GENERALIZED SAMPLING WITH STRUCTURAL CONSTRAINTS

Tsvi G. Dvorkind and Yonina C. Eldar

Department of Electrical Engineering
Technion–Israel Institute of Technology, Haifa, Israel
Email: {dvorkind@tx,yonina@ee}.technion.ac.il.

ABSTRACT

We consider a generalized sampling scheme, for finite-dimensional sampling and reconstruction spaces where the correcting transformation must satisfy certain structural constraints. Several criteria are considered, which aim to minimize the squared error norm between the input signal and its approximation within the reconstruction space. By restricting the correcting transformation to a desired form, it is possible to obtain efficient implementations of the correction stage which may be vital in real time applications.

1. INTRODUCTION

Recent approaches to sampling and reconstruction treat the sample sequence as inner products of the input signal x in some Hilbert space $\mathscr{H}$ with a set of sampling vectors, which form a frame for the sampling subspace $\mathscr{S} \subseteq \mathscr{H}$. The problem then is to reconstruct x from these samples, using a set of vectors that form a frame for the reconstruction subspace $\mathscr{W} \subseteq \mathscr{H}$, resulting in the reconstructed signal $\hat{x}$ [1, 2, 3].

To obtain a good approximation of the signal in the reconstruction space from its samples, the samples can be processed, prior to reconstruction, with a correcting transformation. The works in [1, 2], propose a linear transformation H_{con} which achieves a *consistent reconstruction* of x, i.e. a reconstruction $\hat{x}$ with the property that it yields the same sample sequence as x. Note, however, that the fact that x and $\hat{x}$ yield the same samples does not necessarily imply that $\hat{x}$ is close to x in the squared norm sense. In [3], different objectives were considered, which are related directly to the squared norm error between x and $\hat{x}$. One of the main conclusions of [3] was that the correcting transformation H, that minimizes the worst possible squared error other all finite energy inputs, results in the reconstruction $\hat{x} = P_{\mathscr{W}} P_{\mathscr{S}} x$. Here, $P_{\mathscr{A}}$ stands for the orthogonal projection onto the subspace $\mathscr{A}$.

We note however, that both for the consistent reconstruction criterion, and the robust criterion of [3], the resultant correcting transformation is not constrained to a predefined structure. In many practical applications we seek an efficient and possibly casual correcting transformation H. Hence, structured forms for H are often required.

In this paper we treat the problem of finite dimensional sampling and reconstruction spaces. Specifically, we assume that there is a set of m samples, which after proper processing, are used to reconstruct a signal using m vectors which form a frame for the reconstruction space. The structure of the correcting transformation H is modeled by a linear combination of some predefined matrices $F_n \in \mathbb{R}^{m \times m}$, *i.e.*, $H = \sum_n F_n h[n]$, where $h[n]$ are the design parameters and F_n are the matrices which define the desired structure. As an example, assume that a casual correcting transformation is called upon. In this case, we wish to restrict the correcting transformation H to have a lower triangular form. This can be achieved by choosing matrices F_n which are zero within their upper triangular part. As another example, consider the case where H is restricted to be cyclic. The motivation for this structure is that it allows applying the correcting stage in an efficient manner. Specifically, using the FFT algorithm, the correcting stage can be applied in $O(m \log m)$ multiplications instead of m^2, as it happens for some general unstructured H.

In this work we address two criteria, both aimed at minimizing an objective related to the norm of the squared error. First, we consider a robust criterion, in which we seek the structurally constrained H that minimizes the norm of the squared error for the worst possible (finite energy) input x. Noticing that the worst possible input can be viewed as an eigenvector corresponding to the maximal eigenvalue of the error operator, we will also consider minimizing the trace of that operator. The latter objective corresponds to optimizing an average performance merit.

The outline of this paper is as follows. Section 2 describes the notations, mathematical preliminaries and problem formulation. In Section 3 we address the average performance objectives, and develop explicit solutions to the problem. Section 4 presents some results on the robust objectives. Specifically, we show that by optimizing the worst case performance, the problem can be restated as a *semi-definite programming* (SDP). Finally, in Section 5 we present a design example.

