CHANNEL DEPENDENT TERMINATION OF THE SEMIDEFINITE RELAXATION DETECTOR

Joakim Jaldén, Björn Ottersten*

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

ABSTRACT

We study the problem of semidefinite relaxation (SDR) for detection of symbols transmitted over a general MIMO channel. In the SDR detector the maximum likelihood detection problem is relaxed into a semidefinite program (SDP) which is solved numerically using an interior-point path-following algorithm. Herein, we provide a criteria which, based on the channel matrix realization, determine the accuracy required by the SDP solver to give a good bit error rate performance of the overall SDR detector. This also reduce the complexity of the SDR detector as it limits the number of interior iterations required in the SDP solver. The performance is demonstrated through simulations.

1. INTRODUCTION

Semidefinite relaxation (SDR) has been presented a near optimal and computationally attractive alternative to exact maximum likelihood (ML) detection for symbols transmitted over a general multiple-input multiple-output (MIMO) channel, see e.g. [1, 2]. The SDR detector approaches the detection problem by first relaxing the constraints of the ML detection problem to obtain an convex optimization problem. Then, from the solution to this convex problem an estimate of the transmitted message is obtained. The near optimality of the SDR detector has been supported both by simulations [1, 2] as well as analytical results [1, 3]. For example, in [1] it is shown that several popular suboptimal detectors can be viewed as further relaxation of the SDR detector and [3] presents necessary and sufficient conditions for the SDR estimate to coincide with the ML estimate.

The algorithmic solutions to the convex problem which forms the basis of the SDR detector do not provide an exact solution but rather an approximate solution which can be iteratively refined to some arbitrary precision. Clearly, the iterative procedure should be continued until no significant further improvement in the estimates are obtained as further iterations will only add to the complexity of the detector. However, most of the analysis of SDR only make statements regarding the *objective value* of the optimization problem and not the *optimization variables* themselves, even though these are more directly related to the bit error rate (BER) of the detector.

Most efficient algorithmic solutions to the semidefinite optimization problem are based on the concept of an interior-point *central path* [4] which may be viewed as a set of perturbed solutions to the optimization problem which generate a sequence of iterates approaching the optimal solution. The central path is given implicitly as the solution to a non-linear system of equations and can in general not be expressed explicitly in closed form. One of the main contributions of this work is to provide a linearization of the central path at the optimal solution corresponding to the noise-free case. This provides a perturbation analysis for the SDR and the basis for a criteria which relates the solution accuracy with the error probability of the SDR detector. It is shown that the accuracy can be adaptively based on the channel matrix realization which ultimately provide a means to limit the complexity of the SDR detector without sacrificing BER performance.

In Section 2 the detection problem, along with the necessary background for semidefinite relaxation is presented. The main analytical result of this work is presented in Section 2.3. The proposed termination criteria is evaluated numerically in Section 3 which is followed by conclusions in Section 4. In what follows, we denote vectors and matrices using boldface characters. No notational difference is made between random variables and their realizations. The notation $[\cdot]_i$ and $[\cdot]_{i,j}$ is used to denote the ith and (i,j)th element of vectors and matrices respectively. The operators $\mathrm{Diag}(\cdot)$ and $\mathrm{diag}(\cdot)$ denote a diagonal matrix with some given entries and the vector of diagonal elements respectively. $\mathbf{X} \succeq \mathbf{0}$ is used to denote that $\mathbf{X}$ is a positive semidefinite and symmetric matrix. Finally, the symbol $\circ$ denotes entrywise (Hadamard) matrix and vector multiplication.

