

# SPARSITY MAXIMIZATION UNDER A QUADRATIC CONSTRAINT WITH APPLICATIONS IN FILTER DESIGN

*Dennis Wei and Alan V. Oppenheim*

Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science  
77 Massachusetts Avenue, Cambridge, MA 02139, USA

## ABSTRACT

This paper considers two problems in sparse filter design, the first involving a least-squares constraint on the frequency response, and the second a constraint on signal-to-noise ratio relevant to signal detection. It is shown that both problems can be recast as the minimization of the number of non-zero elements in a vector subject to a quadratic constraint. A solution is obtained for the case in which the matrix in the quadratic constraint is diagonal. For the more difficult non-diagonal case, a relaxation based on the substitution of a diagonal matrix is developed. Numerical simulations show that this diagonal relaxation is tighter than a linear relaxation under a wide range of conditions. The diagonal relaxation is therefore a promising candidate for inclusion in branch-and-bound algorithms.

*Index Terms*— Sparse filters, least squares methods, signal detection, relaxation methods

## 1. INTRODUCTION

In the efficient implementation of discrete-time filters, it is often desirable to have filters with fewer non-zero coefficients, i.e., sparse filters, as a means of reducing the costs of implementation, whether in the form of computation, hardware, or power consumption. The design of sparse filters under a Chebyshev error criterion in the frequency domain has been examined from a variety of perspectives, including integer programming [1] and heuristic approaches [2–4]. In comparison, the case of a weighted least-squares criterion has not received much attention. As discussed in [5], a weighted least-squares error metric is commonly employed as an alternative to a Chebyshev metric because of greater tractability and an association with signal energy or power.

The approximation of desired frequency responses constitutes one class of filter design problems. Another important context in which filters are used is in the detection of signals in noisy environments, where the objective of filtering is to increase the probability of detection. A widely used measure of performance in detection is the signal-to-noise ratio (SNR) of the filter output. It is well-known that the SNR is monotonically related to the probability of detection in the case of Gaussian noise [6].

In this paper, we consider two problems in sparse filter design, the first involving a weighted least-squares constraint on the frequency response, and the second a constraint on SNR. In Section 2 it is shown that both problems can be formulated in terms of a single quadratic constraint, specifically of the form

$$(\mathbf{b} - \mathbf{c})^T \mathbf{Q} (\mathbf{b} - \mathbf{c}) \leq \gamma, \quad (1)$$

---

This work was supported in part by the Texas Instruments Leadership University Program, BAE Systems PO 112991, Lincoln Laboratory PO 3077828, and the MIT William Asbjornsen Albert Memorial Fellowship.

where  $\mathbf{b}$  is a vector of coefficients,  $\mathbf{Q}$  is a symmetric positive definite matrix,  $\mathbf{c}$  is a vector of the same length as  $\mathbf{b}$ , and  $\gamma > 0$ . This formulation allows for a unified approach to solving not only the two problems stated but also other problems involving performance criteria that can be expressed in the form of (1). One example is the criterion of mean squared error used in estimation, which forms the basis for such techniques as linear prediction, Wiener filtering, and least-mean-square adaptive filtering [7].

The design of sparse filters is related to but distinct from the problem of obtaining sparse solutions to underdetermined linear equations, which occurs for example in compressive sensing [8]. Although a quadratic constraint is sometimes also used in the underdetermined equations setting, for example to model the presence of noise, the matrix corresponding to  $\mathbf{Q}$  is rank-deficient and consequently the set of feasible solutions is qualitatively different from that specified by (1).

