This is shown in the following example for $L = 2$:

$$\mathbf{H} = \begin{pmatrix} h_1 & 0 \\ h_2 & h_1 \\ 0 & h_2 \end{pmatrix} \rightarrow \begin{pmatrix} h_1 & 0 & h_2 \\ h_2 & h_1 & 0 \\ 0 & h_2 & h_1 \end{pmatrix} = \tilde{\mathbf{H}}.$$

If channel matrix \mathbf{H} in problem Eq. (11) is approximated by circular matrix $\tilde{\mathbf{H}}$ we obtain

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}^T \mathbf{x} \leq n} \mathbf{x}^T \tilde{\mathbf{H}}^T \tilde{\mathbf{H}} \mathbf{x} - 2\mathbf{y}^T \mathbf{x}. \quad (21)$$

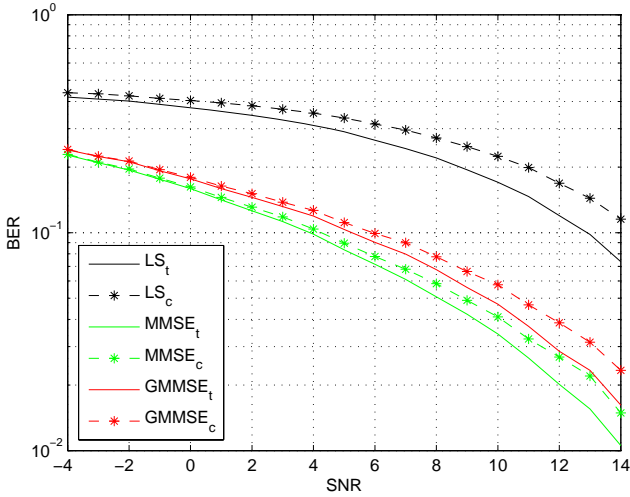


Fig. 2. Performance analysis for Structured LS, MMSE, and GMMSE detectors, $n = 50$ and channel length $L = 7$ (t : Toeplitz case; c : circular case).

We can express the matrix $\tilde{\mathbf{H}}^T \tilde{\mathbf{H}}$ by its eigenvalue decomposition $\tilde{\mathbf{H}}^T \tilde{\mathbf{H}} = \mathbf{F}^T \Lambda \mathbf{F}$, where \mathbf{F} is the discrete Fourier transform matrix (computed by FFT) and $\Lambda = \text{diag}(\mathbf{F} \cdot \tilde{\mathbf{H}}(:, 1))$ in that case problem Eq. (11) can be written as

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}^T \mathbf{x} \leq n} \mathbf{x}^T (\mathbf{F}^T \Lambda \mathbf{F}) \mathbf{x} - 2\mathbf{y}^T \mathbf{x}. \quad (22)$$

The dual problem for problem Eq. (22) is

$$\max_{\lambda \geq 0} -\mathbf{y}^T \left((\mathbf{F}^T \Lambda \mathbf{F}) + \lambda \mathbf{I} \right)^{-1} \mathbf{y} - \lambda n \quad (23)$$

and the gradient descent algorithm in the circular case takes the form

$$\bar{\lambda}(t+1) = \bar{\lambda}(t) + \mu \left(\mathbf{y}^T \mathbf{F}^T (\Lambda + \bar{\lambda}(t) \mathbf{I})^{-2} \mathbf{F} \mathbf{y} - n \right). \quad (24)$$

After getting the optimal value λ^* , the GMMSE solution in the circular case is

$$\mathbf{x}^* = \mathbf{F}^T (\Lambda + \lambda^* \mathbf{I})^{-1} \mathbf{F} \mathbf{y}. \quad (25)$$

Again, besides computing $\mathbf{F} \mathbf{y}$ (IFFT) only diagonal matrices must be inverted in Eqs. (24) and (25). Most important, no EVD computation is required in the circular case, since the EVD of a circular matrix is easily obtained using FFT/IFFT. Therefore, in this case the additional effort (compared to MMSE) is given by the iteration of Eq. (24), i.e. inversions of diagonal matrices and scalar products.

6 Simulation results

The BER performance of the different detectors is discussed. In the simulation we compare the BER performance for LS,

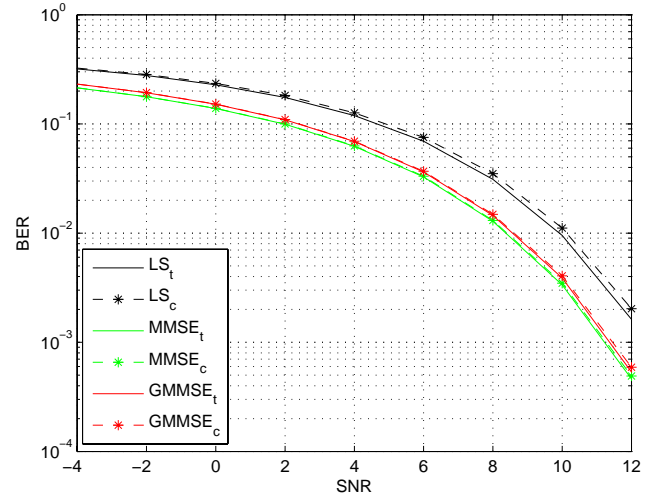


Fig. 3. Performance analysis for Structured LS, MMSE, and GMMSE detectors, $n = 1000$ and channel length $L = 15$ (t : Toeplitz case; c : circular case).