2. DATA MODEL AND DETECTION

The specific detection problem we consider herein is the detection of BPSK symbols transmitted over an AWGN MIMO channel given by

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{v}. \tag{1}$$

In the above, $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{H} \in \mathbb{R}^{n\times m}$ and $\mathbf{v} \in \mathbb{R}^n$ is the vector of received signals, channel matrix, and additive noise respectively. The transmitted symbols, $\mathbf{s} \in \mathcal{B}^m \triangleq \{\pm 1\}^m$, are assumed to be i.i.d. distributed over the signal set and the noise, $\mathbf{v}$, is assumed to be distributed according to a Gaussian distribution with zero mean and unit variance. We will also assume throughout that $n \geq m$ and that $\mathbf{H}$ is full rank. Under these assumptions, the maximum likelihood estimate of $\mathbf{s}$, given $\mathbf{y}$ and $\mathbf{H}$, is well known to be

$$\hat{\mathbf{s}}_{\mathrm{ML}} = \arg\min_{\hat{\mathbf{s}}\in\mathcal{B}^m} \|\mathbf{y} - \mathbf{H}\hat{\mathbf{s}}\|^2, \tag{2}$$

a problem which unfortunately is NP-hard for general $\mathbf{H}$ and $\mathbf{y}$ [5]. This provides motivation for computationally efficient suboptimal approches such as the SDR detector. Note also that while we explicitely consider a real valued channel model the case of a complex channel using QPSK can be treated by writing the model on an equivalent real valued form, see e.g. [6].

*jalden@s3.kth.se

2.1. Semidefinite Relaxation

The optimization problem given by

$$\begin{array}{ll}\min\limits_{\mathbf{X},\,\mathbf{x}} & \mathrm{tr}(\mathbf{L}\mathbf{X}) \\ \text{s.t.} & \mathrm{diag}(\mathbf{X}) = \mathbf{e} \\ & \mathbf{X} = \mathbf{x}\mathbf{x}^{\mathrm{T}}\end{array} \tag{3}$$

where

$$\mathbf{L} = \begin{bmatrix} \mathbf{H}^{\mathrm{T}}\mathbf{H} & -\mathbf{H}^{\mathrm{T}}\mathbf{y} \\ -\mathbf{y}^{\mathrm{T}}\mathbf{H} & \mathbf{y}^{\mathrm{T}}\mathbf{y} \end{bmatrix}, \ \ \mathbf{x} = \begin{bmatrix} \hat{\mathbf{s}} \\ 1 \end{bmatrix}.$$

and $\mathbf{e}$ is the vector of ones is equivalent to (2) in the sense that the solution to either problem is easily obtained from the solution to the other [2]. Essentially, (3) allows the quadric objective of (2) to be replaced by a linear objective by expanding the dimension of the optimization variable from $\mathbb{R}^m$ to $\mathbb{R}^{m+1\times m+1}$. The rank one constraint $\mathbf{X} = \mathbf{x}\mathbf{x}^{\mathrm{T}}$ along with the constraint $\mathrm{diag}(\mathbf{X}) = \mathbf{e}$ enforces a one-to-one correspondence between $\mathbf{X}$ in (3) and $\hat{\mathbf{s}} \in \mathcal{B}^m$ in (2). Also, due to symmetry of the problem the constraint $[\mathbf{x}]_{m+1} = 1$ does not have to be maintained explicitly.

As (2) and (3) are equivalent they are equally hard to solve in a complexity theoretic sense. However, by replacing the rank one constraint, $\mathbf{X} = \mathbf{x}\mathbf{x}^{\mathrm{T}}$, with a less restrictive positive semidefinite constraint, $\mathbf{X} \succeq 0$, a convex, and thus efficiently solvable, optimization problem given by

$$\begin{array}{ll}\min\limits_{\mathbf{X}} & \mathrm{Tr}(\mathbf{L}\mathbf{X}) \\ \text{s.t.} & \mathrm{diag}(\mathbf{X}) = \mathbf{e} \\ & \mathbf{X} \succeq \mathbf{0}\end{array} \tag{4}$$

is obtained. The problem in (4) is referred to as the semidefinite relaxation of (3), or equivalently (2). Whenever the optimal solution, $\mathbf{X}_0$, of (4) is of rank one it is also optimal for (3). The existence of rank one solutions to (4) is however by no means guaranteed and will generally depend on the particular realization of $\mathbf{y}$ and $\mathbf{H}$ [3].