In Sections 3–5, we concentrate on solving the problem of sparse design subject to (1). When  $\mathbf{Q}$  is a diagonal matrix, a maximally sparse design can be easily obtained as described in Section 3. In most other cases, however, the problem is much more difficult and no polynomial-time algorithm is known. Our focus in this paper is on developing relaxations of the problem that are efficiently solvable and lead to strong lower bounds on the true optimal cost, for example within a factor close to unity. Such relaxations are potentially useful as part of a branch-and-bound procedure for solving the problem exactly and are the basis of future work. In Section 4, we discuss the technique of linear relaxation, while in Section 5, we introduce an alternative method, referred to as diagonal relaxation, in which  $\mathbf{Q}$  is replaced by a diagonal matrix. Numerical experiments presented in Section 6 demonstrate that the lower bounds resulting from diagonal relaxations are often significantly tighter than those from linear relaxations.

## 2. FORMULATION OF SPARSE FILTER DESIGN

In this section, we formulate the problems of sparse filter design with a weighted least-squares error criterion and sparse filter design for signal detection under a common framework corresponding to

$$\min_{\mathbf{b}} \|\mathbf{b}\|_0 \quad \text{s.t.} \quad (\mathbf{b} - \mathbf{c})^T \mathbf{Q} (\mathbf{b} - \mathbf{c}) \leq \gamma, \quad (2)$$

where the zero-norm notation  $\|\mathbf{b}\|_0$  refers to the number of non-zero elements in  $\mathbf{b}$ . The constraint in (2) may be interpreted geometrically as specifying an ellipsoid centered at  $\mathbf{c}$ . The eigenvectors and eigenvalues of  $\mathbf{Q}$  determine the orientation and relative lengths of the axes of the ellipsoid while  $\gamma$  determines its absolute size. An alternative form for (1) that is used in this section is

$$\mathbf{b}^T \mathbf{Q} \mathbf{b} - 2\mathbf{f}^T \mathbf{b} \leq \beta \quad (3)$$

with  $\mathbf{f} = \mathbf{Q}\mathbf{c}$  and  $\beta = \gamma - \mathbf{c}^T \mathbf{Q}\mathbf{c}$ .

## 2.1. Weighted least-squares filter design

Consider the design of a causal FIR filter of length  $N$  with coefficients  $b_n$  and frequency response

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} b_n e^{-j\omega n} \quad (4)$$

chosen to meet a squared-error constraint:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) |H(e^{j\omega}) - D(e^{j\omega})|^2 d\omega \leq \delta, \quad (5)$$

where  $D(e^{j\omega})$  is the desired frequency response,  $\delta$  is the desired tolerance, and  $W(\omega)$  is a non-negative and even-symmetric weighting function. The number of non-zero coefficients is to be minimized. Substituting (4) into (5), expanding, and comparing the result with (3), we can identify

$$Q_{mn} = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) \cos((m-n)\omega) d\omega, \quad (6a)$$

$$f_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) D(e^{j\omega}) e^{j\omega n} d\omega, \quad (6b)$$

$$\beta = \delta - \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) |D(e^{j\omega})|^2 d\omega, \quad (6c)$$

where  $m$  and  $n$  range from 0 to  $N-1$ . Equation (6a) defines a positive definite matrix as long as  $W(\omega)$  is non-zero over some interval.

## 2.2. Signal detection

The design of sparse filters for use in signal detection can also be formulated as in (2). We assume that a signal  $s[n]$  is to be detected in the presence of stationary additive noise  $\eta[n]$  having zero mean and autocorrelation  $\phi_{\eta\eta}[m]$ . The received signal is processed with a filter of length  $N$  and sampled at  $n = N-1$ , yielding

$$y[N-1] = \sum_{n=0}^{N-1} b_n (s[N-1-n] + \eta[N-1-n])$$

when the signal is present. The filter coefficients  $b_n$  are chosen such that the SNR is greater than a pre-specified threshold  $\rho$ , where the SNR is defined as the ratio of the mean of  $y[N-1]$  given that the signal is present to the standard deviation of  $y[N-1]$ . By defining  $\mathbf{s} \in \mathbb{R}^N$  and  $\mathbf{R} \in \mathbb{R}^{N \times N}$  according to  $s_n = s[N-1-n]$  and  $R_{mn} = \phi_{\eta\eta}[|m-n|]$ , the problem of sparse design can be expressed as

$$\min_{\mathbf{b}} \|\mathbf{b}\|_0 \quad \text{s.t.} \quad \frac{\mathbf{s}^T \mathbf{b}}{\sqrt{\mathbf{b}^T \mathbf{R} \mathbf{b}}} \geq \rho. \quad (7)$$

