

Mean-Squared Error Sampling and Reconstruction in the Presence of Noise

Yonina C. Eldar, *Member, IEEE*

Abstract—One of the main goals of sampling theory is to represent a continuous-time function by a discrete set of samples. Here, we treat the class of sampling problems in which the underlying function can be specified by a finite set of samples. Our problem is to reconstruct the signal from nonideal, noisy samples, which are modeled as the inner products of the signal with a set of sampling vectors, contaminated by noise. To mitigate the effect of the noise and the mismatch between the sampling and reconstruction vectors, the samples are linearly processed prior to reconstruction. Considering a statistical reconstruction framework, we characterize the strategies that are mean-squared error (MSE) admissible, meaning that they are not dominated in terms of MSE by any other linear reconstruction. We also present explicit designs of admissible reconstructions that dominate a given inadmissible method. Adapting several classical estimation approaches to our particular sampling problem, we suggest concrete admissible reconstruction methods and compare their performance. The results are then specialized to the case in which the samples are processed by a digital correction filter.

Index Terms—Generalized sampling, interpolation, minimax reconstruction, sampling.

I. INTRODUCTION

SIGNAL expansions, in which a signal is represented by a set of coefficients, find many applications in signal processing and communications. The expansion coefficients can be regarded as generalized samples of the underlying signal, so that the signal expansion problem can be formulated within a sampling and reconstruction framework.

In this paper, we focus on signal expansions with a finite number of coefficients. A natural setting in which expansions of this type arise is in the context of sampling a continuous-time signal that lies in a finite-dimensional space. For example, it is well known that a bandlimited periodic signal can be reconstructed from an arbitrary finite set of its samples, as long as the number of samples exceeds the corresponding Nyquist rate [1], [2]. In practice, however, the signal samples are often contaminated by noise. The problem then is to “best” approximate the continuous-time signal from the given noisy samples in some sense.

Manuscript received December 31, 2004; revised January 2, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Sheila S. Hemami. This work was supported by the European Union’s Human Potential Programme, under the contract HPRN-CT-2003-00285 (HASSIP) and by the Israel Science Foundation under grant no. 536/04.

The author is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: yonina@ee.technion.ac.il).

Digital Object Identifier 10.1109/TSP.2006.881266

Here we treat a more general formulation of this problem which allows for nonideal samples that can be described as inner products of the signal with a set of sampling vectors $\{s_i, 1 \leq i \leq m\}$. Furthermore, our development is carried out in an arbitrary Hilbert space \mathcal{H} , which includes the continuous-time model as a special case. The underlying signal x is constrained to a finite-dimensional subspace \mathcal{W} of \mathcal{H} . In our framework, the reconstructed signal \hat{x} is obtained by linearly combining a set of reconstruction vectors that span the signal space \mathcal{W} . The combination coefficients are the result of processing the noisy samples with a linear transformation designed to mitigate the effect of the noise and to compensate for the possible mismatch between the sampling and reconstruction vectors.

The sampling setting we consider here, in which we allow for arbitrary sampling and reconstruction vectors, is an extension of the noise-free framework developed previously in the literature [3]–[8], with the restriction that the underlying signal x lies in the reconstruction space \mathcal{W} . In this case, it was shown in [4]–[6] that perfect reconstruction from noise-free samples is possible, regardless of the sampling method, as long as the sampling space and the reconstruction space satisfy a direct-sum condition. Clearly, in the noisy setting, perfect reconstruction is no longer attainable. Our goal then is to linearly process the noisy samples prior to reconstruction such that the reconstructed signal \hat{x} is close to the original signal x , in some statistical sense.

Previous treatment of reconstruction in the presence of noise includes analysis of the noise effects in existing reconstruction systems [9]–[12], and concrete reconstruction methods from noisy samples [13]–[16]. However, the suggested algorithms tend to focus on the bandlimited setting and are typically not specified to be optimal from the point of view of statistical estimation theory. An exception is a recent work [17] that suggests several concrete reconstruction strategies from noisy filtered samples in shift-invariant spaces. Here, we study the reconstruction problem in a general Hilbert-space setting within a statistical framework, and use statistical-based criteria to provide a complete characterization of all possible linear algorithms.

A popular measure of reconstruction performance is the mean-squared error (MSE), which is the average squared-norm of the estimation error $\hat{x} - x$. In our setting, the original signal x is assumed to be fixed; therefore, the averaging is only over the noise vector, resulting typically in a signal-dependent MSE, which cannot be minimized directly. The MSE performance of different reconstruction methods will also generally depend on x , rendering comparison between different methods a difficult, and often impossible task. A similar issue arises in the context of general estimation theory, when studying the performance of linear estimators in a linear model [18]–[20]. Therefore,

in order to develop a statistical framework for sampling and reconstruction, we suggest formulating our problem in an estimation setup and then adapting recent concepts and results developed in that context to our particular sampling setting.

Two key ideas that are of paramount importance when comparing the performance of different estimation strategies are those of domination and admissibility [19], [20]. A reconstruction \hat{x} *dominates* some other approach \hat{x}_0 in terms of MSE if its MSE is never larger than that of \hat{x}_0 for all values of x , and is strictly smaller for some x ; an *admissible* reconstruction is one that is never dominated by another approach. Thus, although we cannot directly evaluate the MSE performance of different reconstruction strategies, we can characterize those methods that are admissible. Surprisingly, some of the previously proposed approaches are inadmissible, implying that they can be uniformly improved upon in terms of MSE. We therefore suggest concrete designs of admissible reconstructions that dominate a given inadmissible method.

To develop our theory, we adapt recent results on admissible MSE linear estimation to our sampling context [20]. Using these ideas, we derive easily verifiable necessary and sufficient conditions on a reconstruction to be admissible. Our treatment considers admissibility on the entire reconstruction space \mathcal{W} , as well as on the set of norm-bounded signals in \mathcal{W} . Although our results are rooted in [20], they are not a direct extension since in our setting the reconstruction is restricted to a subspace \mathcal{W} and the sampling process projects the signal onto the sampling space. By explicitly taking the structure of our problem into account, we present several examples of admissible and dominating reconstructions. In particular, we show that the consistent approach, previously proposed for the noise-free sampling problem [3]–[6], is admissible on the entire space, but not on a spherical constraint set. For norm-bounded signals, we propose a minimax MSE approach that results in an admissible reconstruction that strictly dominates the consistent method. This strategy is based on some of our prior work on minimax estimation for solving robust estimation problems [21], [22], adapted to our specific setting. We also suggest a regularized least-squares, or Tikhonov [23], approach, which is admissible as long as the regularization parameter is chosen appropriately. However, as we demonstrate through an example, the Tikhonov reconstruction is not guaranteed to dominate the consistent method.

