

# SEMIDEFINITE PROGRAMMING FOR DETECTION IN LINEAR SYSTEMS – OPTIMALITY CONDITIONS AND SPACE-TIME DECODING

*Joakim Jaldén\*, Cristoff Martin, Björn Ottersten*

The Department of Signals, Sensors & Systems  
Royal Institute of Technology (KTH)  
SE-100 44 Stockholm, Sweden

## ABSTRACT

Optimal maximum likelihood detection of finite alphabet symbols in general requires time consuming exhaustive search methods. The computational complexity of such techniques is exponential in the size of the problem and for large problems sub-optimal algorithms are required. In this paper, to find a solution in polynomial time, a semidefinite programming approach is taken to estimate binary symbols in a general linear system. A condition under which the proposed method provides optimal solutions is derived. As an application, the proposed algorithm is used as a decoder for a linear space-time block coding system and the results are illustrated with numerical examples.

## 1. INTRODUCTION

Finding the maximum likelihood (ML) estimates of binary input symbols based on the output of a known linear channel in general requires joint detection of an entire sequence of symbols. On a linear channel under some standard assumptions the ML detection problem can be written as a quadratic optimization problem with integer constraints [1]. Unfortunately this problem is in general non-deterministic polynomial time hard (NP-hard) [2].

Recently, semidefinite (SD) relaxation has been successfully employed to suboptimally solve such detection problems in multiuser detection [3, 4]. These results show that the SD relaxation has superior performance to other well known methods such as the linear-minimum-mean-square-error (LMMSE) and the decorrelating detectors.

In SD relaxation the non-convex ML detection problem is relaxed to a semidefinite program (SDP) which can be efficiently solved in polynomial time [5]. The SD relaxation provides an upper bound on the maximum value of the likelihood function. If this bound is tight the SD relaxation will optimally solve the ML detection problem. If the bound is not tight it will still provide the basis for a heuristic employed to approximate the ML solution.

This paper derives conditions under which the bound is tight. It is shown that under some conditions the probability of solving the ML-detection problem using SD relaxation goes to 1 as the signal to noise ratio (SNR) increase.

The organization of this paper is as follows. In Section 2 the problem formulation is given and SD relaxation is introduced. Section 3 presents the main analytical result of this paper, and Section 4 demonstrates the performance of the proposed method applied as a decoder of a linear space-time block code.

\*jalden@s3.kth.se

## 2. PROBLEM STATEMENT

Let  $\mathcal{B}^n = \{\pm 1\}^n$  be the set of binary  $n$ -tuples. The detection problem considered in this paper is to obtain the ML estimate of  $\mathbf{s} \in \mathcal{B}^n$  given the linear model

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{v} \quad (1)$$

where  $\mathbf{H} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{y} \in \mathbb{R}^m$  and  $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_m)$ . The vector  $\mathbf{y}$  and matrix  $\mathbf{H}$  are assumed to be known. Typically  $\mathbf{y}$  represent a received signal,  $\mathbf{H}$  a channel, and  $\mathbf{v}$  additive white Gaussian noise.

The ML detection problem under the above assumptions may be written as [1]

$$\hat{\mathbf{s}}_{\text{ML}} = \arg \min_{\mathbf{s}} \quad \bar{\mathbf{s}}^T \mathbf{H}^T \mathbf{H} \bar{\mathbf{s}} - 2\mathbf{y}^T \mathbf{H} \bar{\mathbf{s}} \quad (2)$$

s.t.  $\bar{\mathbf{s}} \in \mathcal{B}^n$ .

Unfortunately, in the general case, the computational complexity of solving (2) increase exponentially with  $n$ . A way to efficiently approximate the solution of (2) is to relax the constraints on  $\bar{\mathbf{s}}$ .