While the constraint in (7) cannot be rewritten directly in the form of (3), we show that problems (7) and (2) are nonetheless equivalent. To establish the equivalence, we determine whether feasible solutions to (2) and (7) exist when an arbitrarily chosen subset of coefficients  $b_n$ , represented by the index set  $\mathcal{Z}$ , is constrained to have value zero. Given  $b_n = 0$  for  $n \in \mathcal{Z}$  and with  $\mathcal{Y}$  denoting the complement of  $\mathcal{Z}$ , constraint (3) becomes

$$\mathbf{b}_{\mathcal{Y}}^T \mathbf{Q}_{\mathcal{Y}\mathcal{Y}} \mathbf{b}_{\mathcal{Y}} - 2\mathbf{f}_{\mathcal{Y}}^T \mathbf{b}_{\mathcal{Y}} \leq \beta, \quad (8)$$

where  $\mathbf{b}_{\mathcal{Y}}$  is the  $|\mathcal{Y}|$ -dimensional vector formed from the entries of  $\mathbf{b}$  indexed by  $\mathcal{Y}$  (similarly for  $\mathbf{f}_{\mathcal{Y}}$ ), and  $\mathbf{Q}_{\mathcal{Y}\mathcal{Y}}$  is the  $|\mathcal{Y}| \times |\mathcal{Y}|$  matrix formed from the rows and columns of  $\mathbf{Q}$  indexed by  $\mathcal{Y}$ . We consider minimizing the left-hand side of (8) with respect to  $\mathbf{b}_{\mathcal{Y}}$ . If this minimum is greater than  $\beta$ , then (8) cannot be satisfied for any value of  $\mathbf{b}_{\mathcal{Y}}$  and a feasible solution with  $b_n = 0$ ,  $n \in \mathcal{Z}$  cannot exist. It is straightforward to show by differentiation that the left side is minimized when  $\mathbf{b}_{\mathcal{Y}} = (\mathbf{Q}_{\mathcal{Y}\mathcal{Y}})^{-1} \mathbf{f}_{\mathcal{Y}}$ . Consequently the condition for feasibility is

$$-\mathbf{f}_{\mathcal{Y}}^T (\mathbf{Q}_{\mathcal{Y}\mathcal{Y}})^{-1} \mathbf{f}_{\mathcal{Y}} \leq \beta. \quad (9)$$

We refer to an index set  $\mathcal{Y}$  (equivalently  $\mathcal{Z}$ ) as being feasible if (9) is satisfied. Similarly in the case of problem (7),  $\mathcal{Y}$  is feasible only if the modified constraint

$$\frac{\mathbf{s}_{\mathcal{Y}}^T \mathbf{b}_{\mathcal{Y}}}{\sqrt{\mathbf{b}_{\mathcal{Y}}^T \mathbf{R}_{\mathcal{Y}\mathcal{Y}} \mathbf{b}_{\mathcal{Y}}}} \geq \rho$$

is satisfied when the left-hand side is maximized. The maximizing values of  $\mathbf{b}_{\mathcal{Y}}$  correspond to a whitened matched filter for the partial signal  $\mathbf{s}_{\mathcal{Y}}$  and are proportional to  $(\mathbf{R}_{\mathcal{Y}\mathcal{Y}})^{-1} \mathbf{s}_{\mathcal{Y}}$ . The resulting feasibility condition is  $\sqrt{\mathbf{s}_{\mathcal{Y}}^T (\mathbf{R}_{\mathcal{Y}\mathcal{Y}})^{-1} \mathbf{s}_{\mathcal{Y}}} \geq \rho$ , or after squaring,

$$\mathbf{s}_{\mathcal{Y}}^T (\mathbf{R}_{\mathcal{Y}\mathcal{Y}})^{-1} \mathbf{s}_{\mathcal{Y}} \geq \rho^2. \quad (10)$$

