ROBUST SEMI-DEFINITE RELAXATION MIMO DETECTION IN A NON-GAUSSIAN CHANNEL

Jakob Vovnoboy, Ami Wiesel

Department of Computer Science
The Hebrew University of Jerusalem, Israel

Wing-Kin Ma

Department of Electrical Engineering
The Chinese University of Hong Kong, China

ABSTRACT

Semi-definite relaxation (SDR) is a popular technique for Multi-Input Multi-Output (MIMO) detection. For Binary Phase-Shift Keying (BPSK) and Quadratic Phase-Shift Keying (QPSK), it has been found that SDR can provide a near-optimal Bit Error rate (BER) performance in a Gaussian channel. However if the noise in the channel deviates from the Gaussian model, as it does in many real wireless channels, BER performance drops considerably. In this paper we show that SDR can be applied for detection in a non-Gaussian channel using Huber's M-estimation method for robust regression.

1. INTRODUCTION

In recent years, SDR of the Gaussian Maximum Likelihood (ML) function was used extensively in detection of integer symbols transmitted over Additive White Gaussian Noise (AWGN) MIMO channels [1, 2, 3]. For BPSK it was shown by simulations that SDR gives near-optimal results [4] and later was proven to achieve maximal possible diversity [3].The SDR technique was also extended to higher order Quadrature Amplitude Modulation (QAM) [5, 6].

However, the above is true for AWGN channels only. In many physical channels, mainly due to impulsive nature phenomenons, the ambient noise tends to be non-Gaussian. This was experimentally shown in urban and indoor radio channels [7, 8] as well as in underwater acoustic channels [9]. In such channels performance of the Gaussian ML detector, and with it the SDR of it, degrades considerably even when the deviation is relatively small [10].

A technique based on M-estimation method for robust statistics (i.e. Huber's estimator) is presented in [11] which guarantees that for small unknown deviations from the Gaussian distribution it is the best estimator (in the minimax sense) that can be achieved [12] and gives a robust estimator for non-Gaussian channels. Unfortunately, in the case of integer constrained MIMO the complexity of this technique becomes Non-Polynomial (NP) in the general case. We present several standard methods with polynomial complexity which relax the above and suggest a further improvement by using Semi-definite Relaxed Huber (SDRH) detector. We also show that it is a tighter relaxation of the NP problem in binary constrained MIMO compared to other standard relaxation techniques. In this paper we consider only the case where the transmitted data and channel have real values, but the analysis can be easily extended to the complex case [13].

This paper is organized as follows. Section 2 presents the BPSK MIMO system model with non-Gaussian channel noise and the Huber's robust detector. In Section 3 we present several standard relaxations and derive the SDRH Detector using an alternative form of the Huber's robust detector. In section 4 we prove that in terms of objective value the SDRH is tighter than all the other presented detectors. Section 5 provides simulation results demonstrating the performance gain of the SDRH over the other detectors in a non-Gaussian channel.

The following notation is used. The sets $\mathbb{R}^n$and $\mathbb{R}^{n\times m}$ denote the set of length n vectors and the set of size $n \times m$ matrices. the operator $\|\cdot\|_p$ denotes the L_p norm. The superscript $\boldsymbol{X}^T$ and $\boldsymbol{X}^{-1}$ denotes the transpose and inverse operations. The subscripts $\boldsymbol{x}_i$ and $\boldsymbol{X}_{i,j}$ denotes the i's element in the vector$\boldsymbol{x}$ and the element in row i and column j in matrix $\boldsymbol{X}$. The vectors $\mathbf{1}$ and $\mathbf{0}$ are the all ones and all zeros vectors. The vector $\mathrm{diag}\left(\boldsymbol{X}\right)$ contains the elements of the main diagonal of $\boldsymbol{X}$. We donate the zero mean multivariate Gaussian distribution by $\mathcal{N}\left(0,\, \boldsymbol{\Sigma}\right)$ where $\boldsymbol{\Sigma}$ is the covariance matrix. The set$\{\ldots\}^n$denotes the set of length n vectors where each element belongs to $\{\ldots\}$. The operator $Var\left(\mathcal{G}\right)$ is the variance of distribution $\mathcal{G}$. The operators $\leq$ when applied to vectors and $\preceq$ are an inequality in each element and an inequality in the sense of positive semi-definiteness.