2. SAMPLING FORMULATION

We denote vectors in an arbitrary Hilbert space $\mathscr{H}$ by lowercase letters, and the elements of a sequence $c \in \ell_2$ by $c[n]$. The operator $P_{\mathscr{A}}$ represents the orthogonal projection onto a closed subspace $\mathscr{A}$ of $\mathscr{H}$ and $\mathscr{A}^{\perp}$ is the orthogonal complement of $\mathscr{A}$. The *Moore-Penrose pseudo inverse* and the adjoint of a bounded transformation T are written as $T^{\dagger}$ and T^*, respectively. $M \succeq 0$ means that the matrix M is *positive semi-definite* (PSD). The inner product between vectors $x, y \in \mathscr{H}$ is denoted by $\langle x, y \rangle$, and is linear in the second argument. $\|x\|^2 = \langle x, x \rangle$ is the squared norm of x and $\overline{a}$ stands for the conjugate of the number a. $\mathscr{L}(\mathscr{H})$ represents the set of bounded linear operators from $\mathscr{H}$ to $\mathscr{H}$. A set transformation $V : \ell_2 \to \mathscr{H}$ corresponding to frame vectors $\{v_n\} \in \mathscr{H}$ is defined by $Va = \sum_n a[n] v_n$ for all $a \in \ell_2$. From the definition of the adjoint, if $a = V^* y$, then $a[n] = \langle v_n, y \rangle$. Using a set transformation A which corresponds to vectors that span the subspace $\mathscr{A}$ we can write the orthogonal projection operator $P_{\mathscr{A}}$ as

$$P_{\mathscr{A}} = A(A^*A)^{\dagger} A^*. \tag{1}$$

We consider a general sampling problem in a Hilbert space $\mathscr{H}$, in which the goal is to reconstruct a signal $x \in \mathscr{H}$ from a finite set of samples c. Our formulation of the problem allows for a broad class of sampling strategies where the basic constraint we impose on the sampling process is that it is linear. The samples are modeled as the inner products of the signal x with a set of m sampling vectors $\{s_n\}_{n=1}^m$ that form a frame for a space $\mathscr{S}$, so that $c[n] = \langle s_n, x \rangle$. Denoting by S the set transformation corresponding to the vectors $\{s_n\}_{n=1}^m$, the samples can be written as $c = S^* x$. The problem is to reconstruct x from c, where the reconstruction $\hat{x}$ has the form

$$\hat{x} = \sum_n d[n] w_n = W d, \tag{2}$$

for some coefficients $d = Hc$ that are a linear transformation of c. Here, W is the set transformation corresponding to a set of vectors

This work was supported by the European Union's Human Potential Programme, under the contract HPRN-CT-2003-00285 (HASSIP).

$\{w_n\}_{n=1}^{m}$ that span the reconstruction space $\mathscr{W}$. The sampling and reconstruction scheme is illustrated in Fig. 1.

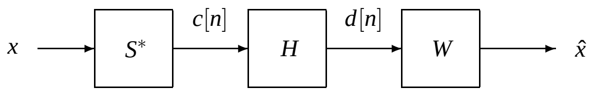

Figure 1: General sampling and reconstruction scheme.

Our problem is to choose the transformation H in Fig. 1 such that $\hat{x}$ is a good approximation of x, in the squared norm sense. The transformation H is restricted to a predefined structure,

$$H = \sum_n F_n h[n], \tag{3}$$

where the matrices $F_n \in \mathbb{R}^{m\times m}$ are specified in advance, and h is the vector of the design parameters.

Aiming to minimize the norm of the squared error, a natural objective is

$$\min_h \|x - WHS^*x\|^2 . \tag{4}$$

Unfortunately, in the general case, perfect reconstruction is not possible[1] and a solution of (4) will depend on x, which is unknown. Instead, we can optimize the correcting transformation H with respect to a class of input signals. Considering the class of finite energy signals, and assuming at first that H is not restricted to a predefined form, we may solve