Condition (10) is identical to (9) for all  $\mathcal{Y}$  with the identifications  $\mathbf{Q} = \mathbf{R}$ ,  $\mathbf{f} = \mathbf{s}$ , and  $\beta = -\rho^2$ . It follows that an index set  $\mathcal{Y}$  is feasible for problem (7) exactly when it is feasible for problem (2), and therefore the optimal index sets for (2) and (7) coincide.

Stationarity is not a necessary condition for equivalence with problem (2). In the absence of stationarity, however, the matrix  $\mathbf{R}$  may vary with time, resulting in a succession of instances of problem (7).

## 3. THE CASE OF DIAGONAL $\mathbf{Q}$

We now shift our focus to solving problem (2) and developing relaxations. This section addresses the case in which the matrix  $\mathbf{Q}$  is diagonal. A diagonal  $\mathbf{Q}$  matrix can arise in least-squares filter design if the weighting in (5) is uniform. In the case of detection,  $\mathbf{R}$  and hence  $\mathbf{Q}$  are diagonal if the noise  $\eta[n]$  is white.

With  $\mathbf{Q}$  diagonal, problem (2) becomes

$$\min_{\mathbf{b}} \|\mathbf{b}\|_0 \quad \text{s.t.} \quad \sum_{n=0}^{N-1} Q_{nn} (b_n - c_n)^2 \leq \gamma. \quad (11)$$

To solve (11), we first determine whether it is feasible to have a solution with  $K$  zero-valued elements. Extending the argument made in Section 2.2, if the constraint in (11) is not met when the left-hand side is minimized over all  $\mathbf{b}$  with  $K$  zero-valued entries, then it cannot be met for any choice of  $\mathbf{b}$  with  $K$  zero-valued entries. The minimum is achieved by setting  $b_n = 0$  for  $n$  corresponding to the  $K$  smallest values of  $Q_{nn} c_n^2$  and  $b_n = c_n$  otherwise. This yields the feasibility condition

$$\Sigma_K(\{Q_{nn} c_n^2\}) \leq \gamma, \quad (12)$$

where  $\Sigma_K(\{Q_{nn} c_n^2\})$  denotes the sum of the  $K$  smallest  $Q_{nn} c_n^2$ . A similarly compact condition is not possible in the case of non-diagonal  $\mathbf{Q}$  with no special structure. Based on (9), the corresponding condition is

$$\min_{|\mathcal{Y}|=N-K} \left\{ -\mathbf{f}_{\mathcal{Y}}^T (\mathbf{Q}_{\mathcal{Y}\mathcal{Y}})^{-1} \mathbf{f}_{\mathcal{Y}} \right\} \leq \beta.$$

The number of sets  $\mathcal{Y}$  of size  $N - K$  is  $\binom{N}{K}$ , which can be very large, and an efficient way of minimizing over all choices of  $\mathcal{Y}$  is not apparent.

Problem (11) can be solved by checking the condition in (12) for successively increasing values of  $K$  starting with  $K = 0$ . The minimum zero-norm is given by  $N - K^*$ , where  $K^*$  is the largest value of  $K$  for which (12) holds. One particular optimal solution results from setting  $b_n = c_n$  for  $n$  corresponding to the  $N - K^*$  largest  $Q_{nn}c_n^2$ , and  $b_n = 0$  otherwise. This solution has an intuitive interpretation in the context of detection in white stationary noise. In this case, we have  $\mathbf{Q} = \mathbf{R} \propto \mathbf{I}$  and  $c_n \propto f_n = s[n]$ , and therefore the solution is to match only the  $N - K^*$  largest-magnitude values of the signal  $s[n]$ . If  $\eta[n]$  is white but non-stationary,  $\mathbf{Q}$  remains diagonal and the solution takes into account any weighting due to a time-varying variance.

