1 2

First (corresponding) author: M. Chrétien is with the Laboratoire de Mathématiques, UMR CNRS 6623 and Université de Franche Comté, 16 route de Gray, 25030 Besançon Cedex, France. Tel: +33 3 81 66 64 65. Fax: +33 3 81 66 66 23. Email: stephane.chretien@math.univ-fcomte.fr

Second author: M. Corset is with LabSAD Université Pierre Mendes France, 1251 Avenue centrale. BP47 38040 Grenoble cedex 9, France. Tel: +33 4 76 82 57 69. Fax: +33 4 76 82 56 65. Email: franck.corset@upmf-grenoble.fr

Using the Eigenvalue Relaxation for Binary Least-Squares Estimation Problems: A survey

St´ephane Chr´etien and Franck Corset

Abstract

The goal of this paper is to survey the properties of the eigenvalue relaxation for least squares binary problems. This relaxation is a convex program which is obtained as the Lagrangian dual of the original problem with an implicit compact constraintand as such, is a convex problem with polynomial time complexity. The necessary tools from convex analysis are recalled and shown at work for handling the problem of exactness of this relaxation. Two applications are described. The first one is the problem of binary image reconstruction and the second is the problem of multiuser detection in CDMA systems.

THE goal of this paper is to study the eigenvalue relaxation of binary least squares estimation problems. Penalized binary least squares estimation problems are problems of the form

$$\min_{x\in\mathbb{R}^n} \|y - Ax\|^2 + \nu x^t P x \quad \text{s.t.} \quad x \in \{-1, 1\}^n, \tag{-.1}$$

where the vector $y \in \mathbb{R}^m$ is the observed data, the matrix $A \in \mathbb{R}^{m\times n}$ represents the "filter", the vector $x \in \mathbb{R}^n$ is the signal, or parameter vector, that has to be estimated, and the term $\nu x^t P x$ is a penalization term that can often be interpreted as an *a priori* information in terms of Bayesian statistics.

This problem belongs to the larger class of minimization of quadratic forms over binary vectors which is known to be $\mathcal{NP}$-hard. Much work has been devoted to constructing Semi Definite Programming (SDP) based relaxations for general quadratic binary problems. SemiDefinite programs are linear optimization problems over symmetric matrices with real coefficients with the additional convex constraint of positive semidefiniteness; see for instance [5] or [1] for excellent books on convex programming presenting SDP. SDP methods have already played an important role in various topics inside signal processing problems and we refer to [2] for a nice survey on possible applications. A common feature of essentially all the existing relaxations is that they can be obtained using Lagrange duality which is a general methodology for obtaining lower bounds to hard minimization problems, as overviewed in [3] and [4].

M. Chr´etien is with the Laboratoire de Math´ematiques, UMR CNRS 6623 and Universit´e de Franche Comt´e, 16 route de Gray, 25030 Besançon Cedex, France. Email: stephane.chretien@math.univ-fcomte.fr

M. Corset is with LabSAD Universit´e Pierre Mendes France, 1251 Avenue centrale. BP47 38040 Grenoble cedex 9, France. Email: franck.corset@upmf-grenoble.fr

The goal of the paper is to survey what is known about another Lagrangian duality based relaxation, namely the eigenvalue relaxation, for this problem. This relaxation was first proposed by Delorme and Poljak [18] for the max-cut problem. See also the work of Poljak, Rendl and Wolkowics [6] for more details. The main advantage of the eigenvalue relaxation over the SDP relaxation is that the optimum value of the former is often much faster to solve than the latter, as reported for instance in [7], [8] and [9]. This remarkable computational tractability of the eigenvalue relaxation is the main motivation for writing this detailed survey.

The content of the paper is as follows. The first section is devoted to a rapid presentation of the relaxation and its relationship with Lagrangian duality. We also recall a simple and well known certificate for exactness of the relaxation, i.e. the fact that an globally optimal binary solution is obtained.

The second second details the relationships between the SemiDefinite Relaxation and the eigenvalue relaxation. The main result of this section is the following: in addition to the fact that the eigenvalue relaxation is much faster to solve than the SDP relaxation, a solution of the SDP relaxation can be recovered from the solution of the eigenvalue relaxation. The case of inexact solutions to the eigenvalue relaxation is also studied.

In the last section, we propose simulation experiments in the case of binary image denoising and CDMA Multiuser Detection problems. The first of these problems has been previously approached by stochastic methods based on Markov chains like simulated annealing and Metropolis Hastings schemes; see for instance [16] and the more recent work of Gibbs [13]. The approach discussed here was presented in [14]. Recently, the same and a lot more problems have been addressed using the SDP relaxation in [15]. The results obtained so far are quite encouraging and the approach performs well on very dirty images. Passing to the second problem, our Monte Carlo experiments show that even for a small number of users, the eigenvalue relaxation outperforms the SDP relaxation in terms of average computational complexity.

Notations. In the sequel we will use the following notations. The inner product on $\mathbb{R}^n$ is denoted by $\langle \cdot, \cdot \rangle$, the set of real symmetric matrices of order n are denoted by $\mathbb{S}_n$. The partial order $\succcurlyeq$ denotes the Lowner ordering, i.e. for A and B in $\mathbb{S}_n$, $A \succcurlyeq B$ means that $A - B$ is positive semidefinite. For a set S in $\mathbb{R}^n$, $\mathrm{conv}(S)$ denotes the convex hull of S and $\overline{S}$ denotes its closure. For a matrix A in $\mathbb{S}_n$, $d(A)$ denotes its diagonal vector and for a in $\mathbb{R}^n$, $D(a)$ denotes the diagonal matrix whose diagonal vector is a. If an equation number $\#$ corresponds to an optimization problem, then $\mathrm{opt}(\#)$ will denote the optimum value for this problem.

I. The eigenvalue relaxation

We first introduce the eigenvalue relaxation and at the same time, we propose a quick refresher on Lagrangian duality, collecting all the results that will play an essential role in the sequel. The proofs of almost all the results presented here can be found in [17].

A. The Lagrangian dual and the eigenvalue relaxation

The least-squares estimation problem (-.1) is in fact equivalent to the homogenized problem

$$\max_{x\in\mathbb{R}^{n+1}} -x^t \begin{bmatrix} A^tA+\nu P & -A^ty \\ -y^tA & y^ty \end{bmatrix} x \text{ s.t. } x\in\{-1,1\}^{n+1}. \tag{I-A.1}$$

Indeed, if we add the constraint $x_{n+1}=1$ in (I-A.1), we obtain (-.1) (modulo the minus sign). Now, if x^* is a solution of (I-A.1), then $-x^*$ is again a solution of (I-A.1), thus adding the constraint $x_{n+1}=1$ is in fact redundant, which proves the claimed equivalence. Set

$$M=\begin{bmatrix} A^tA+\nu P & -A^ty \\ -y^tA & y^ty \end{bmatrix}. \tag{I-A.2}$$

Notice further that the constraint $x_i\in\{-1,1\}$ is equivalent to $x_i^2=1$ for all $i=1,\cdots,n+1$. Thus, to problem (-.1), we can associate the Lagrangian function

$$\begin{aligned} L(x,u) &= -x^tMx+\textstyle\sum_{i=1}^{n+1}u_i(x_i^2-1) \\ &= x^t(D(u)-M)x-u^te. \end{aligned} \tag{I-A.3}$$