$$\min_H \max_{\|x\|\le L} \|x - WHS^*x\|^2 . \tag{5}$$

This objective aims to minimize the norm of the worst squared error between x and the reconstruction $\hat{x} = WHS^*x$. Here, L is an arbitrary positive constant, which does not effect the solution H. In [3] it was proved that $H = 0$ solves (5). As such, this useless solution is valid for our structurally constrained forms as well (as evident from (3), by setting $h = 0$). Since the objective (5) is over-conservative, it was suggested in [3] to examine the difference between $\hat{x} = WHS^*x$, and $P_{\mathscr{W}}x$ which is the best approximation to x in $\mathscr{W}$, over all bounded norm signals. This leads to the regret criterion:

$$\min_H \max_{\|x\|\le L} \|P_{\mathscr{W}}x - WHS^*x\|^2 . \tag{6}$$

The minimax regret solution was shown to be given by

$$H_{reg} = (W^*W)^\dagger W^*S(S^*S)^\dagger , \tag{7}$$

and the resulting reconstruction is simply the application of two orthogonal projections $\hat{x} = WH_{reg}S^*x = P_{\mathscr{W}}P_{\mathscr{S}}x$. The objective (6) and its solution leads us to consider the following two structurally constrained counterparts:

- Minimize the regret criterion, using a structurally constrained transformation (as stated in (3)):

$$\min_h \max_{\|x\|\le L} \|P_{\mathscr{W}}x - WHS^*x\|^2 . \tag{8}$$

- Minimize the error between the unconstrained reconstruction (which is obtained by H_{reg}) and the one attainable with a structurally constrained H:

$$\min_h \max_{\|x\|\le L} \|P_{\mathscr{W}}P_{\mathscr{S}}x - WHS^*x\|^2 . \tag{9}$$

[1]Note that we restrict the reconstructed signal $\hat{x} = WHS^*x$ to lie in the reconstruction space $\mathscr{W}$. If $x \notin \mathscr{W}$ to begin with, perfect reconstruction is impossible, regardless of the choice of H.

Problems (8) and (9) are addressed in Section 4. Defining the error operators

$$\mathscr{E}_1 = P_{\mathscr{W}} - WHS^*; \quad \mathscr{E}_2 = P_{\mathscr{W}}P_{\mathscr{S}} - WHS^*, \tag{10}$$

we can view (8) as equivalent to the spectral norm minimization of $\mathscr{E}_1^*\mathscr{E}_1$. Similarly, (9) is equivalent to the spectral norm minimization of $\mathscr{E}_2^*\mathscr{E}_2$. Instead of optimizing the worst case performance, we may consider minimizing the trace of the error operators $\mathscr{E}_1^*\mathscr{E}_1$ and $\mathscr{E}_2^*\mathscr{E}_2$, *i.e.,*

$$\min_h \mathrm{Tr}\{\mathscr{E}_1^*\mathscr{E}_1\}, \tag{11}$$

and

$$\min_h \mathrm{Tr}\{\mathscr{E}_2^*\mathscr{E}_2\}. \tag{12}$$

This corresponds to an average performance measure, where the input x is assumed to be in the direction of all eigenvectors, with equal weights. Problems (11) and (12) are treated in Section 3.

Before stating results, we mention some theorems and definitions, relevant to the notion of the trace.

Theorem 1. *[4, Thr. VI.18] Let $A \in \mathscr{L}(\mathscr{H})$ be a positive operator on a separable Hilbert space $\mathscr{H}$. Then the trace of A is defined by $Tr\{A\} = \sum_n \langle\varphi_n, A\varphi_n\rangle$ for any orthonormal basis $\{\varphi_n\}$ of $\mathscr{H}$.*

Definition 1. *[4, pp. 207] An operator $A \in \mathscr{L}(\mathscr{H})$ is called trace class iff $Tr\{(A^*A)^{0.5}\} < \infty$. The existence of the square root operator is guaranteed for all positive linear bounded operators, by [4, Thr. VI.9].*

Finally, we quote a theorem which guarantees that the trace is unchanged under cyclic shifts, for trace class operators.