Note that it is not a restriction to only consider the real case in (1). In the case of complex data such that  $\mathbf{s} \in \{\pm 1 \pm j\}^n$ ,  $\mathbf{H} \in \mathbb{C}^{m \times n}$  and  $\mathbf{v}$  being an additive white circularly symmetric complex Gaussian noise, the problem may be rewritten in the form of (1) by expanding the problem dimensionality.

### 2.1. Semidefinite Relaxation

Problem (2) can be rewritten as [4]

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{x}} \quad & \text{tr}(\mathbf{L}\mathbf{X}) \\ \text{s.t.} \quad & \text{diag}(\mathbf{X}) = \mathbf{e} \\ & \mathbf{X} = \mathbf{x}\mathbf{x}^T \end{aligned} \quad (3)$$

where  $\mathbf{e}$  is the vector of all ones,

$$\mathbf{L} = \begin{bmatrix} \mathbf{H}^T \mathbf{H} & -\mathbf{H}^T \mathbf{y} \\ -\mathbf{y}^T \mathbf{H} & 0 \end{bmatrix} \text{ and } \mathbf{x} = \begin{bmatrix} \bar{\mathbf{s}} \\ 1 \end{bmatrix}. \quad (4)$$

Due to symmetry, the constraint  $x_{n+1} = 1$  does not have to be explicitly maintained. The SD relaxation of (3) is

$$\begin{aligned} \min_{\mathbf{X}} \quad & \text{tr}(\mathbf{L}\mathbf{X}) \\ \text{s.t.} \quad & \text{diag}(\mathbf{X}) = \mathbf{e} \\ & \mathbf{X} \succeq \mathbf{0} \end{aligned} \quad (5)$$

where  $\mathbf{X} \succeq \mathbf{0}$  means that  $\mathbf{X}$  is symmetric and positive semidefinite. In (5) the non-convex rank 1 constraint of (3) has been relaxed to a positive semidefinite constraint on  $\mathbf{X}$ . Problem (5) is a semidefinite program and standard methods exist to solve it in polynomial time [5]. If the solution,  $\mathbf{X}^*$ , of (5) is of rank 1 then it will also be an optimal solution of (3).

## 2.2. Practical Considerations

When  $\mathbf{X}^*$  is not of rank 1,  $\hat{\mathbf{s}}_{\text{ML}}$  needs to be approximated from  $\mathbf{X}^*$  in some way. In this paper  $\hat{\mathbf{s}}_{\text{ML}}$  is approximated by the sign of the singular vector corresponding to the largest singular value of  $\mathbf{X}^*$ . That is, let  $\mathbf{u}$  be the singular vector corresponding to the largest singular value where  $\mathbf{u}$  is chosen such that  $u_{n+1} \geq 0$ . Then  $\hat{s}_i = 1$  if  $u_i \geq 0$  or  $\hat{s}_i = -1$  otherwise.

There exist schemes [3] that provide better approximations of  $\hat{\mathbf{s}}_{\text{ML}}$ . The simple method presented above, for the examples provided in this paper, performs close to ML and therefore more complicated schemes are not considered in this paper.

## 3. CONDITIONS FOR A RANK 1 SOLUTION

This section presents the main analytical contribution of this paper by giving necessary and sufficient conditions under which problem (5) has a rank 1 solution.

### 3.1. Main Result

**Theorem 1** Given  $\mathbf{s}$ , let the convex set  $\mathcal{V}_s$  be defined as

$$\mathcal{V}_s = \{\mathbf{v} \mid \mathbf{H}^T \mathbf{H} + \mathbf{S}^{-1} \text{Diag}(\mathbf{H}^T \mathbf{v}) \succeq \mathbf{0}\} \quad (6)$$

where  $\mathbf{S} = \text{Diag}(\mathbf{s})$ . The SD relaxation (5) has a rank 1 solution corresponding to  $\mathbf{s}$  if and only if  $\mathbf{v} \in \mathcal{V}_s$ .