Now we can add to the problem the implicit spherical constraint

$$\mathcal{S}_{n+1}=\{x\in\mathbb{R}^{n+1}\mid x^tx=n+1\},$$

which is redundant with the binary constraints. Then, optimizing over this sphere, we obtain the Lagrangian dual function, i.e.

$$\begin{aligned} \theta(u) &= \max_{x\in\mathcal{S}_{n+1}} x^t(D(u)-M)x-u^te \\ &= \max_{x\in\mathcal{S}_{n+1}} x^t(D(u)-M)x-\tfrac{u^te}{n+1}x^tx \\ &= \max_{x\in\mathcal{S}_{n+1}} x^t\Big(D(u)-M-\tfrac{u^te}{n+1}I\Big)x \end{aligned} \tag{I-A.4}$$

which, using Raleigh-Ritz variational formulation of the largest eigenvalue of symmetric matrices, can be written

$$\theta(u)=(n+1)\lambda_{\max}\Big(D(u)-M-\frac{u^te}{n+1}I\Big). \tag{I-A.5}$$

Finally, the dual problem, i.e. the eigenvalue relaxation, is given by

$$\min_{u\in\mathbb{R}^{n+1}}\theta(u). \tag{I-A.6}$$

B. Properties of the dual relaxation

1) Convexity: It is important to notice first that the dual function $\theta(u)$ is a convex, since it is the attained supremum over a family of linear function in the variable u parametrized by $x \in \mathcal{S}_{n+1}$.

2) Weak duality: The main classical property of the Lagrangian dual is weak duality, i.e.

$$\min_{u \in \mathbb{R}^{n+1}} \theta(u) \geq \mathrm{opt}(I - A.1), \tag{I-B.1}$$

where opt denotes the optimal value. Thus, we get

$$-\min_{u \in \mathbb{R}^{n+1}} \theta(u) \leq \mathrm{opt}(-.1). \tag{I-B.2}$$

This property explains in part why Lagrange duality is used : it provides a bound on the optimal value. When equality holds in the weak duality property, we say that strong duality holds. Sometimes, like in the case of the Max-Cut problem, the bound can be proved to be proportional to the optimal original value. More precisely, Goemans and Williamson proved that the optimum value of the eigenvalue relaxation (in fact the equivalent SDP formulation; see the original paper and Section II below) is greater than or equal to the optimal original value (this is just weak duality), which itself is always greater than or equal to .876 times the eigenvalue relaxation's optimal value. A quite similar but less tight bound, proved by Nesterov applies directly to the present problem. We will recall this bound in section III-A below.

3) Existence of dual solutions: It is well known that there exists an optimal dual solution. This was proved by Poljak and Wolkowicz in [19]. The proof given here is more direct.

Proposition 2.1: The dual function admits a minimizer.

Proof: Let $\theta^* = \inf_{u \in \mathbb{R}^{n+1}} \theta(u)$. Make the change of variable $v = u - \frac{1}{n+1} \sum_{i=1}^{n+1} u_i$, i.e. define

$$\eta(v) = (n+1)\lambda_{\max}(D(v) - M) = \theta(u). \tag{I-B.3}$$

We now have the property that $\sum_{i=1}^{n+1} v_i = 0$. We prove that η is coercive. Take any sequence $(v^k)_{k\in\mathbb{N}}$ with $\|v_k\| \to +\infty$ as $k \to +\infty$. We can assume that $v_i^k \to +\infty$ for some i because otherwise, the fact that $\|v_k\| \to +\infty$ implies that there must exists a sequence $(v_j^k)_{k\in\mathbb{N}}$ with $v_j^k \to -\infty$ and the fact that $\sum_{i=1}^{n+1} v_i = 0$ gives a contradiction. Now, the Gershgorin circle around the diagonal element $M_{i,i} + v_i^k$ has a constant radius, say r and its center goes to $+\infty$. Since $|M_{i,i} + v_i^k - \lambda_{\max}(D(v) - M)|$, this implies that $\lambda_{\max}(D(v) - M) \to +\infty$. Thus η is coercive and since it is continuous, it admits a minimizer that we will denote by v^*. Now, for all $u \in \mathbb{R}$, $v = u - \frac{1}{n+1}\sum_{i=1}^{n+1} u_i$, we have

$$\theta^* \leq \theta(v^*)$$

But, on the other hand, $\theta(v^*) = \eta(v^*) \leq \eta(v) = \theta(v) = \theta(u)$. Therefore,

$$\theta(v^*) \leq \theta^*$$

and the proof is complete. ■

4) Subdifferential's description and exactness criterion: The subdifferential $\partial\theta(u)$ of the eigenvalue relaxation has been much studied. Recall that for any convex function $f : \mathbb{R}^m \mapsto \mathbb{R}$, the subdifferential $\partial f(u)$ is defined by

$$\partial f(u) = \Big\{ g \in \mathbb{R}^m \mid f(u') \geq f(u) + g^t(u' - u) \Big\}.$$

The analysis of $\partial\theta(u)$ is based on the following general theorem.

Theorem 2.2: [17] Let $A : \mathbb{R}^m \mapsto \mathbb{S}_n$ be an affine operator defined by $A(u) = \mathcal{A}u + B$ for some linear operator $\mathcal{A} : \mathbb{R}^m \mapsto \mathbb{S}_n$ and some matrix $B \in \mathbb{S}_n$. Then, we have

$$\partial(\lambda_{\max} \circ A)(u)) = \mathcal{A}^* \partial\lambda_{\max}(A(u))$$

with

$$\partial\lambda_{\max}(X) =$$
$$E_{\max}\Big\{ Z \in \mathbb{S}_{r_{\max}} \mid Z \succcurlyeq 0 \text{ and } \mathrm{trace}(Z) = 1 \Big\} E_{\max}^t$$

where $\mathcal{A}^*$ is the adjoint of $\mathcal{A}$, $r_{\max}$ denotes the multiplicity of $\lambda_{\max}$ at $X \in \mathbb{S}_n$ and $E_{\max}$ is a matrix whose columns form any orthonormal basis of the eigenspace of X associated to $\lambda_{\max}$.

Now, if we set $A(u) = D(u) - M - \frac{u^t e}{n+1} I$, we get $B = -M$, $\mathcal{A}u = D(u) - \frac{u^t e}{n+1} I$ and $\mathcal{A}^* X = d(X) - \frac{1}{n+1}\mathrm{trace}(X)e$. For $d \in \mathbb{N}$, let $\mathcal{Z}_d$ be defined by

$$\mathcal{Z}_d = \Big\{ Z \in \mathbb{S}_d \mid Z \succcurlyeq 0 \text{ and } \mathrm{trace}(Z) = 1 \Big\}.$$

Using the previous theorem, we obtain

Corollary 2.3: The subdifferential $\partial\theta(u)$ of the dual function θ is given by

$$\partial\theta(u) = (n+1)d(E_{\max}\mathcal{Z}E_{\max}^t) - \mathrm{trace}(E_{\max}\mathcal{Z}E_{\max}^t)e$$

Following Oustry [7], the formula for $\partial\lambda_{\max}(X)$ in theorem 2.2 is proved by showing that the maximum eigenvalue function $\lambda_{\max}(X)$ on $\mathbb{S}_n$ is nothing but the support function $\sigma_{\mathcal{Z}_n}(X)$ of $\mathcal{Z}_n$, defined by

$$\sigma_{\mathcal{Z}_n}(X) = \sup_{Z\in\mathcal{Z}_n} \langle X, Z\rangle \qquad \text{(I-B.4)}$$

with the scalar product defined by $\langle X, Z\rangle = \mathrm{trace}(X, Z)$. By definition, the face $F_{\mathcal{Z}_n}(X)$ of $\mathcal{Z}_n$ exposed by X is the set of maximizers in (I-B.4), i.e.

$$F_{\mathcal{Z}_n}(X) = \Big\{ Z \in \mathcal{Z}_n \mid \lambda_{\max}(X) = \langle X, Z\rangle \Big\}.$$

Knowing that the subdifferential of a support function of a set is exactly the exposed face of this set, we finally get

$$\partial\lambda_{\max}(X) = \Big\{ Z \in \mathcal{Z}_n \mid \lambda_{\max}(X) = \langle X, Z\rangle \Big\}$$

the formula follows after some linear algebra.