A class of reconstruction methods that was proposed in [14], [16] in the context of bandlimited sampling is based on convolving the noisy samples with a finite-impulse-response (FIR) filter prior to reconstruction. Under certain conditions on the sampling and reconstruction spaces, we show that FIR reconstruction is admissible if and only if the filter impulse response is symmetric, and its Fourier transform satisfies $0 \leq G(\omega) \leq 1$. Using these results, we show that one of the methods proposed in [14] is inadmissible. We then construct an alternative admissible reconstruction that strictly dominates it for all $x \in \mathcal{W}$.

Before proceeding to the detailed development, we stress that although the ideas presented in this paper are based on [20], there are several important differences. One of the main contributions of the current work is in setting up the reconstruction problem in an estimation context and systematically classifying

reconstruction strategies, which has not been the conventional approach in the sampling literature. In fact, by following this route, we prove that some of the standard strategies to reconstruction from noisy samples are in fact inadmissible and can be uniformly improved upon using our methods. From a mathematical perspective, the sampling framework is different than the standard linear estimation setting since the sampling setup restricts the problem to two subspaces: the sampling space and the reconstruction space. This added twist renders the mathematics associated with both problems somewhat different, as becomes evident from the mathematical derivations in this paper. Finally, there are aspects of the problem that have not been investigated in [20] but are interesting in the context of sampling, such as FIR filtering. Here again we demonstrate that standard approaches may be inadmissible and can therefore be improved using our framework.

The paper is organized as follows. In Section II, we introduce our general sampling framework, and show that both problems of admissibility and constructing dominating methods can be treated by solving a certain convex optimization problem. Dominating reconstructions and necessary and sufficient admissibility conditions on a bounded norm constraint set and the entire space are developed in Sections III and V, respectively. In Section IV, we present several reconstruction methods on a bounded norm set by adapting known solutions to our particular sampling context: least-squares, Tikhonov, and minimax MSE reconstruction. Finally, in Section VI, we study FIR reconstruction and compare our results with the method of [14].

II. SAMPLING FRAMEWORK

We denote vectors in an arbitrary Hilbert space \mathcal{H} by lowercase letters, and vectors in \mathbb{C}^m by boldface lowercase letters. The i th element of a vector \mathbf{a} is written as a_i . Matrices are represented by uppercase boldface letters and arbitrary linear transformations on \mathcal{H} by uppercase letters. The orthogonal projection operator onto a space \mathcal{S} is denoted by $P_{\mathcal{S}}$, and $\text{diag}(\delta_1, \dots, \delta_m)$ represents an $m \times m$ diagonal matrix with diagonal elements δ_i . Given a transformation T , $\mathcal{N}(T)$, $\mathcal{R}(T)$ are its null space and range space, respectively, and T^\dagger , T^* are the Moore–Penrose pseudoinverse [24] and the adjoint, respectively. The inner product between vectors $x, y \in \mathcal{H}$ is denoted by $\langle x, y \rangle$, and is linear in the second argument, and $\|x\|$ is the norm of x . For an operator A , $A \succ 0$ ($A \succeq 0$) means that A is Hermitian and positive (nonnegative) definite, and $A \succeq B$ means that $A - B \succeq 0$.

The *set transformation* $S: \mathbb{C}^m \rightarrow \mathcal{H}$ corresponding to a set of vectors $\{s_i, 1 \leq i \leq m\}$ is defined by $S\mathbf{a} = \sum_{i=1}^m a_i s_i$ for any $\mathbf{a} \in \mathbb{C}^m$. From the definition of the adjoint $S^*: \mathcal{H} \rightarrow \mathbb{C}^m$ it follows that if $\mathbf{a} = S^*x$, then $a_i = \langle s_i, x \rangle$.

A. Noise-Free Nonideal Sampling

We treat the problem of recovering a signal x from a finite-set of its noisy samples. The signal is assumed to lie in an m -dimensional subspace \mathcal{W} of a Hilbert space \mathcal{H} : examples include finite-length discrete-time signals $x[i]$, $0 \leq i \leq m$, or continuous-time bandlimited periodic signals $x(t)$. To fix ideas, we first treat the reconstruction problem in the noise-free setting.

Given a basis $\{w_i, 1 \leq i \leq m\}$ for \mathcal{W} , any signal $x \in \mathcal{W}$ can be written uniquely as a linear combination of these vectors:

$$x = \sum_{i=1}^m a_i w_i = W \mathbf{a}, \quad (1)$$

where W is the set transformation corresponding to the vectors w_i , and

$$\mathbf{a} = (W^*W)^{-1}W^*x. \quad (2)$$

Denoting by $\{h_i, 1 \leq i \leq m\}$ the vectors corresponding to $H = W(W^*W)^{-1}$, each coefficient a_i can be written as the inner product $a_i = \langle h_i, x \rangle$. The expansion (1) can then be interpreted within the sampling framework as perfect reconstruction of any signal $x \in \mathcal{W}$ from its generalized samples a_i using the reconstruction vectors w_i .

The samples (2) are perfectly matched to the reconstruction vectors w_i . More generally, we can consider samples $c_i = \langle s_i, x \rangle$ that are given as the inner products of x with an arbitrary set of linearly-independent sampling vectors $\{s_i, 1 \leq i \leq m\}$ which span a subspace $\mathcal{S} \subseteq \mathcal{H}$. Although the samples are no longer matched to the reconstruction vectors, it was shown in [4] and [5] that perfect reconstruction of any $x \in \mathcal{W}$ from these samples is still possible as long as

$$\mathcal{W} \cap \mathcal{S}^\perp = \{0\}, \quad (3)$$

where \mathcal{S}^\perp is the orthogonal complement of \mathcal{S} in \mathcal{H} . Since \mathcal{W} and \mathcal{S} have the same finite dimension, (3) is equivalent to the direct-sum condition $\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^\perp$. If (3) is satisfied, then perfect reconstruction can be obtained by first processing the samples c_i with the linear transformation

$$\mathbf{G}_{\text{CON}} = (S^*W)^{-1}. \quad (4)$$

The corresponding reconstruction

$$\hat{x}_{\text{CON}} = W\mathbf{G}_{\text{CON}}\mathbf{c} = W(S^*W)^{-1}S^*x, \quad (5)$$

is referred to as a *consistent reconstruction* [3], since it has the property that it yields the same samples as x for any¹ $x \in \mathcal{H}$. The existence of the inverse in (4) follows from the following proposition.

Proposition 1: Let $\{w_i, 1 \leq i \leq m\}$ be a basis for a space \mathcal{W} with set transformation W and let $\{s_i, 1 \leq i \leq m\}$ be a basis for a space \mathcal{S} satisfying $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$ with set transformation S . Then, we have the following:

- 1) W^*S is invertible;
- 2) $W(S^*W)^{-1} = P_{\mathcal{W}}S(S^*P_{\mathcal{W}}S)^{-1}$.