#### 4. LINEAR RELAXATION

In the remainder of the paper, we focus on the case of non-diagonal  $\mathbf{Q}$  for which an efficient solution to (2) is not available. In this section, we derive a linear relaxation of (2) after first reformulating it as a mixed integer optimization problem. Toward this end, we express each coefficient  $b_n$  in terms of its positive and negative parts as  $b_n = b_n^+ - b_n^-$ , where  $b_n^+$  and  $b_n^-$  are non-negative and at least one is equal to zero. Each pair  $b_n^+, b_n^-$  is assigned a corresponding pair of binary-valued indicator variables  $i_n^+, i_n^-$  with the property that  $i_n^\pm = 0$  if  $b_n^\pm = 0$  and  $i_n^\pm = 1$  otherwise. Hence problem (2) is equivalent to

$$\begin{aligned} \min_{\mathbf{b}^+, \mathbf{b}^-, \mathbf{i}^+, \mathbf{i}^-} & \sum_{n=0}^{N-1} (i_n^+ + i_n^-) \\ \text{s.t.} & (\mathbf{b}^+ - \mathbf{b}^- - \mathbf{c})^T \mathbf{Q} (\mathbf{b}^+ - \mathbf{b}^- - \mathbf{c}) \leq \gamma, \\ & 0 \leq b_n^+ \leq B_n^+ i_n^+, \quad 0 \leq b_n^- \leq B_n^- i_n^- \quad \forall n, \\ & i_n^+ \in \{0, 1\}, \quad i_n^- \in \{0, 1\} \quad \forall n. \end{aligned} \quad (13)$$

The first constraint is the quadratic constraint (1) rewritten in terms of  $\mathbf{b}^+$  and  $\mathbf{b}^-$ . The second line of constraints ensures that  $i_n^+$  and  $i_n^-$  behave as indicator variables. In addition, an optimal solution to (13) must have at least one of  $b_n^+, b_n^-$  equal to zero for every  $n$ , as otherwise both could be decreased by  $\min\{b_n^+, b_n^-\}$  without affecting the quadratic constraint while allowing one of  $i_n^+, i_n^-$  to be decreased to zero.

The positive constants  $B_n^+$  and  $B_n^-$  appearing in (13) must be large enough so that the set of feasible vectors  $\mathbf{b}$  is unchanged from that in (2). Specifically, this requires

$$\begin{aligned} B_n^+ & \geq \max \{b_n : (\mathbf{b} - \mathbf{c})^T \mathbf{Q} (\mathbf{b} - \mathbf{c}) \leq \gamma\} \\ & = \sqrt{\gamma (\mathbf{Q}^{-1})_{nn}} + c_n, \\ B_n^- & \geq -\min \{b_n : (\mathbf{b} - \mathbf{c})^T \mathbf{Q} (\mathbf{b} - \mathbf{c}) \leq \gamma\} \\ & = \sqrt{\gamma (\mathbf{Q}^{-1})_{nn}} - c_n. \end{aligned} \quad (14a) \quad (14b)$$

We assume without loss of generality that it is feasible for each  $b_n$  to take a value of zero, and hence  $B_n^+$  and  $B_n^-$  are non-negative. The closed-form solutions to the optimization problems in (14) can be derived from the associated Karush-Kuhn-Tucker conditions [9].