There is a different path to the subdifferential's formula, which is perhaps more *a propos* in the context of duality: it is proved in [17, Chapter XII] that

$$\partial\theta(u) = \overline{\mathrm{conv}}\Big\{ (x_1^2 - 1, \cdots, x_{n+1}^2 - 1)^t \mid L(x, u) = \theta(u) \Big\}, \qquad \text{(I-B.5)}$$

where $\overline{\mathrm{conv}}$ denotes the closure of the convex hull. This fact is in fact true for general continuous constrained problems in the case where the underlying space is compact (for example) [1] and the associated technical condition is called the *filling property*. The following proposition provides a useful sufficient condition for proving that the relaxation is exact, i.e. strong duality applies.

Proposition 2.4: Let u^* be a minimizer of the dual eigenvalue relaxation. Then, if $\lambda_{\max}(A(u^*))$ has multiplicity one, then

$$\min_{u\in\mathbb{R}^{n+1}} \theta(u^*) = \mathrm{opt}(I - A.1)$$

and any eigenvector x of $A(u^*)$ whose squared norm is $n + 1$ is a binary solution of (I-A.1) (and of course also of (-.1)).

The proof is a direct consequence of the theory developed in [17, Chapter XII]. We provide a specialized proof here because it is short and instructive.

[1] which is the case here since we optimize over the sphere $\mathcal{S}_{n+1}$

Proof: Since the multiplicity of $\lambda_{\max}(A(u^*))$ is one, the subdifferential of $\lambda_{\max} \circ A$ at u^* is a single vector. Thus, θ is differentiable at u^* and its gradient is simply

$$\nabla\theta(u^*) = ({x_1^*}^2 - 1, \cdots, {x_{n+1}^*}^2 - 1)^t$$

for any x^* in $\mathcal{S}_{n+1}$ such that $\theta(u^*) = L(x^*, u^*)$. Since, u^* minimizes θ, we must have $\nabla\theta(u^*) = 0$. This implies that ${x_i^*}^2 = 1$ for all $i = 1, \cdots, n+1$. Thus, using weak duality

$$\mathrm{opt}(I - A.1) \leq \theta(u^*) = {x^*}^t(-M)x^* \leq \mathrm{opt}(I - A.1)$$

which proves that x^* solves the original problem (I-A.1). ■

We now have a nice criterion for deciding whether our relaxation worked exactly and we also know how to use it in order to recover a binary solution. This approach works for any quadratic binary problem and is extensively used for approximating combinatorial problems. However, the question remains on what to do when the relaxation is not exact, i.e. when the multiplicity at the optimum is greater than one. The next two sections will help answer this crucial question.

II. From eigenvectors to SDP solutions

The purpose of the next two sections is to describe how to recover primal binary solutions from the eigenvector solutions of the dual eigenvalue problem. It was first shown that good binary solution can be generated at random using the SDP solution by Goemans and Williamson [12] in the case of the Max-Cut problem in graph theory. Their results were then extended by Nesterov to the case of indefinite quadratic binary programming [10]. Those results allowed to conclude that both eigenvalue and SDP relaxations are in a certain precise sense very efficient. However, both relaxations are not equivalent from the computational point of view. Recall that one of the main motivations for using the eigenvalue relaxation is its manageable practical complexity which is often favorable compared to the one of solving the SDP relaxation. But what is not clear is how to generate good (primal) binary solutions in average with the eigenvalue relaxation only ? The first natural approach to this question is of course to try and recover an optimal SDP solution from the eigenvalue relaxation. Thus, we devote this section to this problem. It can be solved as follows : an appropriate convex combination of rank one matrices obtained from a set of optimal eigenvectors is shown to be a solution we are looking for. Our approach simplifies the presentation of [20]. The adaptation of the randomized algorithm of Goemans and Williamson and the associated bound established by Nesterov will be discussed in the next section.

A. The SDP relaxation

In order to obtain the Semi-Definite Programming (SDP) relaxation of the the homogenized problem (I-A.1), we begin with the following equivalence relating our problem to a problem on symmetric matrices. We have [2]

$$\mathrm{opt}(I-A.1) = \max_{x\in\mathbb{R}^{n+1}} \mathrm{trace}(-Mxx^t) \text{ s.t. } d(xx^t)=e.$$

This last problem is itself equivalent to

$$\max_{X\in\mathbb{S}_{n+1}} \mathrm{trace}(-MX) \text{ s.t. } d(X)=e,\ \ X\succcurlyeq 0,\ \ \mathrm{rank}X=1.$$

This problem being nonconvex, we drop the rank constraint and obtain the following SDP (convex) relaxation

$$\max_{X\in\mathbb{S}_{n+1}} \mathrm{trace}(-MX) \text{ s.t. } d(X)=e,\ \ X\succcurlyeq 0 \tag{II-A.1}$$

whose value is obviously greater than or equal to $\mathrm{val}(\mathrm{I}-\mathrm{A.1})$.

B. SDP versus maximal eigenvalue : theoretical equivalence

It follows from the subdifferential's formula given in Corollary 2.3 that at any minimizer u^*, we have

$$0\in\partial\theta(u^*) =$$
$$(n+1)d(E^*_{\max}\mathcal{Z}_{r^*_{\max}}{E^*_{\max}}^t)-\mathrm{trace}(E^*_{\max}\mathcal{Z}_{r^*_{\max}}{E^*_{\max}}^t)e.$$

Suppose we have in hand a matrix $Z^*\in\mathcal{Z}_{r^*_{\max}}$ such that

$$0=(n+1)d(E^*_{\max}Z^*{E^*_{\max}}^t)-\mathrm{trace}(E^*_{\max}Z^*{E^*_{\max}}^t)e. \tag{II-B.1}$$

It appears that a good guess for a candidate solution X^* to the SDP relaxation in the general case is

$$X^*=(n+1)E^*_{\max}Z^*{E^*_{\max}}^t.$$

We just need to check the details to see how it works. This result was initially proved in [20] but the proof given here is more direct.

Theorem 2.1: [20] Let u^* be the optimal solution of the eigenvalue relaxation let $E_{\max}$ be a matrix whose columns for an orthonormal basis of the eigenspace associated to $\lambda_{\max}(A(u^*))$ and let Z^* be as in (II-B.1). Then the matrix $X^*=$

[2]Here, we use the fact that $x^tMx=\mathrm{trace}(x^tMx)=\mathrm{trace}(Mxx^t)$

$(n+1)E^*_{\max}Z^*{E^*_{\max}}^t$ is an optimal solution of the SDP relaxation.