An SDR detector solves (4) to obtain an, not necessarily rank one, optimal solution, $\mathbf{X}_0$. From this solution an estimate, $\hat{\mathbf{s}}$, of $\mathbf{s}$ is obtained using some approximation strategy. In [2] it is suggested to either let $\hat{\mathbf{s}}$ be equal to the sign of the first m elements of the last column of $\mathbf{X}_0$ or to base the estimate on the sign of the eigenvector corresponding to the largest eigenvalue of $\mathbf{X}_0$. It can be shown that in both cases $\hat{\mathbf{s}}$ will be equal to $\hat{\mathbf{s}}_{\mathrm{ML}}$ whenever $\mathbf{X}_0$ is rank one. Another approach is to randomly generate several candidates for $\hat{\mathbf{s}}$ based on $\mathbf{X}_0$, which can then be evaluated by (2), see [1]. Numerical evidence suggests that this approach is superior (in terms of final bit error rate) to the previously mentioned. However, in the interest of analytical simplicity, we shall herein only consider the first approach where $[\hat{\mathbf{s}}]_i = \mathrm{sign}([\mathbf{X}_0]_{i,m+1})$.

2.2. Interior-Point Solutions

The optimization problem posed in (4) is typically, and efficiently, solved numerically using a primal-dual interior-point approach [4]. In such an approach a sequence of approximate solutions to the system of equations given by

$$\mathrm{diag}(\mathbf{X}_\tau) = \mathbf{e} \tag{5a}$$

$$\mathbf{Z}_\tau + \mathrm{Diag}(\mathbf{u}_\tau) = \mathbf{L} \tag{5b}$$

$$\mathbf{X}_\tau\mathbf{Z}_\tau = \tau\mathbf{I} \tag{5c}$$

$$\mathbf{X}_\tau, \mathbf{Z}_\tau \succeq \mathbf{0} \tag{5d}$$

are obtained as $\tau \to 0$. In (5) $\tau > 0$ parameterizes a set of unique solutions, $(\mathbf{X}_\tau, \mathbf{Z}_\tau, \mathbf{u}_\tau)$, referred to as the *central path* where $(\mathbf{Z}_\tau, \mathbf{u}_\tau)$ are auxiliary variables corresponding to the dual optimization problem of (4). Further, when $\tau = 0$ (5) is referred to as the Karush-Kuhn-Tucker (KKT) optimality conditions for the SDR and any $\mathbf{X}_0$ for which there exists a corresponding $\mathbf{Z}_0$ and $\mathbf{u}_0$ such that (5) holds for $\tau = 0$ is also optimal in (4). Essentially, an interior-point method is an iterative algorithm which follows the central path towards the optimal solution of (5) for $\tau = 0$ and thus also of (4).

Herein, following [2], we consider the Helmberg-Kojima-Monteiro (HKM) interior-point method [4]. The HKM method produces a series of iterates $(\mathbf{X}^{(k)}, \mathbf{Z}^{(k)}, \mathbf{u}^{(k)})$ which satisfy (5a), (5b) and (5d) (but not necessarily (5c)) and which converge towards the solution of (5) for $\tau = 0$. At each iterate a Newton step is taken towards the solution of (5) for $\tau = \tau^{(k)}$ where

$$\tau^{(k)} \triangleq \sigma\frac{\mathrm{Tr}(\mathbf{X}^{(k)}\mathbf{Z}^{(k)})}{m+1} \tag{6}$$

for some user specified $\sigma \in (0, 1)$. Note that due to an inherent poor conditioning of the system at the optimal solution (i.e. for $\tau = 0$) a strategy of simply taking Newton steps toward this solution will either converge slowly or not converge at all. However, by applying the HKM strategy convergence to a fix tolerance ϵ for which

$$\frac{\mathrm{Tr}(\mathbf{X}^{(k)}\mathbf{Z}^{(k)})}{m+1} \leq \epsilon \tag{7}$$

can be guaranteed in $\mathcal{O}(\sqrt{m})$ iterations. With a computational complexity of $\mathcal{O}(m^3)$ operations per iteration the HKM method offers a ϵ-approximate solution to (4) with a complexity of $\mathcal{O}(m^{3.5})$. For a full description or the algorithm the reader is referred to [2] or [4] and references therein.