2. PROBLEM FORMULATION

We consider the detection of binary symbols transmitted over an $n \times m$ MIMO channel modeled according to

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{v}, \tag{1}$$

where $\boldsymbol{s} \in \mathcal{B}^m \triangleq \{\pm 1\}^m$ is the binary transmitted message, $\boldsymbol{H} \in \mathbb{R}^{n\times m}$ is the channel matrix (which is assumed to be known to the receiver), $\boldsymbol{y} \in \mathbb{R}^n$ is the received message, $\boldsymbol{v} \in \mathbb{R}^n$ is the additive noise which is white but non-Gaussian and

$m \leq n$. In this paper we restrict ourselves to *ϵ-contaminated Gaussian models* so that the distribution of each noise element $\boldsymbol{v}_i$ is given by

$$\mathcal{P}_i = (1-\epsilon)\,\mathcal{N}\left(0,\sigma^2\right) + \epsilon\mathcal{G}_i \tag{2}$$

where $0 < \epsilon < 1$ is fixed, σ^2 is the variance of the nominal Gaussian distribution and $\mathcal{G}_i$ are unknown symmetric distributions which usually represents impulsive phenomenons and outliers ($Var\left(\mathcal{G}_i\right) \gg \sigma^2$).

The ML estimate for $\boldsymbol{s}$ cannot be formulated directly as $\mathcal{G}_i$'s are not known. However, in [11] it is shown that a robust estimation can be achieved using Huber's minimax M-estimator with the following decorrelator

$$\hat{\boldsymbol{s}}^{BCH} = \begin{array}{ll} \underset{\boldsymbol{s}}{\operatorname{argmin}} & \sum_{i=1}^{n} \rho_h\left(\boldsymbol{y}_i - \boldsymbol{H}_{i,\rightarrow}\boldsymbol{s}\right) \\ s.t. & \boldsymbol{s} \in \mathcal{B}^m \end{array} \tag{3}$$

where $\boldsymbol{H}_{i,\rightarrow}$ is the ith row in $\boldsymbol{H}$ and ρ_h is the Huber penalty function

$$\rho_h\left(x\right) = \begin{cases} x^2 & |x| \leq h \\ 2h\,|x| - h^2 & |x| > h \end{cases} \tag{4}$$

where the parameter h is calculated from the following formula

$$\frac{\sigma}{h}\phi\left(\frac{h}{\sigma}\right) - Q\left(\frac{h}{\sigma}\right) = \frac{\epsilon}{2\left(1-\epsilon\right)} \tag{5}$$

where $\phi\left(x\right) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ and $Q\left(t\right) = \frac{1}{\sqrt{2\pi}}\int_t^{\infty} e^{-\frac{x^2}{2}}dx$.

But, in the case of binary symbol transmission, the problem in (3) have NP complexity in the general case and cannot be solved efficiently for large MIMO.

3. THE SDRH DETECTOR

In this section we present several standard methods to approximate (3) and suggest a tighter approximation which can still be calculated efficiently.

One method to approximate (3) is to ignore the constraints on $\boldsymbol{x}$ which leaves us with the non-constrained Huber (NCH) detector, as in [11],

$$\tilde{\boldsymbol{s}} = \underset{\boldsymbol{s}}{\operatorname{argmin}}\, f_{NCH} \triangleq \sum_{i=1}^{n} \rho_h\left(\boldsymbol{y}_i - \boldsymbol{H}_{i,\rightarrow}\boldsymbol{s}\right). \tag{6}$$

It can be calculated efficiently, using standard methods of convex optimization [14]. Then round the result i.e.

$$\hat{\boldsymbol{s}}^{NCH} = \operatorname{sign}\left(\tilde{\boldsymbol{s}}\right) \tag{7}$$

to fit it into the original constrains. However, this relaxation is too loose and does not give optimal results. A tighter result can be achieved by the Hypercube constrained Huber (HCH) detector which is computed by solving (6) with the constraint $\boldsymbol{s} \in [-1,\ 1]^m$ giving

$$\tilde{\boldsymbol{s}} = \begin{array}{ll} \underset{\boldsymbol{s}}{\operatorname{argmin}} & f_{HCH} \triangleq \sum_{i=1}^{n} \rho_h\left(\boldsymbol{y}_i - \boldsymbol{H}_{i,\rightarrow}\boldsymbol{s}\right) \\ s.t. & \boldsymbol{s} \in [-1,\ 1]^m \end{array} \tag{8}$$

and $\hat{\boldsymbol{s}}^{HCH} = \operatorname{sign}\left(\tilde{\boldsymbol{s}}\right)$. But the it is still too loose and does not give much improvement over the unconstrained case. We suggest a tighter relaxation based on an SDR technique which shows a major improvement over the previos two relaxations.