Remark 2.2: We would like to underline at this point that a more elegant proof of the theorem could be obtained using conic duality but we preferred to keep on with elementary arguments since this is possible in the present context.

Proof: Compute the eigenvalue/eigenvector decomposition $Z^* = U\Delta U^t$, set $F = E^*_{\max}U$, $\delta = d(\Delta)$, let r be the multiplicity of $A(u^*)$ and let $f_1, \cdots, f_r$ denote the columns of F. Recall that from the definition of Z^*, we have $\sum_{j=1}^r \delta_j = 1$. Then, we get

$$0 = d(F\Delta F^t) - \frac{1}{n+1}\mathrm{trace}(F\Delta F^t)e.$$

Thus,

$$\mathrm{trace}\Big((D(u^*) - \tfrac{1}{n+1}(u^*)^t eI)F\Delta F^t\Big) =$$
$$(u^*)^t d(F\Delta F^t) - (u^*)^t \tfrac{1}{n+1}\mathrm{trace}(F\Delta F^t)e = 0.$$

Using this fact, we obtain

$$\begin{aligned}
&\mathrm{trace}(-MX^*)\\
&= (n+1)\mathrm{trace}\Big((-M + D(u^*) - \tfrac{1}{n+1}(u^*)^t eI)F\Delta F^t\Big)\\
&= (n+1)\mathrm{trace}(A(u^*)F\Delta F^t)\\
&= (n+1)\mathrm{trace}(A(u^*)\textstyle\sum_{j=1}^r \delta_j f_j f_j^t)\\
&= (n+1)\textstyle\sum_{j=1}^r \delta_j f_j^t A(u^*) f_j\\
&= (n+1)\textstyle\sum_{j=1}^r \delta_j \lambda_{\max}(A(u^*))\\
&= (n+1)\lambda_{\max}(A(u^*)),
\end{aligned}$$

since $\sum_{j=1}^r \delta_j = 1$. Thus, the optimal value of the SDP is greater than or equal to the optimal value of the eigenvalue relaxation. On the other hand, we it is well known that the optimal value of the eigenvalue relaxation is greater than or equal to the one of the SDP relaxation. We provide a proof here for the sake of completeness. Let X^{**} be an optimal solution to the SDP relaxation. Now, for all u in $\mathbb{R}^{n+1}$, we have

$$\mathrm{trace}\Big(X^{**}(D(u) - \frac{e^t u}{n+1}I)\Big) = 0$$

by using the fact that $D(X^{**}) = e$. Now, compute the eigenvalue/eigenvector decomposition $-M + D(u) - \frac{e^t u}{n+1} I = \sum_{i=1}^{n+1} \lambda_i v_i v_i^t$ and let $\lambda_{\max}$ be the greatest of these eigenvalues. Then,

$$\begin{aligned} \mathrm{trace}(-MX^{**}) &= \mathrm{trace}\Big(X^{**}(-M + D(u) - \tfrac{e^t u}{n+1} I)\Big) \\ &= \textstyle\sum_{i=1}^{n+1} \lambda_i v_i^t X^{**} v_i \\ &\leq \lambda_{\max} \textstyle\sum_{i=1}^{n+1} v_i^t X^{**} v_i \\ &= \lambda_{\max} \mathrm{trace}(X^{**} \textstyle\sum_{i=1}^{n+1} v_i v_i^t) \\ &= \lambda_{\max} \mathrm{trace}(X^{**} I) \\ &= (n+1)\lambda_{\max} \end{aligned}$$

Since this is true for all u, we obtain that the eigenvalue relaxation majorates the SDP relaxation. Thus, both optimal values are equal and this completes the proof of the proposition. ■

C. SDP versus maximal eigenvalue: practical implementation

Of course, it can be hard to find a matrix $Z^* \in \mathcal{Z}_{r^*_{\max}} S$ that works. We will try to overcome this problem. We first have to specify how the subgradients are obtained in practice. At each point $u \in \mathbb{R}^{n+1}$, choose an eigenvector x of squared norm equal to $n+1$ associated to $\lambda_{\max}(A(u))$. Then, using the alternative representation of the subdifferential (I-B.5), a subgradient of θ at u is obtained by setting $g = [{x_1}^2 - 1, \ldots, {x_{n+1}}^2 - 1]^t$. Assume that we have a set of subgradients $g_j = [{x_1^j}^2 - 1, \ldots, {x_{n+1}^j}^2 - 1]^t \in \partial\theta(u^j)$ for some u^j, $j = 1, \ldots, p$ and such that

$$\|0 - \sum_{i=1}^{p} \alpha_j g_j\| \leq \epsilon, \qquad \text{(II-C.1)}$$

for some nonnegative α_j's with $\sum_{j=1}^{p} \alpha_j = 1$. This can be performed for ϵ as small as we want by using a bundle method. Such a method will construct in a finite number of iterations, say k, an iterate u^k and a family of u^j's with the desired property, all of them lying in a small neighborhood of u^k. This is one very nice feature of the bundle mechanism which is extensively described in [17, Volume II]. Moreover, it is a well known fact, called Caratheodory's theorem, that only $p = n+2$ subgradients are sufficient in the expression (II-C.1).

Set

$$X^*_\epsilon = \sum_{j=1}^{p} \alpha_j x^j {x^j}^t.$$

Then, we have the following result.

Proposition 3.1: For any $\epsilon > 0$, the matrix X^*_ϵ defined above satisfies

$$\mathrm{trace}(MX^*_\epsilon) \leq \mathrm{opt}(\mathrm{I-A.6}) - \mathcal{O}(\epsilon).$$

Proof: Let u^* be any minimizer of θ. Then, for each $j = 1, \ldots, p$, we have by the definition of the subdifferential

$$\theta(u^*) \geq \theta(u^j) + g_j^t(u^* - u^j).$$

But $\theta(u^j)$ is given by

$$\theta(u^j) = {x^j}^t\Big(D(u^j) - M - \frac{e^t u^j}{n+1} I\Big)x^j.$$

On the other hand, since ${x^j}^t x^j = n+1$,

$$\begin{aligned}
&{x^j}^t\Big(D(u^j) - M - \tfrac{e^t u^j}{n+1} I\Big)x^j \\
&= {x^j}^t M x^j + \textstyle\sum_{i=1}^{n+1} u_i {x_i^j}^2 - \sum_{i=1}^{n+1} u_i \\
&= {x^j}^t M x^j + \textstyle\sum_{i=1}^{n+1} u_i ({x_i^j}^2 - 1) \\
&= {x^j}^t M x^j + g_j^t u^j.
\end{aligned}$$

Thus, we obtain

$$\theta(u^*) \geq \mathrm{trace}(M{x^j}^t x^j) + g_j^t u^*$$

which implies, after multiplying by α_j and summing over $j = 1, \ldots, p$

$$\theta(u^*) \geq \mathrm{trace}(MX^*_\epsilon) + (\sum_{j=1}^p \alpha_j g_j)^t u^*.$$

Using Cauchy-Schwartz inequality, this gives

$$\theta(u^*) \geq \mathrm{trace}(MX^*_\epsilon) + \epsilon\|u^*\|.$$

Since the eigenvalue and the SDP relaxation have equal optimal values, we finally obtain

$$\mathrm{val}(\mathrm{II-A.1}) \geq \mathrm{trace}(MX^*_\epsilon) + \epsilon\|u^*\|$$

which implies the desired result. ∎

D. Comments

It is a common idea that the SDP relaxation contains more information than the eigenvalue relaxation. We hope that the results of this section managed to convince the reader that this is in fact not the case and a good approximate solution can be recovered quite easily using subgradient information at the optimum.