**Proof:** An  $\mathbf{X} \succeq \mathbf{0}$  is an optimal solution to (5) if and only if there are  $\mathbf{Z} \succeq \mathbf{0}$  and  $\mathbf{z} \in \mathbb{R}^{n+1}$  such that [6, Ch 4.2]

$$\begin{aligned} \text{diag}(\mathbf{X}) &= \mathbf{e} \\ \mathbf{Z} + \text{Diag}(\mathbf{z}) &= \mathbf{L} \\ \mathbf{XZ} &= \mathbf{0}. \end{aligned} \quad (7)$$

$\mathbf{Z}$  and  $\mathbf{z}$  are the dual variables and (7) are the so called Karush-Kuhn-Tucker conditions for (5).

Assuming that (5) has a rank 1 solution corresponding to  $\mathbf{s}$ , i.e.  $\mathbf{X} = \mathbf{x}\mathbf{x}^T$  for  $\mathbf{x}^T = [\mathbf{s}^T \ 1]$ , (7) implies

$$\mathbf{XZ} = \mathbf{x}\mathbf{x}^T(\mathbf{L} - \text{Diag}(\mathbf{z})) = \mathbf{0}. \quad (8)$$

Multiplying on the left by  $\mathbf{x}^T$  and taking the transpose yields

$$\mathbf{Lx} = \text{Diag}(\mathbf{z})\mathbf{x} \quad (9)$$

and

$$\mathbf{z} = \text{Diag}(\mathbf{x})^{-1} \mathbf{Lx}. \quad (10)$$

The matrix  $\mathbf{Z}$  is thus explicitly given by

$$\mathbf{Z} = \mathbf{L} - \text{Diag}(\mathbf{x})^{-1} \text{Diag}(\mathbf{Lx}). \quad (11)$$

Using the model  $\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{v}$ , the matrix  $\mathbf{L}$  becomes

$$\mathbf{L} = \begin{bmatrix} \mathbf{H}^T \mathbf{H} & -\mathbf{H}^T \mathbf{H}\mathbf{s} - \mathbf{H}^T \mathbf{v} \\ -\mathbf{s}^T \mathbf{H}^T \mathbf{H} - \mathbf{v}^T \mathbf{H} & 0 \end{bmatrix} \quad (12)$$

and using  $\mathbf{x}^T = [\mathbf{s}^T \ 1]$ , the matrix  $\mathbf{Z}$  can be written as

$$\mathbf{Z} = \mathbf{L} - \text{Diag}(\mathbf{x})^{-1} \text{Diag}(\mathbf{Lx}) = \begin{bmatrix} \mathbf{H}^T \mathbf{H} + \mathbf{S}^{-1} \text{Diag}(\mathbf{H}^T \mathbf{v}) & -\mathbf{H}^T \mathbf{H}\mathbf{s} - \mathbf{H}^T \mathbf{v} \\ -\mathbf{s}^T \mathbf{H}^T \mathbf{H} - \mathbf{v}^T \mathbf{H} & \mathbf{s}^T \mathbf{H}^T \mathbf{H}\mathbf{s} + \mathbf{s}^T \mathbf{H}^T \mathbf{v} \end{bmatrix}. \quad (13)$$

Note that  $\mathbf{Z}$  can be written as

$$[\mathbf{I} \ -\mathbf{s}]^T (\mathbf{H}^T \mathbf{H} + \mathbf{S}^{-1} \text{Diag}(\mathbf{H}^T \mathbf{v})) [\mathbf{I} \ -\mathbf{s}]. \quad (14)$$

This implies that  $\mathbf{Z}$  is positive semidefinite if and only if

$$\mathbf{H}^T \mathbf{H} + \mathbf{S}^{-1} \text{Diag}(\mathbf{H}^T \mathbf{v}) \quad (15)$$

is positive semidefinite. Thus for any  $\mathbf{s} \in \mathcal{B}^n$  the optimal solution to the SD relaxation problem will have a rank 1 solution corresponding to  $\mathbf{s}$  if and only if (15) is positive semidefinite. This concludes the proof. ■