Proof: The first part of the proposition is proved in [4]. To prove the second part, we first show that

$$\mathcal{R}(P_{\mathcal{W}}S) = \mathcal{W}. \quad (6)$$

This follows from the fact that $\mathcal{R}(P_{\mathcal{W}}S) = \mathcal{N}^\perp(S^*P_{\mathcal{W}})$ (because for any bounded transformation, $\mathcal{R}(A) = \mathcal{N}^\perp(A^*)$) and $\mathcal{N}(S^*P_{\mathcal{W}}) = \mathcal{W}^\perp$. Indeed, since $P_{\mathcal{W}}x \in \mathcal{W}$ for any x , and

¹Alternative methods have been proposed that may yield a better reconstruction when $x \in \mathcal{H}$ [7]. Here, however, we restrict our attention to $x \in \mathcal{W}$.

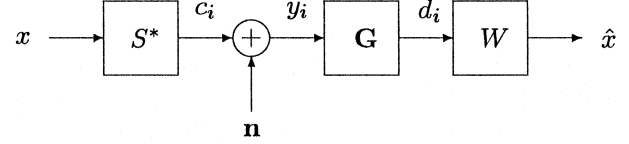


Fig. 1. General sampling and reconstruction scheme.

$\mathcal{N}(S^*) = \mathcal{S}^\perp$, under (3) we have that $S^*P_{\mathcal{W}}x = 0$ only if $P_{\mathcal{W}}x = 0$. Using (6), $P_{\mathcal{W}}S$ can be expressed as $P_{\mathcal{W}}S = W\mathbf{X}$ for some invertible matrix \mathbf{X} . Thus,

$$P_{\mathcal{W}}S(S^*P_{\mathcal{W}}S)^{-1} = W\mathbf{X}(S^*W\mathbf{X})^{-1} = W(S^*W)^{-1}, \quad (7)$$

completing the proof. \blacksquare

We assume throughout the paper that (3) holds.

B. Noisy Sampling

Extending the discussion to the noisy setting, we now treat the case in which the nonideal samples $\mathbf{c} = S^*x$ are corrupted by a zero-mean noise vector \mathbf{n} with positive-definite covariance matrix \mathbf{C} . The noisy samples are then given by

$$\mathbf{y} = S^*x + \mathbf{n}. \quad (8)$$

As in the noise-free setting, we consider reconstructions of the form

$$\hat{x} = \sum_{i=1}^m d_i w_i = W\mathbf{d} = W\mathbf{G}\mathbf{y}, \quad (9)$$

where $\mathbf{d} = \mathbf{G}\mathbf{y}$ is a linear transformation of the noisy samples \mathbf{y} . The sampling and reconstruction scheme is illustrated in Fig. 1. A special case occurs when \mathbf{G} represents an FIR filter, so that reconstruction is obtained by convolving the samples y_i with a properly chosen impulse response.

When the samples are corrupted by noise, the choice $\mathbf{G} = \mathbf{G}_{\text{CON}}$ no longer guarantees perfect reconstruction of $x \in \mathcal{W}$. Our problem then is to choose \mathbf{G} such that \hat{x} is close to x in some sense, for any $x \in \mathcal{W}$. Formulating the problem in a statistical framework, we seek the transformation \mathbf{G} that minimizes the MSE $E\{\|\hat{x} - x\|^2\}$ between x and \hat{x} . However, computing the MSE shows that

$$E\{\|\hat{x} - x\|^2\} = \|(\mathbf{I} - W\mathbf{G}S^*)x\|^2 + \text{Tr}(W^*W\mathbf{G}\mathbf{C}\mathbf{G}^*), \quad (10)$$

which depends in general on x and therefore cannot be minimized. To eliminate the signal dependency of the MSE we need to choose \mathbf{G} such that $W\mathbf{G}S^*x = x$ for all $x \in \mathcal{W}$. Equivalently, \mathbf{G} must satisfy

$$W\mathbf{G}S^*P_{\mathcal{W}} = P_{\mathcal{W}}. \quad (11)$$

Multiplying (11) on the left by S^* and on the right by $W(S^*W)^{-1}$, the unique solution to (11) is $\mathbf{G} = (S^*W)^{-1} = \mathbf{G}_{\text{CON}}$. Thus, the only unbiased strategy is the consistent reconstruction; however, this method does not necessarily result in a small MSE.

Unless we use the consistent approach, the MSE performance of \hat{x} in Fig. 1 will depend on the unknown x , and therefore

cannot be evaluated. Nonetheless, some reconstructions may be better than others in an MSE sense. For example, we will see in Section IV-C that if $\|x\| \leq L$ for some $L < \infty$, then based on a minimax MSE approach we can choose \mathbf{G} such that the resulting MSE is smaller than the MSE of \mathbf{G}_{CON} for all $x \in \mathcal{W}$, $\|x\| \leq L$; thus, the minimax MSE reconstruction *dominates* the consistent method for norm bounded signals. In general, a reconstruction \hat{x}_1 dominates a reconstruction \hat{x}_2 on a set \mathcal{U} of \mathcal{W} if

$$\begin{aligned} E \{ \|\hat{x}_1 - x\|^2 \} &\leq E \{ \|\hat{x}_2 - x\|^2 \}, \quad \text{for all } x \in \mathcal{U}; \\ E \{ \|\hat{x}_1 - x\|^2 \} &< E \{ \|\hat{x}_2 - x\|^2 \}, \quad \text{for some } x \in \mathcal{U}. \end{aligned} \quad (12)$$

Furthermore, \hat{x}_1 *strictly dominates* \hat{x}_2 on \mathcal{U} if

$$E \{ \|\hat{x}_1 - x\|^2 \} < E \{ \|\hat{x}_2 - x\|^2 \}, \quad \text{for all } x \in \mathcal{U}. \quad (13)$$

A reconstruction \hat{x}_1 that dominates \hat{x}_2 is clearly preferable in terms of MSE. However, the fact that \hat{x}_1 dominates \hat{x}_2 does not preclude the possibility that another linear strategy exists that dominates \hat{x}_1 , and is therefore preferable, from an MSE perspective, over both \hat{x}_1 and \hat{x}_2 . An *admissible* reconstruction has the property that it is not dominated by any other linear method. If a reconstruction is inadmissible, then there exists another approach that leads to better MSE performance on \mathcal{U} . This discussion raises two interesting questions.

- 1) Given a correction matrix \mathbf{G}_0 , can we verify whether it leads to an admissible reconstruction on a set \mathcal{U} of \mathcal{W} ?
- 2) If \mathbf{G}_0 results in an inadmissible reconstruction, then can we develop a systematic approach for constructing an admissible method \mathbf{G} that dominates \mathbf{G}_0 on \mathcal{U} ?

Based on the results of [20], a general answer to both of these problems is given in the following theorem.