III. Recovering primal binary solutions

We now are in position to answer our main question of how to recover a satisfactory although suboptimal primal binary solution. The main result in this direction is Nesterov's bound which initially strongly relied on the SDP relaxation. The procedure is mainly based on Goemans and Williamson's randomized rounding algorithm. In this section, we explain how this randomized scheme can be recovered from the eigenvector information.

A. Goemans and Williamson's algorithm and Nesterov's bound

The method relies on the Cholesky factorization of the optimal solution X^* of the SDP relaxation,

$$X^* = V^t V.$$

From Theorem 2.1 we see that $V \in \mathbb{R}^{(n+1)\times r_{\max}}$ where $r_{\max}$ is the multiplicity of $\lambda_{\max}(A(u^*))$ at the chosen corresponding solution u^* of the eigenvalue relaxation. This factorization is important, since it allows to write $X^*_{ij} = v_i^t v_j$ where v_i is the transpose of i^{th} row vector of V. Let ξ be a random variable with uniform distribution on the unit sphere in $\mathbb{R}^{r_{\max}}$.

Procedure 1.1: **(Goemans and Williamson's algorithm)**

1. Find the Choleski factorization $X^* = V^t V$.

Let ζ be a random vector with uniform distribution on the unit sphere of $\mathcal{S}(0,1)$. The random cut is defined by

$$Z = V^t \zeta.$$

where the sign function is defined coordinate-wise.

2. Draw n samples from Z, say $z^1, \ldots, z^n$ and choose the sample giving the best value of ${z^i}^t M z^i$.

The key result is that, in average, the vector Z gives a good binary solution to the original problem. Since the best sample will have greater cut value than the average with overwhelming probability, the above procedure should work well. This is made precise by Nesterov's theorem.

Theorem 1.2: Define

$$f^* = \max_{x\in\mathbb{R}^{n+1}} x^t M x \text{ s.t. } x \in \{-1,1\}^{n+1}$$

and

$$f_* = \min_{x \in \mathbb{R}^{n+1}} x^t M x \text{ s.t. } x \in \{-1, 1\}^{n+1}$$

then, we have

$$\frac{f^* - E[z^t M z]}{f^* - f_*} \leq \frac{2}{\pi}.$$

This result is remarkable despite the fact that the bound $\frac{2}{\pi}$ is rather large. Indeed, the fact that there exists such a bound is known to be not true for any combinatorial optimization problem. Moreover, it can be expected that such a bound is far from reality in practical problems. For this reason, an important issue for future research is to study such type of bounds for particular subclasses of problems in hope of improving Nesterov's result.

B. *The eigenvector viewpoint*

The main drawback of the former presentation is that using the uniform variable ξ is quite hard to motivate from an optimization viewpoint. Let us take a slightly different perspective. Assume that we have a solution u^* of the eigenvalue relaxation. As before, let $E_{\max}$ be a matrix whose columns form an orthonormal bases of the eigenspace associated to $\lambda_{\max}(A(u^*))$. Moreover, we may require that

$$0 = \mathcal{A}^*(E_{\max} \Delta E_{\max}^t), \tag{III-B.1}$$

where Δ is some diagonal matrix with $\alpha = d(\Delta)$, $\alpha \geq 0$ and $\sum_{i=1}^{r_{\max}} \alpha_i = 1$. In the case where the multiplicity at the optimum is one, the optimal eigenbasis reduces to a unique vector and we saw in Proposition 2.4 that multiplying this vector by $n+1$ gives a binary solution. Now let us turn to the case where there are $r_{\max} > 1$ eigenvectors. To each unit norm eigenvector e^j, we associate a subgradient $g_j = [(n+1)(e_1^j)^2 - 1, \ldots, (n+1)(e_{n+1}^j)^2 - 1]^t$. Then, (III-B.1) implies that

$$0 = \sum_{j=1}^{r_{\max}} \alpha_j g_j. \tag{III-B.2}$$

Now one natural strategy might be the following: pick the best eigenvector, i.e. the eigenvector $\sqrt{n+1}e^{j_0}$ whose associated coefficient α_{j_0} in expression (III-B.2) is the *greatest* and round its coordinates to the nearest binary values. There is a second strategy : draw random linear combinations of the $\sqrt{n+1}e^j$'s giving preference to the components with higher associated coefficient in (III-B.2). This can be done by sampling vectors of the type

$$\sum_{j=1}^{r_{\max}} \zeta_j \sqrt{n+1} e^j$$

where the ζ_j's are independent random variables with distribution $\mathcal{N}(0, \alpha_j)$. For each sample, a feasible solution is obtained by rounding off the components to the nearest binary. We sum up this procedure as follows.

Procedure 2.1: **(Randomized algorithm based on optimal eigenvectors)** 1. Find the matrix $E_{\max}$ whose columns form an orthonormal eigenbasis associated to $\lambda_{\max}(A(u^*))$ such that (III-B.2) holds for some α_j's satisfying $\alpha \geq 0$ and $\sum_{j=1}^{r_{\max}} \alpha_j = 1$.

2. Let ζ be a random vector with distribution $\mathcal{N}(0, D(\alpha))$. The random cut is defined by

$$z = \sqrt{n+1} E_{\max} \chi.$$

3. Draw n samples from z, say $z^1, \ldots, z^n$ and choose the sample giving the best value of ${z^i}^t M z^i$.

The important result is that this second strategy is equivalent to Goemans and Williamson's randomized procedure.

Proposition 2.2: Procedure 2.1 is equivalent to Goemans and Williamson's algorithm.

Proof: Set $V = E_{\max} D(\alpha)^{\frac{1}{2}}$. Then Theorem 2.1 and equation III-B.2 imply that $X^* = VV^t$, thus retrieving the Cholesky factorization of X^*. Let $\xi = D(\alpha)^{-\frac{1}{2}} \chi$. It is clear that ξ has distribution $\mathcal{N}(0, I)$. This proves that the cut $\mathrm{sign}(z)$ obtained by Procedure 2.1 is equivalently the output of Goemans and Williamson's procedure. ∎

C. Comments and problems

The eigenvalue point of view allowed us to provide a reasonable explanation for taking a random cut using a uniformly distributed variable on the sphere in Goemans and Williamson's methodology. This procedure provides a relative approximation quality ratio via Nesterov's bound. However, as we shall see now, it is quite unnatural in the context of least squares estimation if one looks for an absolute approximation factor.