Any $(\mathbf{X}_\tau, \mathbf{Z}_\tau, \mathbf{u}_\tau)$ on the central path satisfies $\mathrm{Tr}(\mathbf{X}_\tau\mathbf{Z}_\tau)/(m+1) = \tau$. Although the iterates of the HKM method will *follow* the central path none of them will, strictly speaking, lie directly *on* the central path (trying to enforce this would inevitably lead to an inefficient algorithm). It is however still reasonable to assume that iterates satisfying (7) will behave similarly to central path solutions for $\tau = \epsilon$. Therefore, results regarding τ will be used to obtain statements regarding ϵ and although it is hard to give stringent mathematical support for this procedure, it will be shown numerically that the termination criterion presented herein (which is based on this notion) does in fact lead to a significant reduction in complexity without sacrificing the BER of the detector.

2.3. Channel Adaptive Termination

It can be shown [4] that any $\mathbf{X}$ satisfying (5a), (5b), (5d) and (7) is ϵ-optimal in the sense that $\mathrm{Tr}(\mathbf{L}\mathbf{X}_0) \geq \mathrm{Tr}(\mathbf{L}\mathbf{X}) - \epsilon$. This is however a statement regarding the objective value of the solution and not the solution itself. Ultimately, it is the distance between $\mathbf{X}$ and $\mathbf{X}_0$ which will affect the performance of the SDR detector in terms of its BER and ϵ should be selected to to give an acceptable tradeoff between the BER and complexity (or number of iterations). In what follows we argue that ϵ can be set adaptively based on the channel matrix, $\mathbf{H}$, and give a heuristic for selecting ϵ without sacrificing BER performance.

At the heart of this heuristic is the realization that the sensitivity of the central path solutions $(\mathbf{X}_\tau, \mathbf{Z}_\tau, \mathbf{u}_\tau)$ to a non-zero τ, i.e. $\tau > 0$, indicate to which tolerance the SDR problem of (4) needs to be solved. Recall that the final estimate, $\hat{\mathbf{s}}$, of the SDR detector is based on the sign of the last column of $\mathbf{X}$. Thus, ideally we would like to express the elements of the last column as functions of τ. While no exact result exists in closed form the first order approximations around the noise-free (i.e. when $\mathbf{v} = 0$) points are given by a relatively simple expression. This is formalized by the following proposition (proved in Section 5) which forms the main analytical contribution of this work.

Proposition 1 *Consider the solution,* $\mathbf{X}_\tau$*, of* (5) *in the zero noise (*$\mathbf{v} = \mathbf{0}$*) case. Let*

$$\mathbf{X}_\tau \triangleq \begin{bmatrix} \mathbf{\Psi}_\tau & \boldsymbol{\psi}_\tau \\ \boldsymbol{\psi}_\tau^{\mathrm{T}} & 1 \end{bmatrix}$$

for $\mathbf{\Psi} \in \mathbb{R}^{m \times m}$ *and* $\boldsymbol{\psi} \in \mathbb{R}^m$*. Then*

$$\boldsymbol{\psi}_\tau = \mathbf{s} - \frac{\tau}{2}\mathbf{s} \circ \operatorname{diag}(\mathbf{Q}^{-1}) + \mathcal{O}(\tau^2) \tag{8}$$

where $\mathbf{Q} \triangleq \mathbf{H}^{\mathrm{T}}\mathbf{H}$.