A linear relaxation of (13) is obtained by allowing  $i_n^+$  and  $i_n^-$  to take on a continuous range of values between 0 and 1. The resulting

problem may be simplified to

$$\begin{aligned} \min_{\mathbf{b}^+, \mathbf{b}^-} & \sum_{n=0}^{N-1} \left( \frac{b_n^+}{B_n^+} + \frac{b_n^-}{B_n^-} \right) \\ \text{s.t.} & (\mathbf{b}^+ - \mathbf{b}^- - \mathbf{c})^T \mathbf{Q} (\mathbf{b}^+ - \mathbf{b}^- - \mathbf{c}) \leq \gamma, \\ & \mathbf{b}^+ \geq \mathbf{0}, \quad \mathbf{b}^- \geq \mathbf{0}. \end{aligned} \quad (15)$$

Thus the linear relaxation in (15) is a quadratically constrained linear program and its optimal value is a lower bound on the optimal value of (13). More precisely, since the optimal value of (13) must be an integer, the ceiling of the optimal value of (15) is also a lower bound. Note also that the optimal value of (15) is at its highest when  $B_n^+$  and  $B_n^-$  are set to their minimal values as given in (14).

#### 5. DIAGONAL RELAXATION

As an alternative to linear relaxations, this section discusses relaxations of problem (2) in which the matrix  $\mathbf{Q}$  is replaced by a positive definite diagonal matrix  $\mathbf{D}$ , an approach we refer to as diagonal relaxation. The quadratic constraint (1) is changed to

$$(\mathbf{b} - \mathbf{c})^T \mathbf{D} (\mathbf{b} - \mathbf{c}) = \sum_{n=0}^{N-1} D_{nn} (b_n - c_n)^2 \leq \gamma. \quad (16)$$

As seen in Section 3, the problem of sparse design is straightforward in the diagonal case, thus making it attractive as a relaxation of the problem when  $\mathbf{Q}$  is non-diagonal.

Geometrically, constraint (16) corresponds to an ellipsoid with axes that are aligned with the coordinate axes. Since the relaxation is intended to provide a lower bound for the original problem, we require that this axis-aligned ellipsoid enclose the ellipsoid specified by (1). It can be shown that the nesting of the ellipsoids is equivalent to  $\mathbf{Q} - \mathbf{D}$  being positive semidefinite, which we write as  $\mathbf{Q} - \mathbf{D} \succeq \mathbf{0}$  or  $\mathbf{Q} \succeq \mathbf{D}$ . Because of symmetry, the two ellipsoids can be made concentric without any loss in the quality of the relaxation.

For every  $\mathbf{D}$  satisfying  $\mathbf{0} \preceq \mathbf{D} \preceq \mathbf{Q}$ , minimizing  $\|\mathbf{b}\|_0$  subject to (16) results in a lower bound for problem (2). Thus the set of diagonal relaxations is parameterized by  $\mathbf{D}$ . To determine the tightest diagonal relaxation possible, i.e., a matrix  $\mathbf{D}^*$  such that the minimum zero-norm associated with  $\mathbf{D}^*$  is maximal, the following optimization problem is solved starting with  $K = 0$ :

$$\max_{\mathbf{D}} \Sigma_K(\{D_{nn}c_n^2\}) \quad \text{s.t.} \quad \mathbf{0} \preceq \mathbf{D} \preceq \mathbf{Q}, \quad \mathbf{D} \text{ diagonal.} \quad (17)$$

If the optimal value of (17) is less than or equal to  $\gamma$ , then the condition in (12) holds for every  $\mathbf{D}$  satisfying the constraints in (17). As argued in Section 3, it follows that a feasible solution  $\mathbf{b}$  with  $K$  zero-valued elements exists for every such  $\mathbf{D}$ . We conclude that no diagonal relaxation can give a minimum zero-norm greater than  $N - K$ . The value of  $K$  is then incremented by 1 and (17) is resolved. If on the other hand the optimal value of (17) is greater than  $\gamma$  for some  $K = K^* + 1$ , then there exists a  $\mathbf{D}^*$  for which it is not feasible to have a solution with  $K^* + 1$  zero elements. When combined with the conclusions drawn for  $K \leq K^*$ , this implies that the minimum zero-norm with  $\mathbf{D} = \mathbf{D}^*$  is equal to  $N - K^*$ . Consequently  $N - K^*$  is the tightest lower bound achievable with a diagonal relaxation.