The proof of Nesterov's result heavily relies on the fact that

$$\mathrm{arcsin} X \succcurlyeq X \tag{III-C.1}$$

for all $X \succcurlyeq 0$ whose entree's absolute value is less than or equal to one. Then, it was shown by Nesterov that the expectation of a random binary solution obtained by the randomized algorithm is $\frac{2}{\pi}\mathrm{trace}(-M \mathrm{arcsin}(X^*))$, where X^* is a solution of the SDP relaxation. When M is negative semi-definite, we obtain using (III-C.1) that the expected value of a random binary solution is greater than or equal to $\frac{2}{\pi}\mathrm{trace}(-MX^*)$, which gives

$$\mathrm{opt}(\mathrm{II}-\mathrm{A.1}) \geq \max_{x \in \{-1,1\}^{n+1}} -x^t M x \geq \frac{2}{\pi} \mathrm{opt}(\mathrm{II}-\mathrm{A.1})$$

which proves the important fact that the optimum values admits a $\frac{2}{\pi}$ approximation factor. When M is positive definite, we obtain that the expected value of a random binary solution is less than or equal to $\frac{2}{\pi}\mathrm{trace}(-MX^*)$, which brings no information since we know that it is less than or equal to $\mathrm{trace}(-MX^*)$. Therefore, the usual proof does not apply when the matrix M is positive semidefinite. Moreover, we are not aware of any study of the approximation factor of the randomized procedure in the present case of interest.

IV. Two application examples

In this section, we provide some experimental results for the problems of image denoising and the problem of multiuser detection in CDMA systems.

A. Image denoising

The first set of simulations is devoted to the denoising problem, in which A is simply the identity matrix. This is the problem considered in [21], [13] and [16] for instance. The original binary image as 26 rows and 62 columns which gives a total number of 1612 variables.

For this problem, the penalization matrix P is chosen so as to smooth the image. This is achieved by requiring neighboring pixels to be similar in the sense that if i and j are indices of neighbor pixels, then, we would like the least square cost to be penalized by the quantity $|x_i - x_j|^2$. Thus, P is the matrix associated to the quadratic form

$$\sum_{i\sim j} |x_i - x_j|^2,$$

where $i \sim j$ denotes the property of being neighbor indices.

The experiments reported on below were performed for the case of quite noisy original images. The noise was taken additive, independent identically distributed and Gaussian $\mathcal{N}(0, 2)$ and was applied to the symmetrized image with pixel values in $\{-1, 1\}$. In order to show the influence of the smoothing parameter ν, we displayed the percentage of misspecified bits vs values of ν. The recovered image is the one with the choice of ν giving the best percentage of bits recovered.

We found the results very encouraging. Indeed, even when the observed image is very noisy, we still recover an image which is readable. This suggested that an appropriate postprocessing might easily allow to recover the original written word of sentence, by comparing the letters to a given dictionary. The only problem is the choice of an appropriate value of ν. This issue is a problem that has been addressed in the statistics literature, where a Bayesian approach motivates different possible decisions. Cross validation can also be used. We will not discuss this problem here. Instead, it seems reasonable to argue that the choice of ν can just be made *a posteriori* since it consists in tuning the method until a satisfactory solution is obtained. This reduces the hard combinatorial initial problem to a simpler one parameter knobing procedure. The displayed experiment and the numerous simulations not presented here confirm that robust intervals for the values of ν are not very difficult to identify in practice.

B. Multiuser detection in CDMA systems

This problem was studied by [22] using the maximum likelihood approach. As we will see, the resulting optimization problem is of the same form as the binary least squares problem. The main difference here is that $A \neq I$ and $P = 0$.

A synchronous K users DS-CDMA system is considered with a common single path additive white Gaussian noise (AWGN) channel. The signature waveform of the kth user is denoted by $s_k(t)$, a function taking nonzero values in $[0, T]$ and being equal to zero outside this interval, and x_k is the information bit transmitted by user k. The overall recieved signal is therefore of the form

$$y(t) = \sum_{k=1}^{K} a_k x_k s_k(t) + n(t)$$

where a_k is the amplitude of the kth user's signal and $n(t)$ is an additive white Gaussian white noise with zero mean and variance σ^2. The signal y is then filtered using a bank of K matched filters. The output of the kth matched filter is given by

$$y_k = \int_0^T y(t) s_k(t) dt.$$

In matrix form, this can be written

$$y = RAx + \nu$$

where $y = [y_1, \ldots, y_k]^t$, R is the correlation matrix whose components are given by $R_{ij} = \int_0^T s_i(t) s_j(t) dt$, $A = D(a)$ and ν is the vector with components $\nu_k = \int_0^T n(t) s_j(t) dt$.

Since the gaussian vector has a correlation matrix equal to $\sigma^2 R$, the ML estimator is obtained by simply solving the following combinatorial optimization problem.

$$\min_{x \in \mathbb{R}^n} x^t ARAx - 2y^t Ax$$

(IV-B.1)

$$\text{s.t. } x_i \in \{-1, 1\}, \quad i = 1, \ldots, K.$$

The SDP approach seems to have been first applied for the DS-CDMA detection problem in [24]. Since then numerous contributions have appeared using the SDR and comparing it to other methods as in [26] and [27]. Extension to M-ary phase shift keying symbol constellations is proposed in [28]. The issue of accelerating the speed of the method is addressed in [29]. However, as for the former problem, the main drawback of the standard primal semidefinite relaxation is that the size of the problem is greatly increased by using $K \times K$ matrices instead of vectors of size K. In order to overcome this problem, a better approach using semidefinite programming duality was recently proposed in [30]. From this respect, the eigenvalue relaxation seems to be equally well suited to this problem. In order to verify this point, we performed Monte Carlo simulations over 1000 random problems for a number a users varying from 10 to 35. These computational experiments are reported in Figure V where the number of users is on the x-axis and the average computation time is on the y-axis. The computations where performed using the Scilab software [31]. The SDP solver called *Semidef* interfaces Boyd and Vandenberghe's sp.c program. The eigenvalue relaxation was solved using the solver *Optim* with the ”nd” option for possibly nondifferentiable costs as is the case here. The curves in Figure V interpolate the average computation times for messages taken to be sequences of uniform and independant variables taking values in $\{0, 1\}$ vs. the number of users. The curve with dashed style is for the results of the SDP relaxation while the curve with plain style is for the eigenvalue relaxation. Our computations suggest that the eigenvalue relaxation has lower complexity growth as the number of users increases exactly as expected.

V. Conclusion

In this paper, we surveyed the main properties of the eigenvalue relaxation for binary least squares problem. A full connection with the standard SDP relaxation was presented and we showed how to recover a solution of the Semi-Definite program from the solution of the eigenvalue minimization problem. Although the original binary least squares problem is $\mathcal{NP}$-hard, the randomized procedure adapted from Goemans and Williamson's allowed to recover binary solutions with garanteed relative approximation ratio. Two applications were presented: binary image denoising and detection in multiuser CDMA systems. In the first applications we showed that the penalized binary least squares approach could be successfuly used in such imaging problems. In the second application, SDP was already known to perform well and we showed that the eigenvalue approach could outperform the SDP relaxation in terms of average computational complexity.

References

[1] Boyd, Stephen and Vandenberghe, Lieven *Convex optimization*. Cambridge University Press, Cambridge, 2004

[2] Luo, Zhi-Quan, *Applications of convex optimization in signal processing and digital communication*. ISMP, 2003 (Copenhagen). Math. Program. 97 (2003), no. 1-2, Ser. B, 177–207.

[3] Lemar´echal, Claude and Oustry, Franois *SDP relaxations in combinatorial optimization from a Lagrangian viewpoint*. Advances in convex analysis and global optimization (Pythagorion, 2000), 119–134, Nonconvex Optim. Appl., 54, Kluwer Acad. Publ., Dordrecht, 2001.

[4] Wolkowicz, Henry and Anjos, Miguel F. *Semidefinite programming for discrete optimization and matrix completion problems*. Workshop on Discrete Optimization, DO'99 (Piscataway, NJ). Discrete Appl. Math. 123 (2002), no. 1-3, 513–577.