Theorem 2. *[4, Thr. VI.25] $Tr\{AB\} = Tr\{BA\}$ if A is a trace class and $B \in \mathscr{L}(\mathscr{H})$.*

3. OPTIMIZING AVERAGE PERFORMANCE

We start by examining the trace of the operators $\mathscr{E}_1^*\mathscr{E}_1$, $\mathscr{E}_2^*\mathscr{E}_2$. For that, we state the following result:

Theorem 3. *The solutions of (11) and (12) are the same.*

Proof. Note that $\mathscr{E}_1 = \mathscr{E}_2 + P_{\mathscr{W}}P_{\mathscr{S}^\perp}$. Due to finite dimensionality of $\mathscr{W}$, $\mathscr{E}_1$ is a trace class operators, which justifies the cyclic shift within the trace. Hence, $\mathrm{Tr}\{\mathscr{E}_1^*\mathscr{E}_1\} = \mathrm{Tr}\{\mathscr{E}_1\mathscr{E}_1^*\}$. Finally, since the range of $\mathscr{E}_2^*$ is within $\mathscr{S}$ we have that $\mathscr{E}_1\mathscr{E}_1^* = \mathscr{E}_2\mathscr{E}_2^* + P_{\mathscr{W}}P_{\mathscr{S}^\perp}P_{\mathscr{W}}$. Using the linearity of the trace we conclude $\mathrm{Tr}\{\mathscr{E}_1\mathscr{E}_1^*\} = \mathrm{Tr}\{\mathscr{E}_2\mathscr{E}_2^*\} + \mathrm{Tr}\{P_{\mathscr{W}}P_{\mathscr{S}^\perp}P_{\mathscr{W}}\}$, from which the theorem follows. □

Theorem 3 shows that while dealing with the trace criterion, it is sufficient to obtain solutions for (12). We note that by using (1) and (7), $\mathscr{E}_2 = W(H_{reg} - H)S^*$. The last expression also suggests a way to solve (12):

$$\min_h \mathrm{Tr}\{\mathscr{E}_2^*\mathscr{E}_2\} = \min_h \mathrm{Tr}\{(H_{reg} - H)^*W^*W(H_{reg} - H)S^*S\}. \tag{13}$$

As the objective in (13) is convex in h, a closed form solution can be obtained by differentiating and equating to zero. This results in the following set of equations:

$$y = Mh, \tag{14}$$

where

$$\begin{aligned} y[i] &= \mathrm{Tr}\{F_i^*W^*WH_{reg}S^*S\} \\ M_{i,j} &= \mathrm{Tr}\{F_i^*W^*WF_jS^*S\} \end{aligned} \tag{15}$$

are the i'th element of y and the i, j'th element of the $m \times m$ matrix M. Note that (13) is convex in h, but might not be strictly convex[2]. Thus, a solution of (14) exists, but is not necessarily unique. Choosing the minimal norm solution, leads to $h_{opt} = M^{\dagger} y$. We note however, that regardless of the choice of the coefficients h which solve (14), the resulting transformation WHS^* is unique. A proof of this can be found in Appendix A.

3.1 Design Examples

Example 1: Diagonal matrix. Consider a simple special case where we restrict H to be a diagonal matrix. This might be vital in real time applications, if the digital correcting stage must be maintained very simple. Assuming that (W^*W) and (S^*S) are of full rank, it can be shown that (14) takes the form:

$$\begin{pmatrix} (W^*S)_{1,1} \\ (W^*S)_{2,2} \\ \vdots \\ (W^*S)_{m,m} \end{pmatrix} = (W^*W).*(S.^*S)h,$$

where $(W^*W).*(S.^*S)$ stands for the element-wise multiplication of these two matrices. Accordingly, when in addition we have orthogonal systems, the optimal coefficients are

$$h_{opt}[n] = \frac{(W^*S)_{n,n}}{(W^*W)_{n,n}(S^*S)_{n,n}};\ n = 1, \ldots, m.$$

Note that this solution corresponds to Frobenius norm minimization of $H_{reg} - H$. In fact, regardless of the choice of the structure matrices $\{F_n\}$, if in addition $W^*W = S^*S = I$, problem (13) is always the Frobenius norm minimization of $H_{reg} - H$.