Note that from the expression for $\boldsymbol{\psi}_\tau$ it can be seen that the optimal solution in the zero noise case is equal to the transmitted message, i.e. $\boldsymbol{\psi}_0 = \mathbf{s}$. This is simply stating that if there is no noise present, then the SDR detector will always recover the correct message (at least if (4) were to be solved exactly). Further, it states that the sensitivity of the ith entry of $\boldsymbol{\psi}_\tau$ to a non-zero τ depends directly on the ith diagonal entry of $\mathbf{Q}^{-1} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}$.

In order to further interpret this result, consider the example where $\mathbf{H} = \sqrt{\rho}[\mathbf{h}_1\mathbf{h}_2]$, $\|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 1$ and $\mathbf{h}_1^{\mathrm{T}}\mathbf{h}_2 = \xi$. In this case the diagonal entries of $\mathbf{Q}^{-1}$ are given by $[\mathbf{Q}^{-1}]_{i,i} = \rho^{-1}(1-\xi^2)^{-1}$. Thus, due to an increasing value of $[\mathbf{Q}^{-1}]_{ii}$ the solution to (5) will be more sensitive to a non-zero τ when either the SNR is low (when ρ is small) or when the channel is poorly conditioned (when ξ is close to ± 1). This corresponds well with the intuition that (2) is harder to solve or approximate in these cases.

Based on (8) we propose selecting $\epsilon = \epsilon(\mathbf{H})$ according to

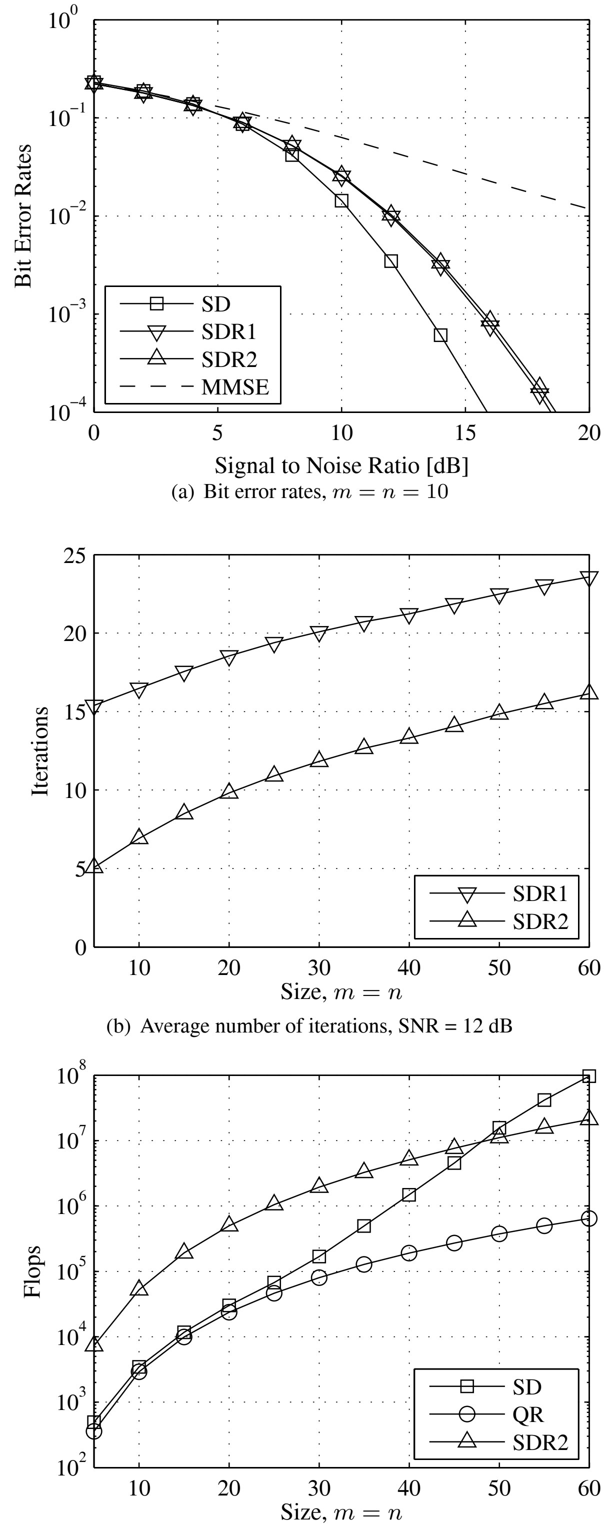
$$\epsilon(\mathbf{H}) = \frac{c}{\max_i [\mathbf{Q}^{-1}]_{i,i}} \tag{9}$$