[5] Ben-Tal, Aharon and Nemirovski, Arkadi *Lectures on modern convex optimization. Analysis, algorithms, and engineering applications*. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2001.

[6] Poljak, S., Rendl, F. and Wolkowicz, H. *A recipe for semidefinite relaxation for* $(0,1)$*-quadratic programming*. J. Global Optim. 7 (1995), no. 1, 51–73.

[7] Oustry, Franois *A second-order bundle method to minimize the maximum eigenvalue function*. Math. Program. 89 (2000), no. 1, Ser. A, 1–33

[8] Helmberg, Christoph and Oustry, Francois *Bundle methods to minimize the maximum eigenvalue function*. Handbook of semidefi nite programming, 307–337, Internat. Ser. Oper. Res. Management Sci., 27, Kluwer Acad. Publ., Boston, MA, 2000

[9] Helmberg, C. and Rendl, F. *A spectral bundle method for semidefinite programming*. SIAM J. Optim. 10 (2000), no. 3, 673–696

[10] Nesterov, Yu. *Semidefinite relaxation and nonconvex quadratic optimization*. Optim. Methods Softw. 9 (1998), no. 1-3, 141–160.

[11] Nesterov, Yuri, Wolkowicz, Henry and Ye, Yinyu *Semidefinite programming relaxations of nonconvex quadratic optimization*. Handbook of semidefi nite programming, 361–419, Internat. Ser. Oper. Res. Management Sci., 27, Kluwer Acad. Publ., Boston, MA, 2000.

[12] Goemans, Michel X. and Williamson, David P. *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*. J. Assoc. Comput. Mach. 42 (1995).

[13] Gibbs, Alison L. *Bounding the convergence time of the Gibbs sampler in Bayesian image restoration*. Biometrika 87 (2000), no. 4, 749–766.

[14] Chrtien, Stphane; Corset, Franck *Least squares reconstruction of binary images using eigenvalue optimization*. COMPSTAT 2002 (Berlin), 419–424, Physica, Heidelberg, 2002.

[15] Keuchel, J.; Schnorr, C.; Schellewald, C.; Cremers, D.; *Binary partitioning, perceptual grouping, and restoration with semidefinite programming*, Pattern Analysis and Machine Intelligence, IEEE Transactions on **25**, (2003), no. 11, 1364–1379.

[16] Besag, Julian *On the statistical analysis of dirty pictures.* J. Roy. Statist. Soc. Ser. B 48 (1986), no. 3, 259–302.

[17] Hiriart-Urruty, J.-B.; Lemarchal, C. *Convex analysis and minimization algorithms. II. Advanced theory and bundle methods.* Grundlehren der Mathematischen Wissenschaften, 306. Springer-Verlag, Berlin, 1993.

[18] Delorme, C.; Poljak, S. *Laplacian eigenvalues and the maximum cut problem.* Math. Programming 62 (1993), no. 3, Ser. A, 557–574.

[19] Poljak, Svatopluk and Wolkowicz, Henry Convex *relaxations of* $(0,1)$*-quadratic programming*. Math. Oper. Res. 20 (1995), no. 3, 550–561.

[20] Poljak, Svatopluk and Rendl, Franz *Nonpolyhedral relaxations of graph-bisection problems*. SIAM J. Optim. 5 (1995), no. 3, 467–487.

[21] Nikolova M., *Estimation of binary images using convex criteria*, Proc. of IEEE Int. Conf. on Image Processing, Oct. 1998.

[22] Verdù, Sergio *Minimum probability of error for asynchronous Gaussian multiple-access channels*. IEEE Trans. Inform. Theory 32 (1986), no. 1, 85–96.

[23] Goemans, Michel X.and Williamson, David P. *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*. J. Assoc. Comput. Mach. 42 (1995), no. 6, 1115–1145.

[24] Peng Hui Tan and Lars K. Rasmussen, *The application of semidefinite programming for detection in CDMA*, IEEE J. Select. Areas in Comm. 18 (2001), no. 8, 1442–1448.

[25] Garey, M. R. and Johnson, D. S. *Computers and intractability. A guide to the theory of NP-completeness*. A Series of Books in the Mathematical Sciences. W. H. Freeman and Co., San Francisco, Calif., 1979.

[26] Fumihiro Hasegawa, Jie Luo, Krishna R. Pattipati, Peter Willett and David Pham, *Speed and accuracy comparison of techniques for multiuser detection in synchronous CDMA*, IEEE Trans. Comm. 52 (2004), no. 4, 540–545.

[27] Peng Hui Tan and Lars K. Rasmussen, *Multiuser detection in CDMA–A comparison of relaxations, exact and Heuristic search methods*, IEEE Trans. Wireless Comm. 3, (2004), no. 5, 1802–1809.

[28] Wing-Kin Ma, Pak-Chung Ching and Zhi Ding, *Semidefinite relaxation based multiuser detection for M-ary PSK multiuser systems*, IEEE Trans. Sig. Proc. 52, (2004), no. 10, 2862–2872.

[29] Moussa Abdi, Hassan El Nahas, Alexandre Jard and Eric Moulines, *Semidefinite positive relaxation of the maximum likelihood criterion applied to the multiuser detection in a CDMA context*, IEEE Sig. Proc. Letters, 9, (2002), no. 6, 165–167.

[30] X. M. Wang, W. S. Lu and A. Antoniou, *A near optimal multiuser detector for DS-CDMA systems using semidefinite programming relaxation*, IEEE Trans. Sig. Proc. 51, (2003), no. 9, 2446–2450.