Example 2. Cyclic matrix. Another interesting case is when H is constrained to be a cyclic matrix. In such circumstances, the transformed samples $d = Hc$ can be calculated efficiently, requiring only $O(m \log m)$ multiplications. Though the system (14) defines a way to obtain the solution, it provides no insight. Instead, we will further assume that W^*W and S^*S are full rank cyclic matrices, and present an efficient algorithm for obtaining the coefficients h. Though we assume cyclic structure for W^*W and S^*S, we assume nothing for the structure of W^*S. Before showing an algorithm for computing a cyclic H, which minimizes the trace objective, we note that in such circumstances, H_{reg} itself can be efficiently computed as well. This is since the inversion of the cyclic matrices W^*W and S^*S can be accomplished using the FFT algorithm.

Let $\{F_i\}$, $i = 1, \ldots, m$ be the appropriate structure matrices. Specifically, we set $F_1 = I$, and set F_n by shifting the columns of F_{n-1} upwards in a cyclic manner. For example, when $m = 3$, we have

$$F_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad F_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad F_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

The following algorithm finds the proper coefficients h.

1. Set $q[n]$, $n = -(m-1), \ldots, (m-1)$ to be the mean value along the n'th diagonal of W^*S. Here, $n = 0$ corresponds to the main diagonal, $n = 1$ to one diagonal above it, while $n = -1$ to the diagonal below it, etc.
2. Multiply the sequence q by a triangular window $g[n]$
$$v[n] = q[n]g[n],$$
where
$$g[n] = \frac{m - |n|}{m}; \quad n = -(m-1), \ldots, m-1.$$
3. Calculate the *discrete-time Fourier transform* (DTFT) of $v[n]$ at angles $\frac{2\pi k}{m}, k = 0, \ldots, m-1$ using
$$V[k] = \text{DFT}\left(\sum_{l=-1}^{0} v[n+lm]\right)[k],$$
where $\text{DFT}(f[n])[k]$ stands for the *discrete Fourier transform* of $f[n]$, *i.e.*, $\sum_{n=0}^{m-1} f[n] e^{-j\frac{2\pi nk}{m}}$.
4. Let r_{ww} be the first row of the matrix W^*W. Similarly, define r_{ss} to be the first row of the matrix S^*S. Calculate the DTFT of r_{ww} and r_{ss} at the same angles using
$$R_{ww}[k] = \text{DFT}(r_{ww}[n])[k], \quad R_{ss}[k] = \text{DFT}(r_{ss}[n])[k].$$
5. The optimal coefficients are given by
$$h[n] = \text{IDFT}\left(\frac{V[k]}{R_{ww}[k] R_{ss}[k]}\right)[n],$$
where $\text{IDFT}(A[k])[n]$ stands for the inverse DFT, *i.e.*, $\frac{1}{m}\sum_{k=0}^{m-1} A[k] e^{j\frac{2\pi nk}{m}}$.

Note that as a special case, when $W^*W = S^*S = I$ (and thus $R_{ww}[k] = R_{ss}[k] = 1\ \forall k$) we have an algorithm for obtaining the best (in the Frobenius norm sense) cyclic approximation of some general matrix W^*S. This approximation can be seen as a two stage procedure; first project the matrix into the subspace of Toeplitz matrices by calculating the mean value along each diagonal. Then, further project into the smaller subspaces of cyclic matrices (this is obtained by the multiplication with the triangular window and the summation, as stated within steps 2 and 3 of the algorithm).

4. OPTIMIZING WORST-CASE PERFORMANCE

In this section we consider objectives (8) and (9). As mentioned previously, these problems are equivalent to minimizing the maximal eigenvalue of the operators $\mathscr{E}_1^*\mathscr{E}_1$ and $\mathscr{E}_2^*\mathscr{E}_2$. First, we wish to relate the two problems. For that we state the following lemma.

Lemma 1. *Let $A, B \in \mathscr{L}(\mathscr{H})$. The non zero point spectrum of the operator AB is equal to that of the operator BA.*

Proof. Assume $BAx = \lambda x$ for some $0 \neq x \in \mathscr{H}$. Define $y = Ax$. Assume $y \neq 0$ (else, we would have $\lambda = 0$). Since $ABy = A(\lambda x) = \lambda Ax = \lambda y$ we established that λ is also an eigenvalue of AB. □

We can use Lemma 1 to claim that $\lambda_{max}(\mathscr{E}_1^*\mathscr{E}_1) = \lambda_{max}(\mathscr{E}_1\mathscr{E}_1^*)$ and similarly for $\mathscr{E}_2$. Here we implicitly assume that the maximal eigenvalue is not zero, *i.e.*, that $\mathscr{E}_1$ and $\mathscr{E}_2$ are not the zero operators, which is true in the general case.