**Note:** Corresponding to  $\mathbf{s}$  in Theorem 1 is vital since the SD relaxation may have a rank 1 solution even if  $\mathbf{v} \notin \mathcal{V}_s$ . As stated earlier the SD relaxation estimate equals the ML estimate for all rank 1 solutions but the important distinction here is that in such a case  $\hat{\mathbf{s}}_{\text{ML}} \neq \mathbf{s}$ . In other words, if  $\mathbf{v} \notin \mathcal{V}_s$  and  $\mathbf{X}^* = \mathbf{x}\mathbf{x}^T$  then  $\hat{\mathbf{s}} = \hat{\mathbf{s}}_{\text{ML}}$  where  $\hat{\mathbf{s}}$  is the solution of (5) but  $\hat{\mathbf{s}}_{\text{ML}} \neq \mathbf{s}$ .

### 3.2. Comments

From Theorem 1 it follows that for all  $\mathbf{y} \in \mathcal{Y}$  where

$$\mathcal{Y} = \bigcup_{\mathbf{s} \in \mathcal{B}^n} \mathcal{Y}_s \quad \text{and} \quad \mathcal{Y}_s = \mathbf{H}\mathbf{s} + \mathcal{V}_s \quad (16)$$

the SD relaxation is guaranteed to produce the ML estimate of  $\mathbf{s}$ . The sets  $\mathcal{Y}_s$  will be denoted the rank 1 regions in analogy with the ML decision regions.

When the matrix  $\mathbf{H}$  is of rank  $n$ , the product  $\mathbf{H}^T \mathbf{H}$  will be strictly positive definite and the  $\mathbf{0}$  vector will lie in the interior of  $\mathcal{V}_s$ . This implies that whenever  $\mathbf{H}$  is of rank  $n$  the probability of obtaining the ML estimate by SD relaxation goes to 1 as the SNR increases.

Assuming that  $\mathbf{s} \in \mathcal{B}^n$  was sent, a necessary but not sufficient condition for  $\hat{\mathbf{s}}_{\text{ML}} = \mathbf{s}$  is

$$\mathbf{h}_i^T \mathbf{h}_i + s_i^{-1} \mathbf{h}_i^T \mathbf{v} \geq 0 \quad i \in \{1, \dots, n\}. \quad (17)$$

where  $\mathbf{h}_i$  is the  $i$ th column of  $\mathbf{H}$ . This condition is easily obtained by stating that for  $\hat{\mathbf{s}}_{\text{ML}} = \mathbf{s}$  the sent vector  $\mathbf{s}$  should have a lower objective value in (2) than  $\mathbf{s}' \in \mathcal{B}^n$  where  $\mathbf{s}' = \mathbf{s} - 2s_i \mathbf{e}_i$  and  $\mathbf{e}_i$  is the  $i$ th unit vector. If  $\mathbf{H}$  has orthogonal columns then  $\mathbf{H}^T \mathbf{H}$  is diagonal and  $\mathbf{v} \in \mathcal{V}_s$  reduces to (17). Thus for any  $\mathbf{H}$  with orthogonal columns SD relaxation always obtains the ML estimate. Of course, in such a case ML decoding is already a trivial problem.

Fig. 1 shows the rank 1 regions as a function of the received vector  $\mathbf{y}$  for a simple example where

$$\mathbf{H} = \begin{bmatrix} 1 & 0.2 \\ 0.2 & 1 \end{bmatrix}. \quad (18)$$

The four white regions show  $\mathcal{Y}_s$  for all different values of  $\mathbf{s} \in \mathcal{B}^2$ . The boundary of the ML decision regions are shown by the bold lines. From this picture it is clear how the rank 1 regions approximate the ML decision regions. Note that  $\mathcal{Y}_s$  are subsets of the ML decision regions. This is an obvious consequence of the SD relaxation in Section 2.1. Also note that the rank 1 regions are only dependent on  $\mathbf{H}$  and not on the SNR.