for some c, i.e. if the diagonal entries of $\mathbf{Q}^{-1}$ are large then the solution is sensitive to a perturbation in the central path parameter τ and the semidefinite program of (4) needs to be solved to a higher tolerance. Alternatively, for high SNR and well conditioned channels it is less crucial to obtain an exact solution and fewer iterations are required when solving the SDR optimization problem in (4).

3. SIMULATIONS RESULT

In this section a numerical evaluation of the proposed method is given. For the purpose of the simulations, realizations of the $\mathbf{H}$ matrix were generated for the $m = n$ case with i.i.d. (real valued) Gaussian elements of variance ρ/n where ρ is the SNR of the channel. The results reported in Fig. 1 were averaged over both channel realization as well as noise and symbol realizations. Further, initialization of the SDR detector was done according to the procedure outlined in [2].

First, Fig. 1(a) shows the bit error rate of the SDR detector compared to the ML detector (implemented using the sphere decoder (SD)) as well as the MMSE detector. Version one, SDR1, of the SDR detector solve (4) to a high precision (of $\epsilon = 10^{-6}$ which is the same as used in [2]) while version 2, SDR2, use the adaptive termination criteria of (9) for $c = 1$. As can be seen from Fig. 1(a) the loss in terms of BER which can be attributed to an unprecise solution of (4) is minor. Also, the gap to the ML detector can be further reduced by considering more advanced procedures of converting the SDR solution to a symbol estimate, $\hat{\mathbf{s}}$, see e.g. [1]. Naturally, the



(a) Bit error rates, $m = n = 10$

(b) Average number of iterations, SNR = 12 dB

(c) Average floating point operations, SNR = 12 dB

Fig. 1. Performance of the SDR detector

proposed termination criteria will be useful also in this case, even though it was motivated using a simpler strategy.

Fig. 1(b) shows the number of iterations required by the two SDR versions to converge to the specified tolerance. It is worthwhile to note that the number of iterations required by an implementation based on a fixed tolerance, ϵ, can be made closer to the one of SDR2 by selecting ϵ based on the channel statistics and SNR (which could be accomplished by extensive simulations). However, such an approach will inevitably be targeted at the specific scenario and experience a loss in either BER of complexity performance when the statistics of the problem are changed and lack the ability to handle poorly conditioned channel realizations. Further, note that it is only the number of iterations and not the computational complexity per iteration which is affected by the termination criteria. Thus, only a complexity reduction by a constant factor (as is apparent in Fig. 1(b)) should be expected.

Finally, the complexity in terms of floating point operations for the implementation of the SDR2 detector is shown in Fig. 1(c). As a comparison the complexity of the popular sphere decoder algorithm (implemented according to the Schnorr-Euchner strategy with an initial radius tangent to the Babai-estimate [7]) along with the QR-factorization pre-processing step of this decoder is displayed. As can be seen from the figure, for small problems the complexity of the sphere decoder is dominated by the QR-factorization which is of less computational complexity than the SDR detector. However, due to the (average and worst case) exponential complexity of the sphere decoder [8] the (polynomial time) SDR detector is more computationally attractive for large scale problems. For this particular example, the cross over point is at around $m \approx 48$ but this usually varies with the statistics of the channel model and the SNR. Note also that the average complexity of the semidefinite relaxation detector, as well as the sphere decoder, can be further reduced using the techniques outlined in [6].

4. CONCLUSIONS

Based on the algorithmic solution to a semidefinite program forming an integral part of the SDR detector we derive a sensitivity analysis on which a channel adaptive termination criteria is based. This criteria is used to lower the computational complexity of the SDR detector and its effectiveness is illustrated through numerical simulations.