Since, $\mathscr{E}_1\mathscr{E}_1^* = \mathscr{E}_2\mathscr{E}_2^* + P_{\mathscr{W}} P_{\mathscr{S}^\perp} P_{\mathscr{W}}$ we conclude that

$$\lambda_{max}(\mathscr{E}_1^*\mathscr{E}_1) \leq \lambda_{max}(\mathscr{E}_2^*\mathscr{E}_2) + \lambda_{max}(P_{\mathscr{W}} P_{\mathscr{S}^\perp} P_{\mathscr{W}}).$$

Furthermore, since $\lambda_{max}(P_{\mathscr{W}} P_{\mathscr{S}^\perp} P_{\mathscr{W}}) = \lambda_{max}(P_{\mathscr{W}} P_{\mathscr{S}^\perp}) \leq 1$, we have

$$\lambda_{max}(\mathscr{E}_1^*\mathscr{E}_1) \leq \lambda_{max}(\mathscr{E}_2^*\mathscr{E}_2) + 1.$$

As a result, the value of the objective (9) can be used to upper bound (8). Yet, the question remains, how to solve the two problems at hand. The next theorem shows that an efficient numerical solution can be obtained by recasting each objective as an SDP problem.

Theorem 4. 1 *The solution of $\arg\min_h \lambda_{max}(\mathscr{E}_2^*\mathscr{E}_2)$ can be obtained by solving the following convex optimization problem:*

$$\arg\min_{h,t} \quad t \tag{16}$$
$$s.t. \quad \begin{pmatrix} I & B \\ B^* & tI \end{pmatrix} \succeq 0$$

where

$$B = (W^*W)^{\frac{1}{2}}(H_{reg} - H)(S^*S)^{\frac{1}{2}}. \tag{17}$$

[2] Strict convexity can be assured if W^*W and S^*S are of rank m. The latter holds when the sampling and reconstruction vectors form Riesz bases for $\mathscr{S}$ and $\mathscr{W}$, respectively.

2 *The solution of* $\arg\min_h \lambda_{max}(\mathscr{E}_1^*\mathscr{E}_1)$ *can be obtained by solving the following convex optimization problem:*

$$\begin{aligned} \arg\min_{h,t} \quad & t & (18)\\ s.t. \quad & \begin{pmatrix} I & B^* \\ B & tI - C \end{pmatrix} \succeq 0 \end{aligned}$$

where B is defined by (17) and

$$C = (W^*W)^{\frac{1}{2}} \left((W^*W)^{\dagger} - H_{reg}S^*W(W^*W)^{\dagger} \right) (W^*W)^{\frac{1}{2}}. \quad (19)$$

Proof. **1** Since $\mathscr{E}_2 = W(H_{reg} - H)S^*$, we have that $\lambda_{max}(\mathscr{E}_2^*\mathscr{E}_2) = \lambda_{max}(S(H_{reg} - H)^*W^*W(H_{reg} - H)S^*)$. Using Lemma 1 and taking square roots, we also obtain $\lambda_{max}(\mathscr{E}_2^*\mathscr{E}_2) = \lambda_{max}(B^*B)$, where B is given by (17). By introducing a slack variable t, we can write an equivalent problem of minimizing the maximal eigenvalue, *i.e.*,

$$\begin{aligned} \arg\min_{h,t} \quad & t & (20)\\ s.t. \quad & tI - B^*B \succeq 0. & (21) \end{aligned}$$

Using the lemma of Schur [5, Lemma 4.2.1] we obtain (16).