[31] http://www.scilab.org

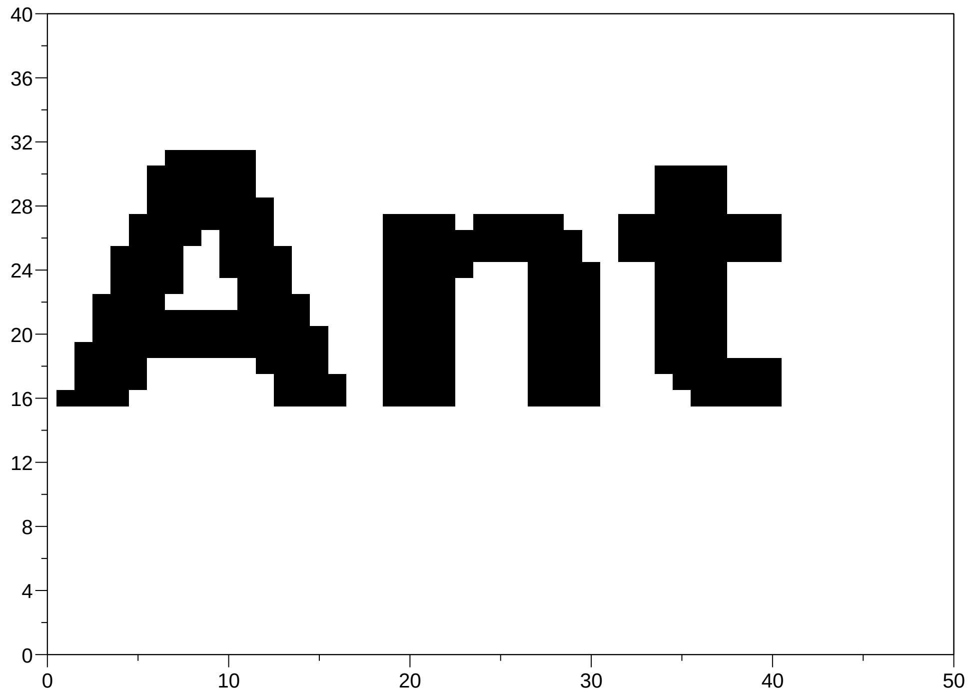

Fig. 1. Original image

Fig. 2. Noisy image: i.i.d. $\mathcal{N}(0, 2)$

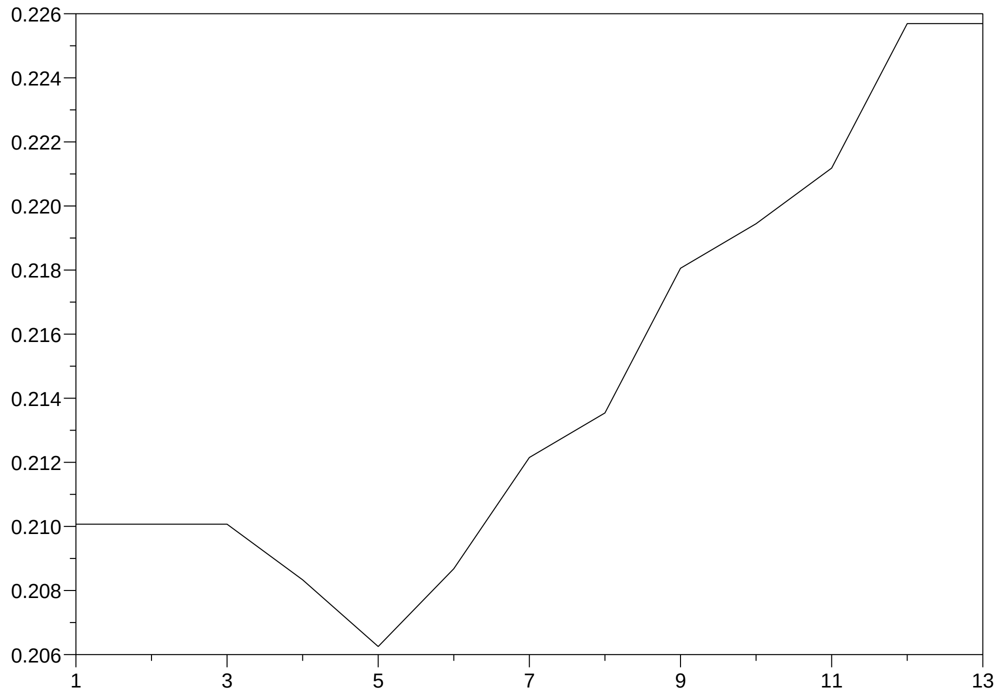

Fig. 3. Percentage of misspecified bits v.s. ν

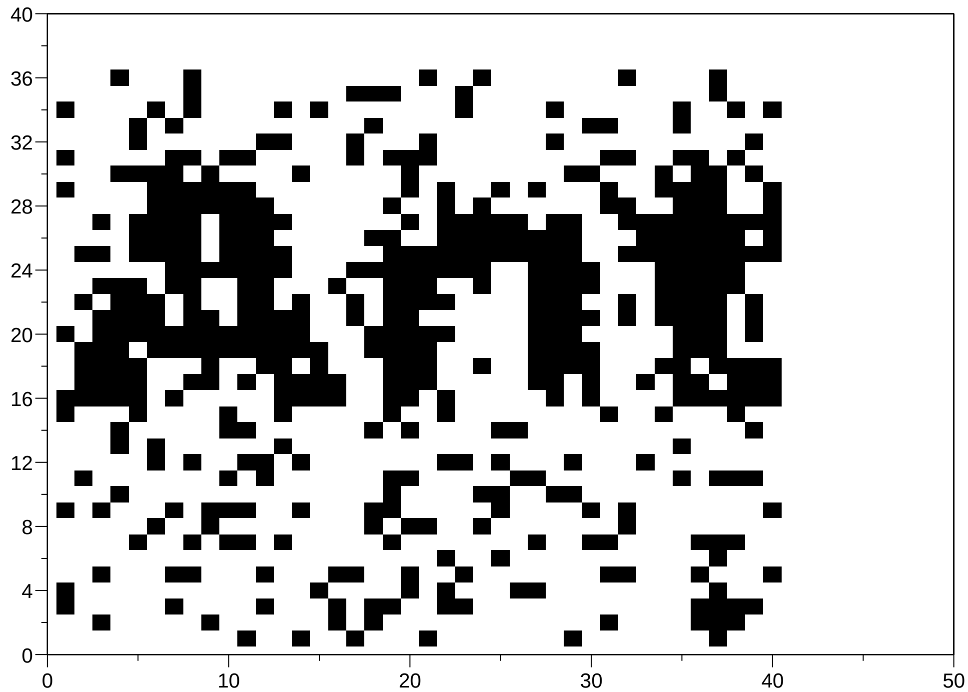

Fig. 4. Recovered image

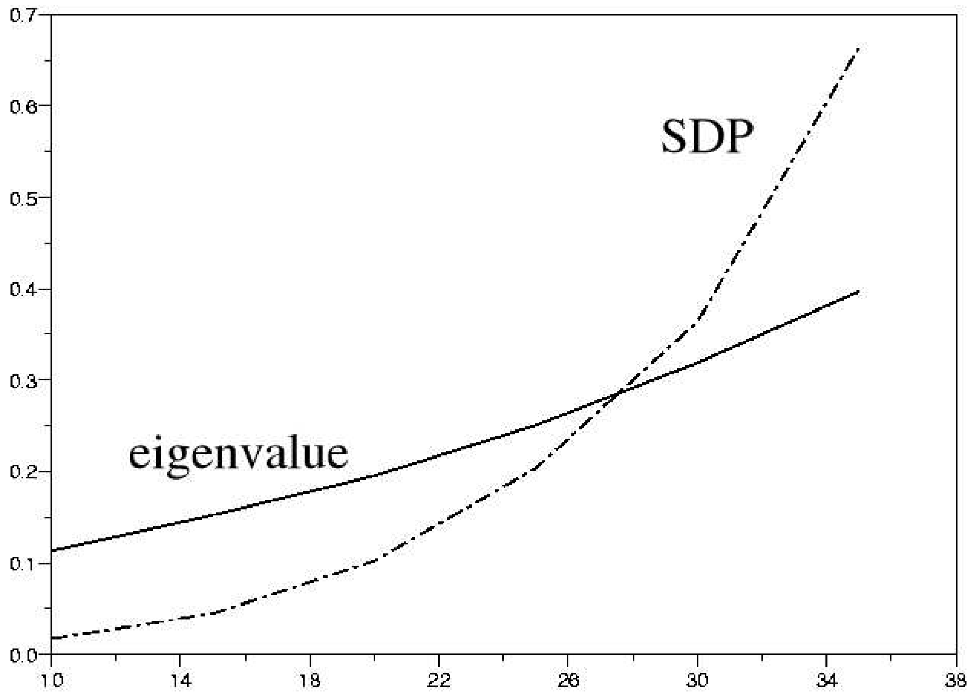

Fig. 5. Comparison of SDP and eigenvalue relaxations for CDMA multiuser detection