5. PROOF OF PROPOSITION 1

When $\mathbf{v} = 0$ the solution of (5) for $\tau = 0$ is given by

$$\mathbf{X}_0 = \begin{bmatrix} \mathbf{s}\mathbf{s}^{\mathrm{T}} & \mathbf{s} \\ \mathbf{s}^{\mathrm{T}} & 1 \end{bmatrix}, \qquad \mathbf{u}_\tau = \mathbf{0}$$

$$\text{and} \qquad \mathbf{Z}_0 = \mathbf{L} \qquad \text{where} \qquad \mathbf{L} = \begin{bmatrix} \mathbf{Q}^{\mathrm{T}} & -\mathbf{Q}\mathbf{s} \\ -\mathbf{s}^{\mathrm{T}}\mathbf{Q} & \mathbf{s}^{\mathrm{T}}\mathbf{Q}\mathbf{s} \end{bmatrix}$$

which can be verified by substituting the solution back into (5). Let $\tilde{\mathbf{X}}_\tau$ and $\tilde{\mathbf{Z}}_\tau$ be the deviation from the optimal solution due to $\tau > 0$, i.e. $\tilde{\mathbf{X}}_\tau \triangleq \mathbf{X}_\tau - \mathbf{X}_0$ and $\tilde{\mathbf{Z}}_\tau \triangleq \mathbf{Z}_\tau - \mathbf{Z}_0$. By expanding (5c) one obtains $\mathbf{X}_\tau\mathbf{Z}_\tau = \mathbf{X}_0\mathbf{Z}_0 + \tilde{\mathbf{X}}_\tau\mathbf{Z}_0 + \mathbf{X}_0\tilde{\mathbf{Z}}_\tau + \tilde{\mathbf{X}}_\tau\tilde{\mathbf{Z}}_\tau = \tau\mathbf{I}$. Using $\mathbf{X}_0\mathbf{Z}_0 = \mathbf{0}$ and $\tilde{\mathbf{X}}_\tau\tilde{\mathbf{Z}}_\tau \in \mathcal{O}(\tau^2)$ it follows that

$$\underbrace{\tilde{\mathbf{X}}_\tau\mathbf{Z}_0 + \mathbf{X}_0\tilde{\mathbf{Z}}_\tau}_{\mathbf{A}} = \tau\mathbf{I} + \mathcal{O}(\tau^2). \tag{10}$$

Now let

$$\tilde{\mathbf{X}}_\tau \triangleq \begin{bmatrix} \tilde{\mathbf{\Psi}}_\tau & \tilde{\boldsymbol{\psi}}_\tau \\ \tilde{\boldsymbol{\psi}}_\tau^{\mathrm{T}} & 0 \end{bmatrix} \qquad \text{and} \qquad \tilde{\mathbf{Z}}_\tau \triangleq \begin{bmatrix} \tilde{\mathbf{\Phi}}_\tau & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & \tilde{\phi} \end{bmatrix}.$$

Note that due to the constraints given by (5a) $\tilde{\mathbf{X}}_\tau$, and $\tilde{\mathbf{\Psi}}_\tau$, have zero elements along the diagonal. Equivalently, by (5b) $\tilde{\mathbf{Z}}_\tau$ must be a diagonal matrix. Expanding (10) using the above parameterizations yields

$$\mathbf{A} = \begin{bmatrix} \tilde{\mathbf{\Psi}}_\tau\mathbf{Q} - \tilde{\boldsymbol{\psi}}_\tau\mathbf{s}^{\mathrm{T}}\mathbf{Q} + \mathbf{s}\mathbf{s}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_\tau & -\tilde{\mathbf{\Psi}}_\tau\mathbf{Q}\mathbf{s} + \tilde{\boldsymbol{\psi}}_\tau\mathbf{s}^{\mathrm{T}}\mathbf{Q}\mathbf{s} \\ \tilde{\boldsymbol{\psi}}_\tau^{\mathrm{T}}\mathbf{Q} + \mathbf{s}^{\mathrm{T}}\tilde{\mathbf{\Phi}}_\tau & -\tilde{\boldsymbol{\psi}}_\tau^{\mathrm{T}}\mathbf{Q}\mathbf{s} + \tilde{\phi}_\tau \end{bmatrix}. \tag{11}$$