2 Since $\lambda_{max}(\mathscr{E}_1^*\mathscr{E}_1) = \lambda_{max}(\mathscr{E}_1\mathscr{E}_1^*)$ and using $\mathscr{E}_1\mathscr{E}_1^* = \mathscr{E}_2\mathscr{E}_2^* + P_{\mathscr{W}}P_{\mathscr{S}^\perp}P_{\mathscr{W}}$ we obtain that

$$\lambda_{max}(\mathscr{E}_1^*\mathscr{E}_1) = \lambda_{max}(\mathscr{E}_2\mathscr{E}_2^* + P_{\mathscr{W}} - P_{\mathscr{W}}P_{\mathscr{S}}P_{\mathscr{W}}).$$

After rewriting the orthogonal projections using the set transformations S and W, applying cyclic shifts and taking square roots, it is easy to show that

$$\lambda_{max}(\mathscr{E}_1^*\mathscr{E}_1) = \lambda_{max}(BB^* + C)$$

where B and C are given by (17) and (19), respectively. We note that the matrix C has the same eigenvalues as the spectrum points of the operator $P_{\mathscr{W}}P_{\mathscr{S}^\perp}P_{\mathscr{W}}$. Applying the lemma of Schur to the equivalent problem

$$\begin{aligned} \arg\min_{h,t} \quad & t & (22)\\ s.t. \quad & tI - (BB^* + C) \succeq 0, \end{aligned}$$

results in (18).

□

Surely, if H_{reg} is obtainable by a linear combination of $\{F_i\}$, then the values of (8) and (9) are $\lambda_{max}(P_{\mathscr{W}}P_{\mathscr{S}^\perp})$ and zero, respectively. We also note that the value of (11) can be used as an upper bound for $\lambda_{max}\left(\mathscr{E}_1^*\mathscr{E}_1\right)$. Similarly, the value of (12) upper bounds $\lambda_{max}\left(\mathscr{E}_2^*\mathscr{E}_2\right)$. As a result, an easy way to obtain a bound for the worst case performance (without actually solving the SDP problems), is by evaluating (13) first.

In the next section we present an example which illustrates the use of a correcting matrix H, constrained to a cyclic form.

5. EXAMPLE

Suppose we sample a continuous-time speech signal $x(t)$ using a non-ideal sampler, so that the samples $c[n]$ are equal to the average of the signal over intervals of length Δ:

$$c[n] = \frac{1}{\Delta}\int_{nT-\Delta}^{nT} x(t)dt. \quad (23)$$

The samples $c[n]$ can be obtained by filtering the signal $x(t)$ with a filter whose impulse response $s(t)$ is given by

$$s(t) = \begin{cases} \frac{1}{\Delta}, & 0 \le t \le \Delta; \\ 0, & \text{otherwise}, \end{cases} \quad (24)$$

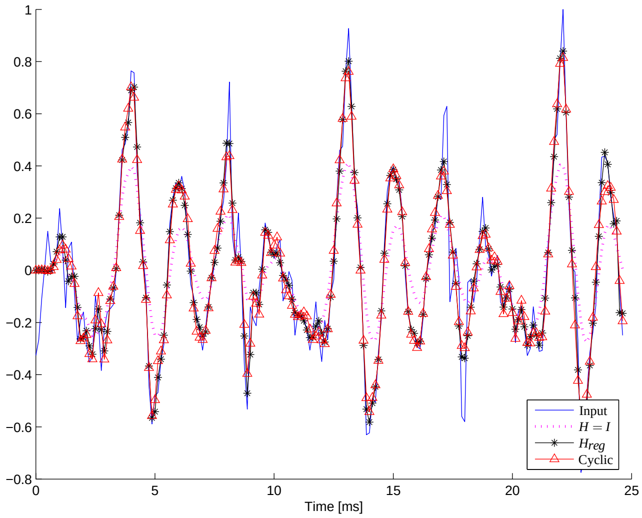

Figure 2: The original speech fragment $x[n]$ and the reconstructed signal using three different methods: not applying correction at all, H_{reg} as a correcting transformation, and a correction with a cyclic matrix.

and then sampling the output at times $t = nT$. The filter $s(t)$ can be viewed as a (non-ideal) low-pass filter (LPF). In the simulations below, we use $T = 4000^{-1}$s and $\Delta = 1.125$ms. The reconstructed output $\hat{x}(t)$ is obtained from the transformed sequence $d = Hc$ using an interpolation kernel $w(t)$, *i.e.*, $\hat{x}(t) = \sum_n d[n]w(t - nT)$. Specifically, we consider $w(t)$ to be a non-ideal LPF with support on $t \in [0, 1.75]$ms which approximates an ideal LPF with cutoff frequency of 2kHz.