SDR is usually applied to quadratically constrained quadratic programs (QCQP) [16]. In order to transform (3) into a QCQP problem we use the following lemma from [15].

Lemma 1. Let $\boldsymbol{z} \in \mathbb{R}^n$. It holds true that

$$\sum_{i=1}^{n} \rho_h\left(\boldsymbol{z}\right) = \min_{\boldsymbol{u}\in\mathbb{R}^n} \quad \|\boldsymbol{z} - \boldsymbol{u}\|_2^2 + 2h\,\|\boldsymbol{u}\|_1. \tag{9}$$

From the lemma it can be deduced that (3) is equivalent to

$$\begin{array}{ll} \underset{\boldsymbol{s},\boldsymbol{u}}{\min} & \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s} - \boldsymbol{u}\|_2^2 + 2h\,\|\boldsymbol{u}\|_1 \\ \text{s.t.} & \boldsymbol{s} \in \mathcal{B}^m,\ \boldsymbol{u} \in \mathbb{R}^n \end{array} \tag{10}$$

in the sense that if $\boldsymbol{s}^*, \boldsymbol{u}^*$ minimizes (10) then $\boldsymbol{s}^*$ minimizes (3). Now, we can use the SDR technique to relax the above problem. First, it can be seen that (10) is equivalent to

$$\begin{array}{ll} \underset{\boldsymbol{S},\boldsymbol{Q},\boldsymbol{U},\boldsymbol{s},\boldsymbol{u}}{\min} & f_{BCH} \triangleq Tr\left(\boldsymbol{L}\begin{pmatrix} \boldsymbol{S} & \boldsymbol{Q} & \boldsymbol{s} \\ \boldsymbol{Q^T} & \boldsymbol{U} & \boldsymbol{u} \\ \boldsymbol{s^T} & \boldsymbol{u^T} & \boldsymbol{1} \end{pmatrix}\right) + 2h\,\|\boldsymbol{u}\|_1 \\ \text{s.t.} & \operatorname{diag}\left(\boldsymbol{S}\right) = \boldsymbol{1} \\ & \begin{pmatrix} \boldsymbol{S} & \boldsymbol{Q} \\ \boldsymbol{Q}^T & \boldsymbol{U} \end{pmatrix} = \begin{pmatrix} \mathrm{s} \\ \boldsymbol{u} \end{pmatrix}\begin{pmatrix} \mathrm{s} \\ \boldsymbol{u} \end{pmatrix}^T \end{array} \tag{11}$$

where

$$\boldsymbol{L} \triangleq \begin{pmatrix} \boldsymbol{H}^T\boldsymbol{H} & \boldsymbol{H}^T & -\boldsymbol{H}^T\boldsymbol{y} \\ \boldsymbol{H} & \boldsymbol{I} & -\boldsymbol{y} \\ -\boldsymbol{y}^T\boldsymbol{H} & -\boldsymbol{y}^T & \boldsymbol{y^T}\boldsymbol{y} \end{pmatrix}. \tag{12}$$

Then, the last constraint can be relaxed to a convex one giving

$$\begin{array}{ll} \underset{\boldsymbol{S},\boldsymbol{Q},\boldsymbol{U},\boldsymbol{s},\boldsymbol{u}}{\min} & f_{BCH} \triangleq Tr\left(\boldsymbol{L}\begin{pmatrix} \boldsymbol{S} & \boldsymbol{Q} & \boldsymbol{s} \\ \boldsymbol{Q^T} & \boldsymbol{U} & \boldsymbol{u} \\ \boldsymbol{s^T} & \boldsymbol{u^T} & \boldsymbol{1} \end{pmatrix}\right) + 2h\,\|\boldsymbol{u}\|_1 \\ \text{s.t.} & \operatorname{diag}\left(\boldsymbol{S}\right) = \boldsymbol{1} \\ & \begin{pmatrix} \boldsymbol{S} & \boldsymbol{Q} \\ \boldsymbol{Q}^T & \boldsymbol{U} \end{pmatrix} \succeq \begin{pmatrix} \mathrm{s} \\ \boldsymbol{u} \end{pmatrix}\begin{pmatrix} \mathrm{s} \\ \boldsymbol{u} \end{pmatrix}^T \end{array} \tag{13}$$