Theorem 1: Let $\mathbf{y} = S^*x + \mathbf{n}$ denote noisy samples of a signal x in an m -dimensional subspace $\mathcal{W} \subseteq \mathcal{H}$, where S is a given set transformation and \mathbf{n} is a zero-mean random vector with covariance $\mathbf{C} \succ 0$. Suppose that $\hat{x}_0 = W\mathbf{G}_0\mathbf{y}$ is a linear reconstruction of x in \mathcal{W} , where W is a given set transformation corresponding to a basis for \mathcal{W} . Denote the MSE of \hat{x}_0 by $\epsilon(\mathbf{G}_0, x) = E\{\|W\mathbf{G}_0\mathbf{y} - x\|^2\}$, and let $\mathcal{U} \subseteq \mathcal{W}$. Then, we have the following:

- 1) \hat{x}_0 is admissible on \mathcal{U} if and only if $\hat{\mathbf{G}} = \mathbf{G}_0$, or equivalently, if and only if $\mathcal{T}(\mathbf{G}_0) = 0$, where $\hat{\mathbf{G}}$ is the unique solution to

$$\min_{\mathbf{G}} \sup_{x \in \mathcal{U}} \{ \epsilon(\mathbf{G}, x) - \epsilon(\mathbf{G}_0, x) \}, \quad (14)$$

and

$$\mathcal{T}(\mathbf{G}_0) = \min_{\mathbf{G}} \sup_{x \in \mathcal{U}} \{ \epsilon(\mathbf{G}, x) - \epsilon(\mathbf{G}_0, x) \}; \quad (15)$$

- 2) if $\mathcal{T}(\mathbf{G}_0) < 0$ then $\hat{x} = W\hat{\mathbf{G}}\mathbf{y}$ strictly dominates \hat{x}_0 on \mathcal{U} ;
- 3) $\hat{x} = W\hat{\mathbf{G}}\mathbf{y}$ is admissible on \mathcal{U} .

Note that the minimum in (14) and (15) is well defined since the objective is continuous and coercive [25].

Proof: The proof follows immediately from the proof of [20, Theorem 1] by noting that $\epsilon(\mathbf{G}, x)$ is continuous, coercive, and strictly convex in \mathbf{G} . ■

Theorem 1 provides a general recipe for determining admissibility of a linear reconstruction and for constructing admissible and strictly dominating methods, by solving a convex optimization problem (the problem (14) is convex for arbitrary constraint sets \mathcal{U} since the supremum of a convex function over any set \mathcal{U} is convex). As noted in [20], an interesting consequence of the theorem is that if a reconstruction \hat{x}_0 is inadmissible, then there exists a method that *strictly* dominates it.

C. Relation With Linear Estimation in Linear Models

The sampling problem of Fig. 1 is very similar to the standard linear regression model, in which the goal is to estimate a parameter vector \mathbf{x} from observations \mathbf{y} which are related through the linear equation

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \quad (16)$$

Here \mathbf{H} is a given model matrix, which is often assumed to have full column rank, and \mathbf{n} is a noise vector. Replacing \mathbf{H} in (16) with $S^*P_{\mathcal{W}}$ results in the sampling model (8) (since we explicitly assume in our development that $x \in \mathcal{W}$).

In this paper, we exploit the relationship between the two models to apply ideas and concepts developed in [20] in the context of linear regression to our sampling problem. In applying these results, there are two main differences which need to be accounted for. The first is that the admissibility and domination conditions of [20] only treat the case in which \mathbf{H} has full column rank. However, when \mathcal{W} is a true subset of \mathcal{H} , the null space of $S^*P_{\mathcal{W}}$ is nonzero which is equivalent to a rank-deficient \mathbf{H} . Another important discrepancy between the two problems is that linear estimators designed to estimate \mathbf{x} in the regression model are typically not restricted. In contrast, our reconstruction \hat{x} is forced to lie in a given subspace \mathcal{W} .

A naive approach to developing admissible reconstruction strategies is to use any admissible estimation method for (16) with \mathbf{H} replaced by $S^*P_{\mathcal{W}}$, and then orthogonally project the estimate onto \mathcal{W} . Unfortunately, the characterization of all admissible methods in [20] only treats the case in which \mathbf{H} has full column rank, which is not relevant to our sampling problem. Nonetheless, in many special cases, starting from an admissible linear estimator for (16) when \mathbf{H} is full rank, substituting $S^*P_{\mathcal{W}}$ for \mathbf{H} and orthogonally projecting the solution results in an admissible reconstruction strategy. As an example, assuming that $\mathbf{C} = \mathbf{I}$, the least-squares estimator

$$\hat{\mathbf{x}}_{\text{LS}} = (\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^*\mathbf{y} \quad (17)$$

is admissible for the model (16) when \mathbf{x} is unrestricted [20]. In order to apply $\hat{\mathbf{x}}_{\text{LS}}$ to (8) we first note that it can be alternatively expressed as

$$\hat{\mathbf{x}}_{\text{LS}} = \mathbf{H}^*(\mathbf{H}\mathbf{H}^*)^\dagger\mathbf{y}. \quad (18)$$

Replacing \mathbf{H} by $S^*P_{\mathcal{W}}$ in (18) and orthogonally projecting the estimator onto \mathcal{W} leads to the reconstruction

$$\hat{x} = P_{\mathcal{W}}S(S^*P_{\mathcal{W}}S)^{-1}\mathbf{y}. \quad (19)$$

From the admissibility conditions we develop in Section V-B (see Theorem 7), it follows that \hat{x} of (19) is indeed an admissible reconstruction strategy.

In the rest of the paper, we use Theorem 1 to systematically develop necessary and sufficient conditions for a reconstruction to be admissible which are easy to verify, as well as an explicit procedure for constructing a reconstruction dominating an inadmissible method. We also propose several admissible strategies by adapting known estimation methods to our particular sampling framework. Motivated by the fact that in many practical scenarios the norm of the signal is bounded, in Sections III and IV we consider the case in which \mathcal{U} consists of vectors $x \in \mathcal{W}$ satisfying $\|x\| \leq L$. Admissibility over the entire space \mathcal{W} is treated in Section V.

III. MSE PERFORMANCE FOR NORM-BOUNDED SIGNALS

Suppose we are given a reconstruction \hat{x}_0 that is inadmissible on the set \mathcal{U} of norm-bounded signals defined by

$$\mathcal{U} = \{x \in \mathcal{W} \mid \|x\| \leq L\}. \quad (20)$$

We would like to construct an admissible reconstruction \hat{x} that strictly dominates \hat{x}_0 . We defer the discussion on conditions under which \hat{x}_0 is inadmissible to Section III-C; for now, we focus on the design problem. Applying Theorem 1, \hat{x} can be constructed as $\hat{x} = W\hat{\mathbf{G}}\mathbf{y}$, where

$$\hat{\mathbf{G}} = \arg \min_{\mathbf{G}} \max_{x \in \mathcal{U}} \{\epsilon(\mathbf{G}, x) - \epsilon(\mathbf{G}_0, x)\}. \quad (21)$$

Here $\epsilon(\mathbf{G}_0, x) = E\{\|W\mathbf{G}_0\mathbf{y} - x\|^2\}$ is the MSE of \hat{x}_0 .