MMSE, and GMMSE detectors, taking into account that we have two different structures, Toeplitz and circular approximation. We applied this simulation for two different simulation scenarios. The equalization problem in the first scenario has a channel impulse response of length $L = 7$ and a transmitted bit vector of length $n = 50$. The second scenario describes the equalization problem with channel impulse response of length $L = 15$ and a transmitted bit vector of length $n = 1000$. Figures (2) and (3) show that GMMSE detector has almost the same performance as MMSE detector but it has the advantage that it does not require the knowledge of σ^2 . Furthermore, we see that the circular approximation only slightly degrades the performance of the detectors.

7 Complexity analysis

The computational complexity of the GMMSE detector is composed of two parts:

(Part 1). The complexity of the solution of the system of equations (Eqs. 16, 20 or 25) which is the same as for LS and MMSE (Eqs. 9 or 10).

(Part 2). The complexity of the iterations required for the gradient descent algorithm (Eqs. 14, 19 or 24).

In part 1, if there is no structure the solution is obtained by the Cholesky algorithm with complexity $n^3/3$. When there is a Toeplitz structure, the solution is given by the Levinson algorithm with complexity $4n^2$ and if we approximate this Toeplitz matrix to a circular structure, the solution is obtained using the FFT decomposition with complexity $3/2(n+L-1)\log_2(n+L-1)$. Therefore, the circular

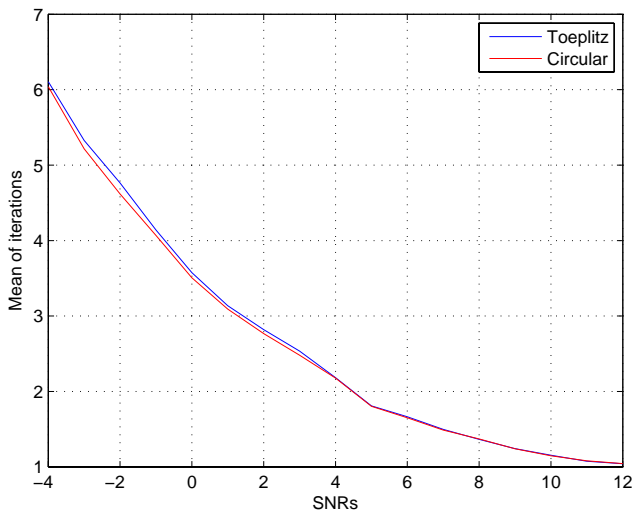


Fig. 4. Iterations for gradient descent algorithm in Toeplitz and circular structure cases for $n = 50$ and channel length $L = 7$.

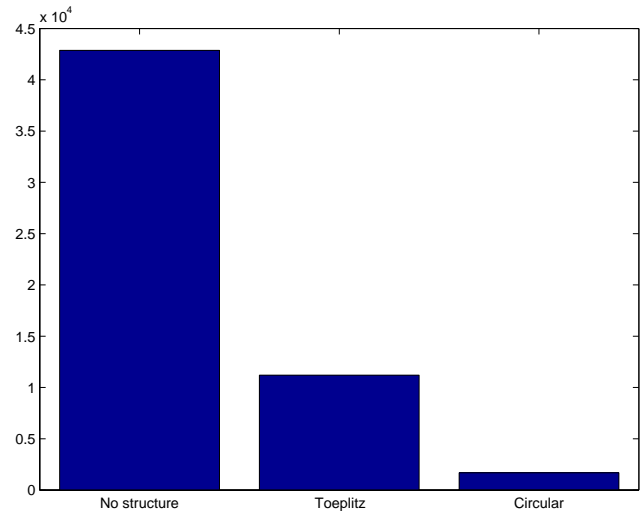


Fig. 6. Computational complexity for $n = 50$ and channel length $L = 7$.

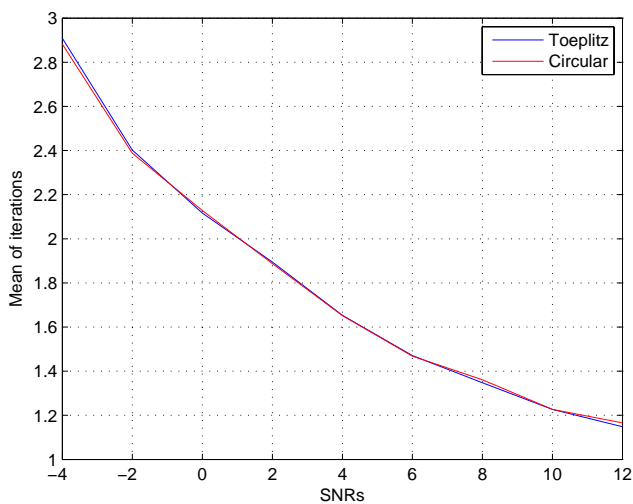


Fig. 5. Iterations for gradient descent algorithm in Toeplitz and circular structure cases for $n = 1000$ and channel length $L = 15$.