For the purpose of simulation we approximate the continuous-time signal $x(t)$ with a discrete sequence $x[n]$ on a fine grid. The signal was chosen as a speech fragment, taken from the Timit database [6], at a sample rate of 8kHz. The continuous time integration kernel $s(t)$ is approximated by the discrete filter

$$s[n] = \begin{cases} \frac{1}{N}, & 0 \le n \le N-1; \\ 0, & \text{otherwise}, \end{cases} \quad (25)$$

with $N = 10$ samples. The ideal sampling is implemented by down-sampling the filter output with a decimation factor of 2. The (non-ideal) LPF followed by decimation can be described by proper construction of the sampling matrix S^*.

To implement the reconstruction we use a linear-phase FIR filter of order 14 (with cutoff frequency 2kHz) as the interpolation kernel. Here as well, the discrete-time interpolation kernel simulates the continuous-time interpolation kernel, by constructing it over the 8kHz fine grid and up-sampling the input sequence d by a factor of 2, prior to filtering. The up-sampling followed by the filtering operation can be described by properly constructing the matrix W. Figure 2 shows an example of an input sequence $x[n]$ and 3 different reconstructed signals, corresponding to $H = I$ (that is, no correcting transformation is applied at all), $H = H_{reg}$ and structurally constrained H to a cyclic form. Specifically, H was optimized as a solution of (13).

The motivation behind the cyclic structure, is that the multiplication Hc can be obtained in an efficient manner (requiring only $m\log_2 m + m$ complex multiplications, instead of m^2).

As can be seen from the figure, when no correcting transformation is applied, the results are poor. On the other hand, the reconstructions obtained by H_{reg} and the cyclic H are much closer to the original speech fragment. Furthermore, and at least for this example, it seems that there is no meaningful increase in the approximation error despite the structural constraint which were imposed on

H. On the other hand, we have gained an efficient implementation of the correcting stage.

A. APPENDIX

Proof. Let h_1 and h_2 be two vectors which solve (14). Let $H_i = \sum_n F_n h_i[n]$, $i = 1,2$ be the corresponding matrices. We will show that $WH_2S^* - WH_1S^*$ must be the zero transformation. It easy to show that A is the zero transformation if and only if $\mathrm{Tr}\{A^*A\} = 0$. Accordingly, we will prove

$$J = \mathrm{Tr}\{(H_2 - H_1)^* W^* W (H_2 - H_1) S^* S\} = 0,$$

where we used a cyclic shift within the trace. Since

$$J = \mathrm{Tr}\left\{(H_2 - H_1)^* W^* W (H_2 - H_{reg} + H_{reg} - H_1) S^* S\right\},$$

we can use the linearity of the trace and the orthogonality principle, as imposed by (14), to show that

$$J = \mathrm{Tr}\left\{(H_2 - H_1)^* W^* W (H_i - H_{reg}) S^* S\right\} = 0; \; i = 1,2,$$

which completes the proof. □

REFERENCES

[1] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Trans. Signal Processing*, vol. 42, no. 11, pp. 2915–2925, Nov. 1994.

[2] Y. C. Eldar, "Sampling and reconstruction in arbitrary spaces and oblique dual frame vectors," *J. Fourier Analys. Appl.*, vol. 1, no. 9, pp. 77–96, Jan. 2003.

[3] Y. C. Eldar and T. Dvorkind, "A minimum squared-error framework for sampling and reconstruction in arbitrary spaces," submitted to *IEEE Trans. Signal Processing*, July 2004.

[4] M. Reed and B. Simon, *Methods of modern mathematical physics. I: Functional Analysis. Revised and Enlarged Edition*, Academic Press, 1980.

[5] A. Ben-Tal and A. Nemirovsky, *Lectures on Modern Convex Optimization. Analysis, Algorithms, and Engineering Applications*, MPS-SIAM, 2001.

[6] National Institute of Standards and Technology, "The DARPA TIMIT acoustic-phonetic continuous speech corpus," CD-ROM NIST Speech Disc 1-1.1, Oct. 1991.