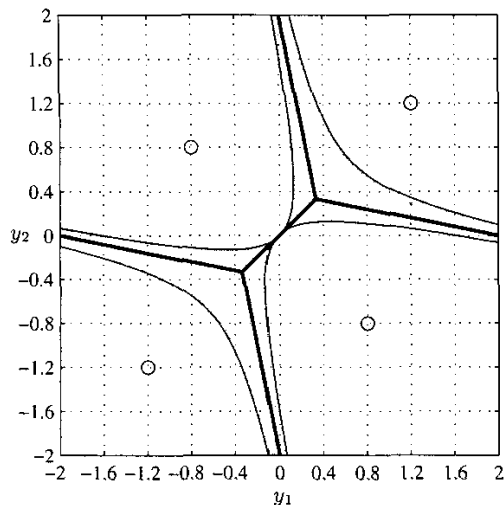


Fig. 1. The rank 1 regions for a simple 2 dimensional example are shown as light regions. The noise free points,  $\mathbf{H}\mathbf{s}$ , are shown as small circles and the boundary of the ML decision regions are shown by bold lines.

#### 4. APPLICATIONS

The model (1) encompass several practical detection problems, especially in a communication context. For example, the problem of multi-user detection in CDMA may be considered with the aforementioned model. Semidefinite programming, as a tool to solve this problem for the synchronous case has appeared in communication literature [3, 4]. In this section we show that the techniques presented in Section 2 may also be used to address the problem of decoding linear space-time block codes (STBC). The performance of the proposed technique in this application is illustrated with numerical results from simulations.

##### 4.1. Linear STBC Decoding

Consider a space-time block coding system consisting of an  $M$ -antenna transmitter and an  $N$ -antenna receiver. The code words are of length  $L$  in time and we assume a narrowband channel between the transmitter and receiver. Such systems are commonly modeled as

$$\mathbf{Y} = \mathbf{G}\mathbf{C} + \mathbf{V}, \quad (19)$$

where  $\mathbf{Y} \in \mathbb{C}^{N \times L}$  is the received data, the elements of  $\mathbf{G} \in \mathbb{C}^{N \times M}$  model the phase and attenuation of the channels between the transmit and receive antennas,  $\mathbf{C} \in \mathbb{C}^{M \times L}$  is the space-time code word transmitted and  $\mathbf{V} \in \mathbb{C}^{N \times L}$  is additive white circularly symmetric complex Gaussian noise. Here we consider a linear space-time block coding [7] scheme where the code words are formed as

$$\mathbf{C} = \sum_{q=1}^Q \mathbf{A}_q s_q. \quad (20)$$

The  $\mathbf{A}_q$  matrices define the code and are known to the receiver while  $s_q \in \{\pm 1\}$  are the data bits to be transmitted. Thus, from (20)  $2^Q$  different code words  $\mathbf{C}$  can be formed.

By stacking the columns of the output data matrix  $\mathbf{Y}$  in a vector  $\mathbf{y}$  it is obvious that the problem is of the form (1),

$$\begin{aligned} \mathbf{y} &= \text{vec } \mathbf{Y} \\ &= (\mathbf{I}_L \otimes \mathbf{G}) \text{vec} \left\{ \sum_{q=1}^Q \mathbf{A}_q s_q \right\} + \text{vec } \mathbf{V} \\ &= (\mathbf{I}_L \otimes \mathbf{G}) \mathbf{A} \mathbf{s} + \mathbf{v} \\ &= \mathbf{H} \mathbf{s} + \mathbf{v}. \end{aligned} \quad (21)$$

Here,  $\mathbf{A} = [\text{vec } \mathbf{A}_1 \dots \text{vec } \mathbf{A}_Q]$ ,  $\mathbf{s} = [s_1 \dots s_Q]^T$ ,  $\mathbf{v} = \text{vec } \mathbf{V}$ ,  $\mathbf{H} = (\mathbf{I}_L \otimes \mathbf{G}) \mathbf{A}$  and  $\otimes$  denotes the Kronecker product.