The term diagonal relaxation will refer henceforth to the tightest diagonal relaxation, and the above procedure will be referred to as solving the diagonal relaxation. Problem (17) can be recast as a

semidefinite program to which efficient interior-point algorithms as well as other simplifications may be applied. A detailed discussion of the solution of (17) is beyond the scope of the current paper.

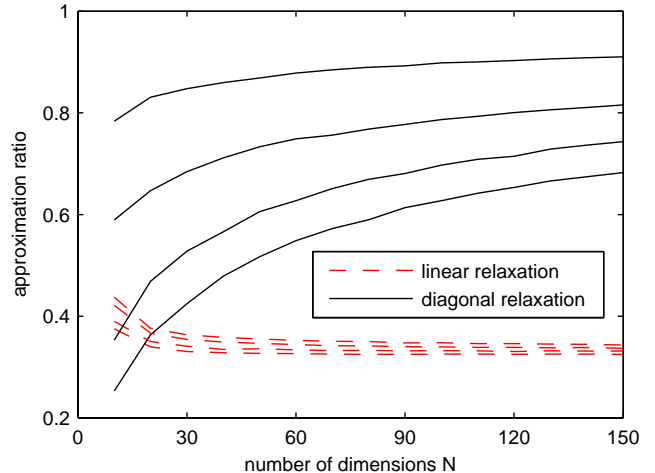
The solution of the diagonal relaxation suggests a heuristic method for generating a feasible solution to the original problem (2). The final matrix  $\mathbf{D}^*$  has the property that the sum of the  $K^*$  smallest  $D_{nn}^* c_n^2$  is no greater than  $\gamma$ . This implies that the index set  $\mathcal{Z}$  corresponding to the  $K^*$  smallest  $D_{nn}^* c_n^2$  is feasible for the relaxed problem. Using (9), we can check whether  $\mathcal{Z}$  (more precisely, its complement  $\mathcal{Y}$ ) is also feasible for problem (2). If it is, an optimal solution to (2) has been found because the zero-norm  $N - K^*$  of the solution is equal to the lower bound provided by the diagonal relaxation. If not,  $\mathcal{Z}$  is reduced in size to correspond to the  $K^* - 1$  smallest  $D_{nn}^* c_n^2$  and the feasibility test is repeated. The size of  $\mathcal{Z}$  is successively decreased in this manner until  $\mathcal{Z}$  becomes feasible, at which point a solution has been obtained with zero-norm equal to  $N - |\mathcal{Z}|$ .

## 6. NUMERICAL EXPERIMENTS

Preliminary numerical experiments were performed to evaluate the quality of the lower bounds resulting from linear and diagonal relaxations. The number of dimensions  $N$  was varied between 10 and 150, and for each value of  $N$ , the condition number  $\kappa(\mathbf{Q})$  was set in turn to  $\sqrt{N}$ ,  $N$ ,  $10N$ , and  $100N$ . One thousand (1000) test cases were created for each pair of  $N$  and  $\kappa(\mathbf{Q})$ . The parameter  $\gamma$  was normalized to 1 throughout. The matrix  $\mathbf{Q}$  was generated by first choosing  $N$  eigenvalues distributed uniformly in the logarithmic domain (i.e.,  $\log \lambda$  is uniformly distributed) and then scaling to match the specified condition number. The eigenvalues were combined with an orthonormal set of eigenvectors oriented randomly and uniformly over the unit sphere. Given the random orientation of eigenvectors, the larger the condition number  $\kappa(\mathbf{Q})$ , the farther  $\mathbf{Q}$  tends to be from being diagonal. Each component  $c_n$  of the ellipsoid center was drawn uniformly from the interval  $[-\sqrt{(\mathbf{Q}^{-1})_{nn}}, \sqrt{(\mathbf{Q}^{-1})_{nn}}]$  to ensure that  $B_n^+$  and  $B_n^-$  in (14) are non-negative.