which is a convex problem. To calculate the computational efficiency of (13), it can be transformed to a standard SDP problem. First a slack variable $\boldsymbol{t} \in \mathbb{R}^n$ is introduced to (13) replacing $\|\boldsymbol{u}\|_1$in the objective by $\boldsymbol{1}^T\boldsymbol{t}$ and adding the constraint $-\boldsymbol{t} \leq \boldsymbol{u} \leq \boldsymbol{t}$. Then, using Schur's lemma [14] and because $\boldsymbol{t} \geq 0$ there exists a positive semi-definite matrix $\boldsymbol{T} \in \mathbb{R}^{n\times n}$ such that $\mathrm{diag}(\boldsymbol{T}) = \boldsymbol{t}$ transforming the problem to a standard SDP problem which can be solved in $O\left((m+2n+1)^{4.5}\right)$ calculations [16]. As $m \leq n$, we conclude that the computational complexity of (13) is $O\left(n^{4.5}\right)$.

Several heuristics can be used to retrieve a binary estimator from the above problem. In this paper we use the Gaussian Randomization which proved to be very effective in the Gaussian case [16]. A vector $\boldsymbol{z}$ is randomly chosen from $\mathcal{N}(0,\boldsymbol{\Sigma})$ where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{X} & \boldsymbol{s} \\ \boldsymbol{s}^T & 1 \end{pmatrix}. \tag{14}$$

Then, the first m components of $\boldsymbol{z}$ are divided by the last component and a sign is taken

$$\hat{\boldsymbol{s}}^{SDRH} = \mathrm{sign}\left(\frac{1}{\boldsymbol{z}_{m+1}}\begin{pmatrix} \boldsymbol{z}_1 & \boldsymbol{z}_2 & .... & \boldsymbol{z}_m \end{pmatrix}^T\right) \tag{15}$$

this process is repeated 50 times and the best estimation is taken.

4. TIGHTNESS OF THE OBJECTIVE VALUE OF SDRH DETECTOR

In this section we prove that the objective value of the SDRH is between the the objective value of the Hypercube constrained Huber (HCH) detector and the objective value of the Binary Constrained Huber detector (BCH). i.e.

$$\min f_{BCH} \geq \min f_{SDRH} \geq \min f_{HCH} \geq \min f_{NCH}. \tag{16}$$

This does not prove that the BER of SDRH is smaller than the BER of BCH but it is a good heuristic to assess that this is correct. In the next section we show by simulations that the BER of SDRH is substantially better then the BER of BCH for high signal to noise ratio (SNR).

The left and right inequalities in (16) are trivial and we focus on the inner inequality. For this purpose, first, let us observe that

$$\begin{pmatrix} \boldsymbol{H}^T\boldsymbol{H} & \boldsymbol{H}^T \\ \boldsymbol{H} & \boldsymbol{I} \end{pmatrix} \succeq 0. \tag{17}$$

This can be proven by Schur's lemma as $\boldsymbol{I} \succ 0$ and $\boldsymbol{H}^T\boldsymbol{H} - \boldsymbol{H}^T\boldsymbol{I}^{-1}\boldsymbol{H} = 0 \succeq 0$. Then, by rewriting the last constraint in (13) we get

$$\begin{pmatrix} \boldsymbol{S} & \boldsymbol{Q} \\ \boldsymbol{Q}^T & \boldsymbol{U} \end{pmatrix} - \begin{pmatrix} \mathrm{s} \\ \boldsymbol{u} \end{pmatrix}\begin{pmatrix} \mathrm{s} \\ \boldsymbol{u} \end{pmatrix}^T \succeq 0. \tag{18}$$

Using $Tr(\boldsymbol{AB}) \geq 0$ for $\boldsymbol{A} \succeq 0$ and $\boldsymbol{B} \succeq 0$, we obtain

$$\begin{gathered} Tr\left\{\begin{pmatrix} \boldsymbol{H}^T\boldsymbol{H} & \boldsymbol{H}^T \\ \boldsymbol{H} & \boldsymbol{I} \end{pmatrix}\begin{pmatrix} \boldsymbol{S} & \boldsymbol{Q} \\ \boldsymbol{Q}^T & \boldsymbol{U} \end{pmatrix}\right\} \geq \\ Tr\left\{\begin{pmatrix} \boldsymbol{H}^T\boldsymbol{H} & \boldsymbol{H}^T \\ \boldsymbol{H} & \boldsymbol{I} \end{pmatrix}\begin{pmatrix} \mathrm{s} \\ \boldsymbol{u} \end{pmatrix}\begin{pmatrix} \mathrm{s} \\ \boldsymbol{u} \end{pmatrix}^T\right\} \end{gathered} \tag{19}$$