Unless the structure of the code provides means for efficient decoding which may only be practical for certain code rates, the computational complexity of a maximum likelihood decoder increases exponentially with  $Q$ . In order to render the decoding problem feasible, techniques such as sphere decoding [8] or sub-optimal methods such as MMSE or the SDP relaxation presented in this paper need to be considered.

##### 4.2. Numerical Example

To illustrate the performance achievable for a linear STBC setup employing the detector described in the above sections, a numerical example is studied. In the example, an  $M = 8$  antenna transmitter is communicating with an  $N = 2$  antenna receiver using a space-time code of length  $L = 8$ . The entries of  $\mathbf{G}$  are drawn independently from a circularly symmetric complex Gaussian distribution. The codes used in this example have been designed using a scheme described in [9] where the code matrices are designed to minimize an upper bound on the resulting error probability of the code. Here, the design algorithm has been modified to handle the linear structure of (20), which is also mentioned in [9]. Fig. 2 and Table 1 show how various parameters are affected by the number of information bits per code word,  $Q$ , and the signal to noise ratio. We define the SNR as

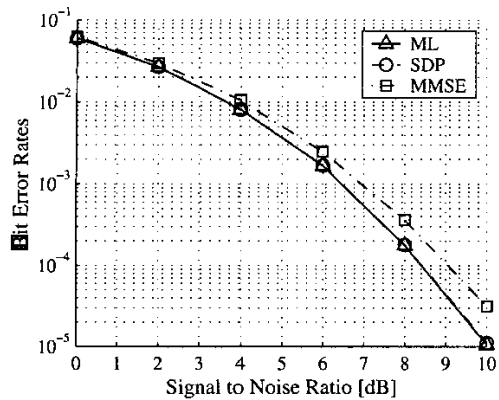
$$\text{SNR} = \frac{E\{\|\mathbf{G}\mathbf{C}\|_F^2\}}{E\{\|\mathbf{V}\|_F^2\}}. \quad (22)$$

In Fig. 2(a) and 2(b) the bit error rate of the semidefinite decoder is studied as a function of the signal to noise ratio and bits per space-time code word,  $Q$ . Results from an MMSE- and an ML-decoder have been included for reference. In these cases, the performance of the SDP decoder is virtually identical with that of the optimal ML decoder. As a rough illustration of the computational complexity of the various methods used, Fig. 2(c) shows the computation time for the various detectors at different  $Q$ . While the execution time is high for the SDP detector at small  $Q$ , as  $Q$  increases the complexity of the ML detector increases much faster.

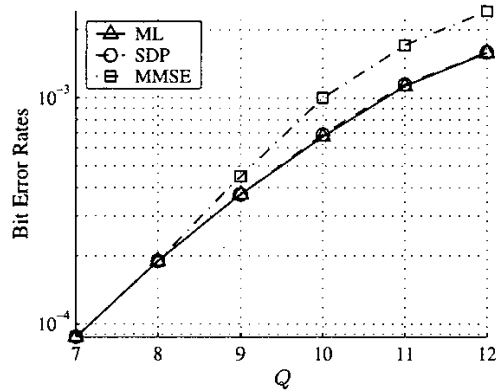
SNR = 6 dB, $Q$ :	7	8	9	10	11	12
Part:	1	1	0.99	0.98	0.96	0.93
$Q = 12$ , SNR [dB]:	0	2	4	6	8	10
Part:	0.57	0.69	0.83	0.93	0.98	1

Table 1. Part of realizations resulting in a rank 1 solution.