The linear relaxation (15) of each test problem was solved using the function `fmincon` in MATLAB. We used a custom solver for the diagonal relaxation; a general-purpose solver such as SDPT3 [10] can also be used to solve (17). In addition, a feasible solution was obtained according to the procedure described in Section 5. The ratio of the optimal cost of each relaxation to the cost of the feasible solution is used to assess the quality of the relaxation. This ratio, referred to as the approximation ratio, is a lower bound on the ratio of the optimal cost of the relaxation to the true optimal cost, the latter of which is difficult to compute.

In Fig. 1 we plot the average approximation ratios for linear and diagonal relaxations as functions of  $N$  and  $\kappa(\mathbf{Q})$ . For linear relaxation, the ratio does not vary much with  $N$  or  $\kappa(\mathbf{Q})$  except for a slight decrease at low  $N$ . In contrast, the ratio for diagonal relaxation is markedly higher for lower  $\kappa(\mathbf{Q})$  as expected since  $\mathbf{Q}$  is on average closer to being diagonal. For  $\kappa(\mathbf{Q}) = \sqrt{N}$ , approximation ratios between 0.78 and 0.91 imply that the lower bounds obtained through diagonal relaxations are quite strong. Moreover, the ratio also improves with increasing  $N$ , so that even for  $\kappa(\mathbf{Q}) = 100N$  the diagonal relaxation outperforms the linear relaxation for  $N \geq 20$ . The difference is substantial at large  $N$  and is reflected not only in the average ratios but also in the distributions; as  $N$  increases, histograms of optimal values for diagonal relaxations become widely separated from corresponding histograms for linear relaxations.



**Fig. 1.** Average approximation ratios for linear and diagonal relaxations. Within each set of curves,  $\kappa(\mathbf{Q}) = \sqrt{N}, N, 10N, 100N$  from top to bottom.

## 7. REFERENCES

- [1] J. T. Kim, W. J. Oh, and Y. H. Lee, "Design of nonuniformly spaced linear-phase FIR filters using mixed integer linear programming," *IEEE Trans. Signal Processing*, vol. 44, pp. 123–126, Jan. 1996.
- [2] J. L. H. Webb and D. C. Munson, "Chebyshev optimization of sparse FIR filters using linear programming with an application to beamforming," *IEEE Trans. Signal Processing*, vol. 44, pp. 1912–1922, Aug. 1996.
- [3] D. Mattered, F. Palmieri, and S. Haykin, "Efficient sparse FIR filter design," in *Proc. ICASSP*, May 2002, vol. 2, pp. 1537–1540.
- [4] T. A. Baran and A. V. Oppenheim, "Design and implementation of discrete-time filters for efficient rate-conversion systems," in *Proc. Asilomar Conf. Signals Syst. Comp.*, Nov. 2007.
- [5] J. W. Adams, "FIR digital filters with least-squares stopbands subject to peak-gain constraints," *IEEE Trans. Circuits Syst.*, vol. 39, pp. 376–388, Apr. 1991.
- [6] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. 1, John Wiley & Sons, New York, 2004.
- [7] M. H. Hayes, *Statistical digital signal processing and modeling*, John Wiley & Sons, New York, 1996.
- [8] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, pp. 489–509, Feb. 2006.
- [9] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 1999.
- [10] K. C. Toh, R. H. Tütüncü, and M. J. Todd, *On the implementation and usage of SDPT3 — a MATLAB software package for semidefinite-quadratic-linear programming, version 4.0*, July 2006, available at <http://www.math.nus.edu.sg/~mattohkc/sdpt3.html>.