In Section III-A, we derive optimality conditions on $\hat{\mathbf{G}}$ and develop closed form solutions for some special cases. A semidefinite programming (SDP) formulation is presented in Section III-B. Using these results, in Section III-C, we provide necessary and sufficient conditions on a reconstruction \hat{x} to be admissible on \mathcal{U} , which are easy to verify. Specific examples of reconstruction strategies are presented in Section IV.

A. Necessary and Sufficient Conditions for Optimality

Substituting (10) into (21), $\hat{\mathbf{G}}$ can be determined as

$$\hat{\mathbf{G}} = \arg \min_{\mathbf{G}} \left\{ \text{Tr}(W^*W\mathbf{G}\mathbf{C}\mathbf{G}^*) + \max_{x \in \mathcal{U}} \langle x, Zx \rangle \right\}, \quad (22)$$

where we defined

$$Z = (I - W\mathbf{G}S^*)(I - W\mathbf{G}S^*)^* - (I - W\mathbf{G}_0S^*)(I - W\mathbf{G}_0S^*)^*. \quad (23)$$

Our approach to solving (22) is to apply a change of variables that reduces the problem to a form similar to that considered in [20] in the context of estimation in linear models. Adapting these results to our setting leads to the optimality conditions given in Theorem 2 below. The detailed proof of the theorem is presented in Appendix A.

Theorem 2: Let $\mathbf{y} = S^*x + \mathbf{n}$ denote noisy samples of a signal x in an m -dimensional subspace $\mathcal{W} \subseteq \mathcal{H}$, where S is the set transformation corresponding to the sampling vectors $\{s_i, 1 \leq i \leq m\}$ which form a basis for an m -dimensional subspace $\mathcal{S} \subseteq \mathcal{H}$ such that $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$, and \mathbf{n} is a zero-mean

random vector with covariance $\mathbf{C} \succ 0$. Let $\{w_i, 1 \leq i \leq m\}$ denote reconstruction vectors with set transformation W that form a basis for \mathcal{W} , let $\mathcal{U} = \{x \in \mathcal{W} \mid \|x\| \leq L\}$, and let $\hat{x}_0 = W\mathbf{G}_0\mathbf{y}$ be a given reconstruction of x with MSE $\epsilon(\mathbf{G}_0, x) = E\{\|W\mathbf{G}_0\mathbf{y} - x\|^2\}$. Then,

$$\hat{\mathbf{G}} = \arg \min_{\mathbf{G}} \max_{x \in \mathcal{U}} \{\epsilon(\mathbf{G}, x) - \epsilon(\mathbf{G}_0, x)\}$$

if and only if $\hat{\mathbf{G}}$ satisfies the following conditions for some $\lambda \geq 0$ with $R = W\hat{\mathbf{G}}S^*P_{\mathcal{W}}$ and $R_0 = W\mathbf{G}_0S^*P_{\mathcal{W}}$:

- 1) $\mathbf{C}^{-1}S^*W\hat{\mathbf{G}} = \hat{\mathbf{G}}^*W^*S\mathbf{C}^{-1}$;
- 2) $0 \preceq \mathbf{C}^{-1}S^*W\hat{\mathbf{G}} \prec \mathbf{C}^{-1}$;
- 3) $(P_{\mathcal{W}} - R)^*(P_{\mathcal{W}} - R) \preceq \lambda P_{\mathcal{W}} + (P_{\mathcal{W}} - R_0)^*(P_{\mathcal{W}} - R_0)$;
- 4) $(P_{\mathcal{W}} - R)^*(P_{\mathcal{W}} - R)R = \lambda R + (P_{\mathcal{W}} - R_0)^*(P_{\mathcal{W}} - R_0)R$;
- 5) $\text{Tr}(W^*W\hat{\mathbf{G}}(\mathbf{C}^{-1} - \mathbf{C}^{-1}S^*W\hat{\mathbf{G}})^{-1}(W^*S)^{-1}) \leq L^2$;
- 6) $\lambda \text{Tr}(W^*W\hat{\mathbf{G}}(\mathbf{C}^{-1} - \mathbf{C}^{-1}S^*W\hat{\mathbf{G}})^{-1}(W^*S)^{-1}) = \lambda L^2$.

Although the conditions of Theorem 2 are difficult to solve in general, they can be used to verify a solution. For example, we can check when $\hat{\mathbf{G}} = \mathbf{G}_0$, which leads to necessary and sufficient conditions for admissibility, as we discuss in Section III-C. Furthermore, in some special cases the conditions can be solved explicitly. We now consider one such class of examples, in which $P_{\mathcal{W}}S\mathbf{C}^{-1}S^*P_{\mathcal{W}}$ and $W\mathbf{G}_0S^*P_{\mathcal{W}}$ have the same eigenvectors.

1) *Special Case:* Let $P_{\mathcal{W}}S\mathbf{C}^{-1}S^*P_{\mathcal{W}}$ have an eigendecomposition

$$P_{\mathcal{W}}S\mathbf{C}^{-1}S^*P_{\mathcal{W}} = U\Sigma U^*, \quad (24)$$

where $U: \mathbb{C}^m \rightarrow \mathcal{H}$ is an orthonormal set transformation satisfying $U^*U = I$ and $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_m)$ with $\sigma_i > 0$, and suppose that

$$W\mathbf{G}_0S^*P_{\mathcal{W}} = U\Delta U^*, \quad (25)$$

where $\Delta = \text{diag}(\delta_1, \dots, \delta_m)$ for some δ_i .

A general case in which (25) is satisfied is when $\mathbf{C} = \sigma^2\mathbf{I}$ for some $\sigma^2 > 0$, $P_{\mathcal{W}}S = W$, and $W^*W = I$ so that $\{w_i\}$ form an orthonormal basis for \mathcal{W} . Indeed, $P_{\mathcal{W}}S\mathbf{C}^{-1}S^*P_{\mathcal{W}} = (1/\sigma^2)P_{\mathcal{W}}$ which commutes with $W\mathbf{G}_0S^*P_{\mathcal{W}}$, and therefore in this setting the two matrices have the same eigenvectors.

Theorem 3 below provides a closed form expression for the optimal matrix $\hat{\mathbf{G}}$ under these conditions.