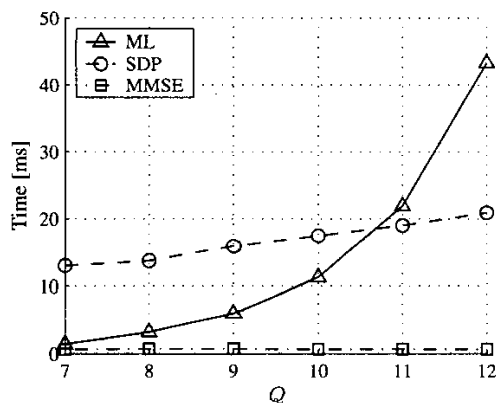
Finally, in Table 1, the rank of the SDP solution is studied for various  $Q$  and signal to noise ratios. Note that for  $Q \leq 8$  all the found SDP solutions are rank 1 and thus the ML-solution and SDP



(a) Bit Error Rates,  $Q = 12$



(b) Bit Error Rates, SNR = 6 dB



(c) Rough Time Estimates, SNR = 6 dB

**Fig. 2.** Performance of the SDP relaxation when used as a decoder for linear STBC,  $M = 8$ ,  $N = 2$  and  $L = 8$ .

solutions are identical. This should be expected as the code words used provide orthogonal  $\mathbf{H}$  for  $Q \leq 8$ , see Section 3.2. The part of the simulation samples that result in rank 1 solutions is largely dependent on the structure of  $\mathbf{H}$  and from simulations we have noticed that this part may be significantly lower in for example multi-user ML-sequence detection experiments. In these cases it may also be necessary to employ more advanced techniques if the solution found,  $\mathbf{X}^*$ , is of rank higher than 1, see e.g. [3].

## 5. CONCLUSIONS

In this paper we have presented semidefinite relaxation as a method for detection of binary data on a general linear channel. Conditions, under which the semidefinite relaxation method equals maximum likelihood detection, have been derived. We have shown that analogous to ML detection regions, we can define rank 1 regions in which the semidefinite relaxation is guaranteed to produce the ML estimate.

In simulations the performance of the SDP decoder has been shown to be virtually identical with the ML decoder in the case of linear space-time block codes. This is true even in the case where the aforementioned conditions do not always hold.

## ACKNOWLEDGMENTS

We would like to thank our colleague George Jöngren for enlightening discussions around space-time codes as well as the design of the actual codes used in Section 4.2.

## 6. REFERENCES

- [1] T. Kailath and H. V. Poor, "Detection of stochastic processes," *IEEE Transactions on Information Theory*, vol. 44, no. 6, Oct. 1998.
- [2] S. Verdú, "Computational complexity of multiuser detection," *Algorithmica*, vol. 4, pp. 303–312, 1989.
- [3] W.-K. Ma, T. N. Davidson, K. M. Wong, Z.-Q. Lou, and P.-C. Ching, "Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous CDMA," *IEEE Transactions on Signal Processing*, vol. 50, no. 4, Apr. 2002.
- [4] P. H. Tan and L. K. Rasmussen, "The application of semidefinite programming for detection in CDMA," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, Aug. 2001.
- [5] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, pp. 49–95, 1996.
- [6] H. Wolkowicz, R. Saigal, and L. Vandenberghe, Eds., *Handbook of Semidefinite Programming*, Kluwer Academic Publishers, 2000.
- [7] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1804–1824, June 2002.
- [8] H. Vikalo and B. Hassibi, "Maximum-likelihood sequence detection of multiple antenna systems over dispersive channels via sphere decoding," *EURASIP Journal on Applied Signal Processing*, vol. 2002, no. 5, pp. 525–531, May 2002.
- [9] G. Jöngren, M. Skoglund, and B. Ottersten, "Utilizing partial channel information in the design of space-time block codes," in *Proceedings of the 5th International Symposium on Wireless Personal Multimedia Communications*, Oct. 2002.