and subsequently using Lemma 1 it can be seen that

$$f_{SDRH}(\boldsymbol{S},\boldsymbol{Q},\boldsymbol{U},\boldsymbol{s},\boldsymbol{u}) \geq \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}-\boldsymbol{u}\|_2^2 + 2h\|\boldsymbol{u}\|_1 \geq f_{BCH}(\boldsymbol{s}) \tag{20}$$

for each feasible point in (13). It remains to prove that $\boldsymbol{s}$ taken from the minimal point in (13) is a feasible point in (8) . Again, using Schur's lemma and the last constraint in (13) it can be deduced that $\boldsymbol{S} \succeq \boldsymbol{s}\boldsymbol{s}^T$. Now, from the definition of positive semi-definiteness and the second constraint in (13) we can see that

$$\boldsymbol{e}_i^T\boldsymbol{S}\boldsymbol{e}_i \geq \boldsymbol{e}_i^T\boldsymbol{s}\boldsymbol{s}^T\boldsymbol{e}_i \Longrightarrow 1 \geq \boldsymbol{s}_i^2$$

where $\boldsymbol{e}_i$ a length m unit vector. So, $\boldsymbol{s}$ taken from the minimal point of (13) is a feasible point in (8) thus concluding the proof.

5. SIMULATION RESULTS

For simulation purposes we use a *two-term Gaussian* distribution: $\mathcal{G}_i = \mathcal{N}\left(0,\kappa\sigma^2\right)$ with $\kappa = 100$ and $\epsilon = 0.01$ which serves as an approximation to Middleton Class A noise model [17]. The elements of $\boldsymbol{H}$ are chosen randomly from $\mathcal{N}\left(0,\frac{1}{mn}\right)$ and are Independent and identically distributed (i.i.d).

Fig. 1 shows the performance of SDRH detector ($\hat{\boldsymbol{s}}^{SDRH}$) in comparison with the non-Constrained Huber (NCH) detector ($\hat{\boldsymbol{s}}^{NCH}$) , HCH detector ($\hat{\boldsymbol{s}}^{HCH}$), the Binary Constrained Huber (BCH) detector ($\hat{\boldsymbol{s}}^{BCH}$) and the Binary Constrained Gaussian (BCG) detector. It can be seen that while all Huber's detectors (NCH, HCH and SDRH) are better than the best Gaussian estimator (BCG), SDRH is much closer to BCH than the others.

To demonstrate further the performance of SDRH, simulations were done on a bigger system where brute force search is intractable. Fig. 2 shows a performance comparison between the three Huber's detectors (NCH, HCH and SDRH) and the Semi-definite Relaxed Gaussian (SDRG) detector for $m = 20,\ n = 60$.

6. CONCLUSIONS

In this paper we developed a robust SDR detector for binary constrained MIMO in weakly non-Gaussian channels with maximal computational complexity of $O\left(n^{4.5}\right)$. We proved that in terms of objective value this detector is tighter than standard relaxation techniques of the binary constrained Huber's detector. We also showed experimentally that it is substantially better than Gaussian detectors and Huber's detector standard relaxation techniques of the binary constrained Huber's detector.

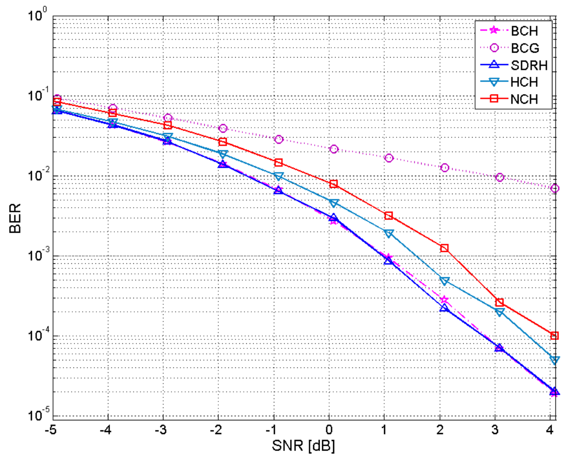

Fig. 1. *Bit Error Ratio V.S. SNR for* $m = 6,\ n = 31$.