Theorem 3: Consider the problem of Theorem 2. Let $P_{\mathcal{W}}S\mathbf{C}^{-1}S^*P_{\mathcal{W}} = U\Sigma U^*$ where $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_m)$ with $\sigma_i > 0$, and suppose that $W\mathbf{G}_0S^*P_{\mathcal{W}} = U\Delta U^*$ where $\Delta = \text{diag}(\delta_1, \dots, \delta_m)$. Then

$$\hat{\mathbf{G}} = (W^*W)^{-1}W^*U\mathbf{D}U^*W(S^*W)^{-1}$$

where $\mathbf{D} = \text{diag}(d_1, \dots, d_m)$ with

$$d_i = \begin{cases} 1 - \sqrt{\eta_i}, & \eta_i < 1; \\ 0, & \eta_i \geq 1, \end{cases} \quad (26)$$

and

$$\eta_i = \lambda + |1 - \delta_i|^2. \quad (27)$$

²Note, that since $\mathcal{S}^\perp \cap \mathcal{W} = \{0\}$, $S^*P_{\mathcal{W}}S$ is invertible, and the rank of $P_{\mathcal{W}}S\mathbf{C}^{-1}S^*P_{\mathcal{W}}$ is equal to m .

The parameter λ is selected according to the following procedure: if $T(0) \leq 0$, where

$$T(\lambda) = \sum_{i:\eta_i < 1} \frac{1}{\sigma_i} \left(\frac{1}{\sqrt{\eta_i}} - 1 \right) - L^2, \quad (28)$$

then $\lambda = 0$. Otherwise, λ is the unique value for which $T(\lambda) = 0$ in the range $(0, \alpha)$ with $\alpha = 1 - \min_i |1 - \delta_i|^2$.

Proof: See Appendix B. \blacksquare

We conclude that if we are given an inadmissible \mathbf{G}_0 satisfying (25), then Theorem 3 can be used to construct an admissible reconstruction strictly dominating \mathbf{G}_0 on \mathcal{U} . An example of an application of this theorem is presented in Section IV-C.

B. Semidefinite Programming Formulation

Theorem 2 can be used to derive explicit solutions to (22) in some special cases. Following [20], we now show that in the general case, the optimal \mathbf{G} can be found numerically by solving an SDP [26], [27], which is the problem of minimizing a linear objective subject to linear matrix inequality constraints. These are matrix constraints of the form $\mathbf{Z}(\mathbf{x}) \succeq 0$, where the matrix \mathbf{Z} depends linearly on the vector \mathbf{x} . The advantage of this formulation is that it readily lends itself to efficient computational methods that are guaranteed to converge to the global optimum within any desired accuracy [26]–[28]. In practice, the solution can be obtained using one of the many available SDP software packages.

To develop the SDP formulation we define the matrices $\mathbf{M} = U^*W\mathbf{G}$, $\mathbf{H} = S^*U$ and

$$\mathbf{A} = U^*(\mathbf{I} - W\mathbf{G}_0S^*)^*(\mathbf{I} - W\mathbf{G}_0S^*)U, \quad (29)$$

where $U: \mathbb{C}^m \rightarrow \mathcal{H}$ is the orthonormal set transformation in the SVD of W (see (78) in Appendix A). As detailed in Appendix A, the problem (22) can be formulated in terms of these matrices as

$$\min_{\mathbf{M}} \left\{ \text{Tr}(\mathbf{MCM}^*) + \max_{\|\mathbf{d}\| \leq L} \mathbf{d}^* ((\mathbf{I} - \mathbf{MH})^*(\mathbf{I} - \mathbf{MH}) - \mathbf{A}) \mathbf{d} \right\}. \quad (30)$$

Now, for any matrix \mathbf{Z} , we have that

$$\max_{\|\mathbf{d}\| \leq L} \mathbf{d}^* \mathbf{Z} \mathbf{d} = L^2 \max(\lambda_{\max}(\mathbf{Z}), 0), \quad (31)$$

where $\lambda_{\max}(\mathbf{Z})$ is the largest eigenvalue of \mathbf{Z} . Expressing $\max(\lambda_{\max}(\mathbf{Z}), 0)$ as the solution to

$$\min_{\lambda \geq 0} \{ \lambda : \mathbf{Z} \preceq \lambda \mathbf{I} \}, \quad (32)$$

equation (30) becomes

$$\min_{\mathbf{M}, \lambda \geq 0} \{ \text{Tr}(\mathbf{MCM}^*) + L^2 \lambda \} \quad (33)$$

subject to

$$(\mathbf{I} - \mathbf{MH})^*(\mathbf{I} - \mathbf{MH}) - \mathbf{A} \preceq \lambda \mathbf{I}, \quad (34)$$

or, equivalently,

$$\min_{\tau, \mathbf{M}, \lambda \geq 0} \tau \quad (35)$$

subject to

$$\begin{aligned} \mathbf{m}^* \mathbf{m} + L^2 \lambda &\leq \tau \\ (\mathbf{I} - \mathbf{MH})^*(\mathbf{I} - \mathbf{MH}) - \mathbf{A} &\preceq \lambda \mathbf{I}. \end{aligned} \quad (36)$$

Here $\mathbf{m} = \text{vec}(\mathbf{C}^{1/2}\mathbf{M})$ is the vector obtained by stacking the columns of $\mathbf{C}^{1/2}\mathbf{M}$. Using Schur's complement [29, p. 472], the constraints (36) can be written as the linear matrix inequalities

$$\begin{aligned} \begin{bmatrix} \tau - L^2 \lambda & \mathbf{m}^* \\ \mathbf{m} & \mathbf{I} \end{bmatrix} &\succeq 0 \\ \begin{bmatrix} \lambda \mathbf{I} + \mathbf{A} & (\mathbf{I} - \mathbf{MH})^* \\ \mathbf{I} - \mathbf{MH} & \mathbf{I} \end{bmatrix} &\succeq 0. \end{aligned} \quad (37)$$

We conclude that the problem (30) is equivalent to the SDP of minimizing τ subject to (37). The solution $\hat{\mathbf{G}}$ to (22) can be obtained from the SDP solution $\hat{\mathbf{M}}$ as $\hat{\mathbf{G}} = (U^*W)^{-1}\hat{\mathbf{M}}$.

C. Admissible Reconstructions

We now use Theorem 2 to develop necessary and sufficient conditions on \hat{x}_0 to be admissible, or equivalently, such that \mathbf{G}_0 is the solution to (21).

Theorem 4: Consider the problem of Theorem 2. Then $\hat{x} = W\hat{\mathbf{G}}\mathbf{y}$ is an admissible reconstruction of x on \mathcal{U} if and only if

- 1) $\mathbf{C}^{-1}S^*W\hat{\mathbf{G}} = \hat{\mathbf{G}}^*W^*S\mathbf{C}^{-1}$;
- 2) $0 \preceq \mathbf{C}^{-1}S^*W\hat{\mathbf{G}} \prec \mathbf{C}^{-1}$;
- 3) $\text{Tr}(W^*W\hat{\mathbf{G}}(\mathbf{C}^{-1} - \mathbf{C}^{-1}S^*W\hat{\mathbf{G}})^{-1}(W^*S)^{-1}) \leq L^2$.