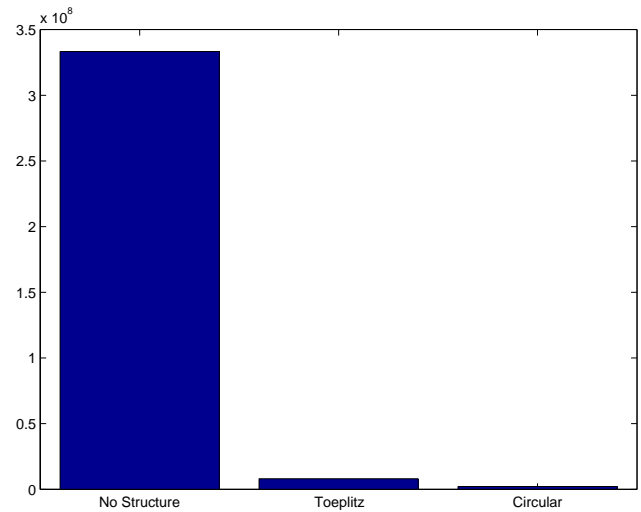


Fig. 7. Computational complexity for $n = 1000$ and channel length $L = 15$.

approximation results in a significantly reduced computational complexity.

In part 2 gradient descent algorithm adds some complexity. However, for the structured cases (Eqs. 19 or 24) the iterations of the gradient descent algorithm are only applied to diagonal matrices (Λ) such that the complexity is only of $O(n)$ per iteration. Figures (4) and (5) show the mean number of required iterations for the Toeplitz case and the circular case in our two scenarios respectively. Obviously, the number of iterations in each scenario for both cases is almost the same. We assume the worst case complexity for part 2 by taking 6 iterations for all cases. Since the required number

of iterations is quite small and the computational complexity is only of $O(n)$ per iteration, the complexity of the gradient descent algorithm is almost negligible compared to part 1.

The overall complexity (part 1 and part 2) for our two cases is shown in Fig. (6) and (7).

8 Conclusions

In this paper, it was shown that the circular approximation of the Toeplitz channel matrix is not only effective to significantly reduce the computational complexity of GMMSE detector using the gradient descent algorithm, but it also keeps

the performance gain compared to LS detector (is almost the same as MMSE) without any requirement to know the noise power value (σ^2).

In future work we will apply the presented technique to various practical problems and evaluate the performance depending on the channel length (L) and the dimension of the transmitted bit vector (n). We will also apply it to some common communication schemes like CDMA and OFDM.

References

- Boyd, S. and Vandenberghe, L.: Convex Optimization. Cambridge Univ. Press, 2004.
- Chang, X.-W. and Han, Q.: Solving box-constrained integer least squares problem, *IEEE Transactions on wireless communications*, 7[1], 277–287, 2008.
- Golub, G. and Loan, C.: Matrix Computations. Johns Hopkins Univ. Press, 3rd edition 1996.
- Hansen, P.: Methods of nonlinear 0–1 programming, In *Annals of Discrete Mathematics 5: Discrete Optimization II*, edited by: Hammer, P. L., Johnson, E. L., and Korte, B. H., Amsterdam, The Netherlands: North Holland, 1979.
- Lupas, R. and Verdu, S.: Linear multiuser detectors for synchronous code-division multiple-access channels, *IEEE Trans. Inform. Theory*, 35, 123–136, 1989.
- Madhow, U. and Honig, M. L.: MMSE interference suppression for direct-sequence spread-spectrum CDMA, *IEEE Trans. Commun.*, 42, 3178–3188, 1994.
- Nash, S. and Sofer, A.: *Linear and Nonlinear Programming*. New York: McGraw-Hill, 1996.
- Tan, P. H. and Rasmussen, L. K.: The application of semidefinite programming for detection in CDMA, *IEEE journal on selected areas in communications*, 19[8], 1442–1449, 2001.
- Verdu, S.: Computational complexity of multiuser detection, *Algorithmica*, 4[4], 303–312, 1989.
- Verdu, S.: *Multiuser Detection*, Cambridge, UK: Cambridge Univ. Press, 1998.
- Vollmer, M., Haardt, M., and Götze, J.: Comparative study of joint-detection techniques for TD-CDMA based mobile radio systems, *IEEE journal on selected areas in communications*, 19[8], 1461–1475, 2001.
- Yener, A., Yates, R., and Ulukus, S.: CDMA multiuser detection: A nonlinear programming approach, *IEEE Trans. Commun.*, 50[6], 1016–1024, 2002.