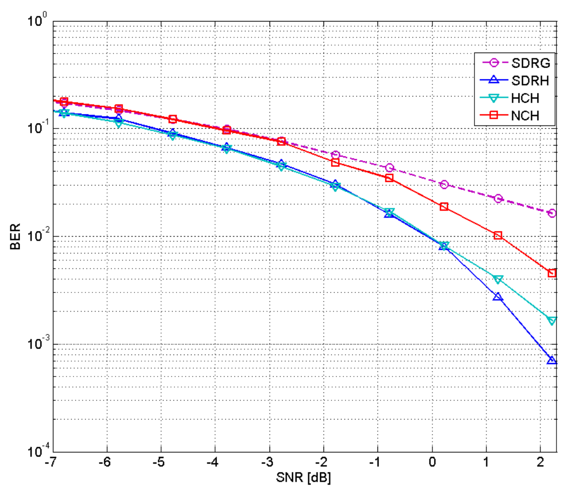

Fig. 2. Bit Error Ratio V.S. SNR for $m = 20,\ n = 60$.

7. REFERENCES

[1] W.-K. Ma, T. N. Davidson, K.M. Wong, Z.-Q Luo, 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, pp. 912–922, Apr. 2002.

[2] 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, pp. 1442–1449, Aug. 2001.

[3] J. Jalden, B. Oterson, "The Diversity Order of the Semidefinite Relaxation Detector" IEEE Transactions on Signal Processing, vol. 54, no. 4, pp. 1405-1422, Apr. 2008.

[4] X. M. Wang, W.-S. Lu, and A. Antoniou, "A near-optimal multiuser detector for DS-CDMA systems using semidefinite programming relaxation," IEEE Trans. Signal Process., vol. 51, no. 9, pp. 2446–2450, Sep. 2003.

[5] A. Wiesel, Y. C. Eldar, and S. Shamai, "Semi-definite relaxation for detection of 16-QAM signaling in MIMO channels," IEEE Signal Process. Lett., vol. 12, no. 9, pp. 653–656, Nov. 2005.

[6] W.-K. Ma, C. -C. Su, J. Jalden, T.-H. Chang, C.-Y. Chi, "The Equivalence of Semi-definite Relaxation MIMO Detectors for Higher-Order QAM" IEEE Journal on Selected Topics In Signal Processing, Vol. 3, no. 6 pp. 1038-1052, Dec. 2009

[7] K. L. Blackard, T. S. Rappaport, and C. W. Bostian, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," IEEE J. Select. Areas Commun., vol. 11, pp. 991–1001, Sept. 1993.

[8] D. Middleton, "Man-made noise in urban environments and transportation systems: Models and measurement," IEEE Trans. Commun., vol. COMM-21, pp. 1232–1241, 1973.

[9] P. L. Brockett, M. Hinich, and G. R. Wilson, "Nonlinear and non- Gaussian ocean noise," J. Acoust. Soc. Amer., vol. 82, pp. 1286–1399, 1987.

[10] H. V. Poor and M. Tanda, "An analysis of some multiuser detectors in impulsive noise," in Proc. 16th GRETSI Symp. Signal Image Process., Grenoble, France, Sept. 1997.

[11] X. Wang, H. V. Poor "Robust Multiuser Detection in Non-Gaussian Channels" IEEE Transactions on Signal Processing, vol. 47, no. 2, pp. 289-305, Feb. 1999.

[12] P. J. Huber, *Robust statistics.* New York: Wiley, 1981

[13] W.-K. Ma, T.-N. Davidson, K. M. Wong, and P. C. Ching, “A block alternating likelihood maximization approach to multiuser detection,” IEEE Trans. Signal Process., vol. 52, no. 9, pp. 2600–2611, Sep. 2004.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004

[15] J. J. Fuchs “An Inverse Problem Approach to Robust Regression” IEEE International Conference on Acoustics, Speech, and Signal Processing. Vol. 4 pp 1809-1812, Mar 1999.

[16] Z. -Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye. S. Zhang, “Semi-definite Relaxation of Quadratic Optimization Problems” IEEE Signal Processing Magazine, May 2010.

[17] S. M. Zabin and H. V. Poor, “Efficient estimation of the class A parameters via the EM algorithm,” IEEE Trans. Inform. Theory, vol. 37, pp. 60–72, Jan. 1991.