Proof: To prove the theorem it is sufficient to show that $\hat{\mathbf{G}}$ satisfies the conditions of Theorem 2 with $\hat{\mathbf{G}} = \mathbf{G}_0$ for some $\lambda \geq 0$ if and only if it satisfies conditions 1–3.

Since conditions 1–3 are a subset of those of Theorem 2, any $\hat{\mathbf{G}}$ satisfying Theorem 2 also satisfies these conditions. On the other hand, if $\hat{\mathbf{G}}$ satisfies conditions 1–3, then it satisfies the conditions of Theorem 2 with $\lambda = 0$. \blacksquare

An immediate consequence of Theorem 4 is that the consistent reconstruction $\hat{x}_{\text{CON}} = W\mathbf{G}_{\text{CON}}\mathbf{y}$ with \mathbf{G}_{CON} given by (4) is inadmissible on the bounded-norm set \mathcal{U} . Indeed, $S^*W\mathbf{G}_{\text{CON}} = \mathbf{I}$, and the second constraint of the theorem is violated. Since \hat{x}_{CON} is inadmissible, an interesting problem is to construct an alternative reconstruction \hat{x} that is admissible and strictly dominates \hat{x}_{CON} . In Section IV-C, we show that a reconstruction with this property is the minimax MSE approach, which minimizes the worst-case MSE on \mathcal{U} .

A general class of admissible reconstructions is given in the following proposition.

Proposition 2: Consider the problem of Theorem 2. Let $P_W S \mathbf{C}^{-1} S^* P_W = U \Sigma U^*$ where $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_m)$ with $\sigma_i > 0$. Then $\hat{x} = W\hat{\mathbf{G}}\mathbf{y}$ is admissible for any $\hat{\mathbf{G}}$ of the form

$$\hat{\mathbf{G}} = (W^*W)^{-1}W^*U\mathbf{D}U^*W(S^*W)^{-1}, \quad (38)$$

where $\mathbf{D} = \text{diag}(d_1, \dots, d_m)$ and the values d_i satisfy

$$0 \leq d_i < 1, \quad 1 \leq i \leq m; \quad \sum_{i=1}^m \frac{d_i}{(1-d_i)\sigma_i} \leq L^2.$$

Proof: In the proof of Theorem 3 we showed that $\hat{\mathbf{G}}$ of the form (38) satisfies conditions 1 and 2 of Theorem 3 if $0 \leq$

$d_i < 1$. We also showed that the trace in condition 3 is equal to $\sum_{i=1}^m [d_i / ((1 - d_i)\sigma_i)]$, from which the proposition follows. ■

IV. RECONSTRUCTION METHODS ON A SPHERE

Theorem 4 implies that there are many possible admissible reconstructions on \mathcal{U} . Therefore, there is a need for a criterion to choose between them. In this section, we consider several strategies and develop the corresponding solutions. The design methods are based on standard estimation approaches for the linear regression model (16), adapted to our particular sampling problem.

A. Least-Squares

The most straightforward approach to designing \mathbf{G} is to consider a least-squares criterion, in which we minimize the (weighted) data error

$$\begin{aligned} \epsilon_{\text{LS}} &= (S^*\hat{x} - \mathbf{y})^* \mathbf{C}^{-1} (S^*\hat{x} - \mathbf{y}) \\ &= (S^*W\mathbf{G}\mathbf{y} - \mathbf{y})^* \mathbf{C}^{-1} (S^*W\mathbf{G}\mathbf{y} - \mathbf{y}). \end{aligned} \quad (39)$$

Clearly, ϵ_{LS} is minimized with $\mathbf{G} = \mathbf{G}_{\text{CON}}$ of (4), and is therefore inadmissible on a sphere.

It is interesting to note that from Proposition 1, the least-squares reconstruction $\hat{x} = W\mathbf{G}_{\text{CON}}\mathbf{y}$ is equal to the reconstruction (19) which is obtained by a projection onto \mathcal{W} of the least-squares estimator for the model (16), with \mathbf{H} replaced by $S^*P_{\mathcal{W}}$.

B. Tikhonov Regularization

The least-squares strategy does not take advantage of the prior information that x has bounded norm. To take this into account, we may seek the reconstructed signal $\hat{x} = W\mathbf{G}\mathbf{y}$ that minimizes the least-squares error (39) subject to the constraint $\|\hat{x}\| \leq L$. The solution \hat{x} can be determined by minimizing the Lagrangian

$$(S^*W\mathbf{G}\mathbf{y} - \mathbf{y})^* \mathbf{C}^{-1} (S^*W\mathbf{G}\mathbf{y} - \mathbf{y}) + \lambda \mathbf{y}^* \mathbf{G}^* W^* W \mathbf{G} \mathbf{y}, \quad (40)$$

where from the Karush–Kuhn–Tucker conditions [25] $\lambda \geq 0$ and at the optimal solution

$$\lambda (\mathbf{y}^* \mathbf{G}^* W^* W \mathbf{G} \mathbf{y} - L^2) = 0; \quad (41)$$

thus, either $\|\hat{x}\| = L$ or $\lambda = 0$.

Differentiating (40) with respect to \mathbf{G} and equating to 0,

$$W^* (S\mathbf{C}^{-1}S^* + \lambda\mathbf{I}) W \mathbf{G} \mathbf{y} \mathbf{y}^* = W^* S \mathbf{C}^{-1} \mathbf{y} \mathbf{y}^*, \quad (42)$$

which is satisfied for all \mathbf{y} if

$$\mathbf{G} = (W^* (S\mathbf{C}^{-1}S^* + \lambda\mathbf{I}) W)^{-1} W^* S \mathbf{C}^{-1}. \quad (43)$$

Using some algebraic manipulations, \mathbf{G} of (43) can be shown to be equivalent to

$$\mathbf{G} = (W^*W)^{-1} W^* S (S^*P_{\mathcal{W}}S + \lambda\mathbf{C})^{-1}, \quad (44)$$

where λ must satisfy (41). Denoting by $\mathbf{G}(\lambda)$ the matrix \mathbf{G} of (44) for fixed λ , this implies the following procedure for selecting λ : if $\mathbf{y}^* \mathbf{G}(0) W^* W \mathbf{G}(0) \mathbf{y} \leq L^2$, then $\lambda = 0$, and using Proposition 1, $\mathbf{G} = (S^*W)^{-1} = \mathbf{G}_{\text{CON}}$. Otherwise, $\lambda > 0$ is a parameter that depends on the noisy observations \mathbf{y} and is chosen such that

$$\mathbf{y}^* \mathbf{G}(\lambda) W^* W \mathbf{G}(\lambda) \mathbf{y} = L^2. \quad (45)$$

To show that such a λ always exists, let

$$\mathcal{G}(\lambda) = \mathbf{y}^* \mathbf{G}(\lambda) W^* W \mathbf{G}(\lambda) \mathbf{y} - L^2, \quad (46)$$

so that λ is a positive root of $\mathcal{G}(\lambda)$. Clearly, $\mathcal{G}(\lambda)$ is monotonically decreasing in λ . In addition, $\mathcal{G}(0) > 0$, and $\mathcal{G}(\lambda) \rightarrow -L^2$ as $\lambda \rightarrow \infty$, so that $\mathcal{G}(\lambda)$ has exactly one positive root.