From the bottom left block of $\mathbf{A}$ in (11), along with the diagonal constraint for $\tilde{\mathbf{\Phi}}_\tau$, it follows that

$$\tilde{\mathbf{\Phi}}_\tau = -\mathrm{Diag}(\mathbf{s})\mathrm{Diag}(\mathbf{Q}\tilde{\boldsymbol{\psi}}_\tau) + \mathcal{O}(\tau^2)$$

where $[\mathbf{s}]_i^2 = 1$ has been used. Inserting this into the top left block of $\mathbf{A}$ yields

$$\tilde{\mathbf{\Psi}}\mathbf{Q} - \tilde{\boldsymbol{\psi}}_\tau\mathbf{s}^{\mathrm{T}}\mathbf{Q} + \mathbf{s}\tilde{\boldsymbol{\psi}}_\tau^{\mathrm{T}}\mathbf{Q} = \tau\mathbf{I} + \mathcal{O}(\tau^2)$$

and

$$\tilde{\mathbf{\Psi}}_\tau - \tilde{\boldsymbol{\psi}}_\tau\mathbf{s}^{\mathrm{T}} - \mathbf{s}\tilde{\boldsymbol{\psi}}_\tau^{\mathrm{T}} = \tau\mathbf{Q}^{-1} + \mathcal{O}(\tau^2)$$

by right multiplication with $\mathbf{Q}^{-1}$. The diagonal elements of the above matrix are given by

$$-2[\mathbf{s}]_i[\tilde{\boldsymbol{\psi}}_\tau]_i = \tau[\mathbf{Q}^{-1}]_{i,i} + \mathcal{O}(\tau^2)$$

(as $\tilde{\mathbf{\Psi}}_\tau$ has a zero diagonal) or equivalently

$$\tilde{\boldsymbol{\psi}}_\tau = -\frac{\tau}{2}\mathbf{s} \circ \mathrm{diag}(\mathbf{Q}^{-1}) + \mathcal{O}(\tau^2)$$

where $([\mathbf{s}]_i)^2 = 1$ has been used again. Combining this with $\boldsymbol{\psi}_0 = \mathbf{s}$ which follows directly from the expression for $\mathbf{X_0}$ yields

$$\boldsymbol{\psi}_\tau = \mathbf{s} - \frac{\tau}{2}\mathbf{s} \circ \mathrm{diag}(\mathbf{Q}^{-1}) + \mathcal{O}(\tau^2)$$

which establishes the proposition. ■

6. 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. Jaldén, C. Martin, and B. Ottersten, “Semidefinite programming for detection in linear systems – optimality conditions and space-time decoding,” in *Proc. IEEE ICASSP'03*, Apr. 2003.

[4] C. Helmberg, F. Rendl, R.J. Vanderbei, and H. Wolkowicz, “An interior-point method for semidefinite programming,” *SIAM Journal on Optimization*, vol. 6, pp. 342–361, 1996.

[5] S. Verdú, “Computational complexity of multiuser detection,” *Algorithmica*, vol. 4, pp. 303–312, 1989.

[6] J. Jaldén, B. Ottersten, and W.-K. Ma, “Reducing the average complexity of ML detection using semidefinite relaxation,” in *Proc. IEEE ICASSP'05*, Mar. 2005.

[7] M. O. Damen, H. El Gamal, and G. Caire, “On maximum-likelihood detection and the search for the closest lattice point,” *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2389–2401, Oct. 2003.

[8] J. Jaldén and B. Ottersten, “On the complexity of sphere decoding in digital communications,” *IEEE Transaction on Signal Processing*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.