We conclude that the Tikhonov matrix \mathbf{G}_{TIK} is given by (47) at the bottom of the page, where $\lambda > 0$ is chosen such that $\mathbf{y}^* \mathbf{G}_{\text{TIK}} W^* W \mathbf{G}_{\text{TIK}} \mathbf{y} = L^2$.

Evidently, Tikhonov reconstruction is in general *nonlinear*, and does not have an explicit solution; the parameter λ does not have a closed form, but is rather determined as the solution of the data-dependent (45). However, instead of choosing λ according to (45) we may select λ as a constant such that the resulting reconstruction is admissible. The possible values of λ satisfying this condition are given in the following proposition.

Proposition 3: The Tikhonov reconstruction $\hat{x}_{\text{TIK}} = W\mathbf{G}_{\text{TIK}}\mathbf{y}$ with $\mathbf{G}_{\text{TIK}} = (W^*W)^{-1} W^* S (S^*P_{\mathcal{W}}S + \lambda\mathbf{C})^{-1}$ is admissible on $\mathcal{U} = \{x \in \mathcal{W} \mid \|x\| \leq L\}$ if and only if $\lambda \geq m/L^2$.

Proof: We prove the proposition by showing that the conditions of Theorem 4 are satisfied if and only if $\lambda \geq m/L^2$.

First, we note that

$$\begin{aligned} \mathbf{C}^{-1} S^* W \mathbf{G}_{\text{TIK}} &= \mathbf{C}^{-1} S^* P_{\mathcal{W}} S (S^* P_{\mathcal{W}} S + \lambda \mathbf{C})^{-1} \\ &= \mathbf{C}^{-1} - \left(\left(\frac{1}{\lambda} \right) S^* P_{\mathcal{W}} S + \mathbf{C} \right)^{-1}, \end{aligned} \quad (48)$$

where we used the matrix inversion lemma [29]. It therefore follows that $\mathbf{C}^{-1} S^* W \mathbf{G}_{\text{TIK}}$ is Hermitian, so that the first condition holds. Since $(1/\lambda) S^* P_{\mathcal{W}} S + \mathbf{C} \succeq \mathbf{C} \succ 0$, the second condition is also satisfied. To verify the last condition, we note that by direct substitution of $\mathbf{G} = \mathbf{G}_{\text{TIK}}$ we have $\text{Tr}(W^* W \mathbf{G} (\mathbf{C}^{-1} - \mathbf{C}^{-1} S^* W \mathbf{G})^{-1} (W^* S)^{-1}) = (1/\lambda) \text{Tr}(\mathbf{I})$. Therefore, the last condition is satisfied if and only if $m/\lambda \leq L^2$, completing the proof. ■

$$\mathbf{G}_{\text{TIK}} = \begin{cases} (S^*W)^{-1}, & \|W(S^*W)^{-1}\mathbf{y}\| \leq L; \\ (W^*W)^{-1}W^*S(S^*P_{\mathcal{W}}S + \lambda\mathbf{C})^{-1}, & \text{otherwise} \end{cases} \quad (47)$$

Similarly to the least-squares design method, the Tikhonov reconstruction can be shown to be equivalent to the orthogonal projection onto \mathcal{W} of the standard Tikhonov estimator for the linear regression model (16). To see this, recall that the Tikhonov estimator [23], [30] for (16) is given by the value of $\hat{\mathbf{x}}$ that minimizes the regularized squared error

$$(\mathbf{H}\hat{\mathbf{x}} - \mathbf{y})^* \mathbf{C}^{-1} (\mathbf{H}\hat{\mathbf{x}} - \mathbf{y}) + \lambda \|\hat{\mathbf{x}}\|^2, \quad (49)$$

and is equal to

$$\hat{\mathbf{x}} = (\mathbf{H}^* \mathbf{C}^{-1} \mathbf{H} + \lambda \mathbf{I})^{-1} \mathbf{H}^* \mathbf{C}^{-1} \mathbf{y} = \mathbf{H}^* (\mathbf{H} \mathbf{H}^* + \lambda \mathbf{C})^{-1} \mathbf{y}. \quad (50)$$

Replacing \mathbf{H} with $S^* P_{\mathcal{W}}$, (50) becomes

$$\hat{\mathbf{x}} = P_{\mathcal{W}} S (S^* P_{\mathcal{W}} S + \lambda \mathbf{C})^{-1} \mathbf{y} = W \mathbf{G}_{\text{TIK}} \mathbf{y}. \quad (51)$$

C. Minimax MSE Reconstruction

From Proposition 3, it follows that we can choose the regularization parameter in the Tikhonov reconstruction such that $\hat{\mathbf{x}}_{\text{TIK}}$ is admissible on \mathcal{U} . However, in general $\hat{\mathbf{x}}_{\text{TIK}}$ does not dominate the consistent reconstruction $\hat{\mathbf{x}}_{\text{CON}}$, as we demonstrate in the context of a concrete example in Section IV-D. Since $\hat{\mathbf{x}}_{\text{CON}}$ is inadmissible on \mathcal{U} , a strictly dominating method exists. We now consider a minimax MSE approach and show that the resulting reconstruction dominates $\hat{\mathbf{x}}_{\text{CON}}$ on \mathcal{U} .

Instead of minimizing a regularized data function, we may seek the reconstruction $\hat{\mathbf{x}}_{\text{MX}} = W \mathbf{G}_{\text{MX}} \mathbf{y}$ that minimizes the worst-case MSE over all bounded-norm signals, so that \mathbf{G}_{MX} is the solution to

$$\mathbf{G}_{\text{MX}} = \arg \min_{\mathbf{G}} \max_{x \in \mathcal{W}, \|x\| \leq L} \epsilon(\mathbf{G}, x). \quad (52)$$

Clearly, if the reconstruction is unique, then it is admissible since it minimizes the MSE at the worst case and therefore cannot be dominated by any other method.

To develop the minimax MSE estimator we note that it can be alternatively expressed as

$$\mathbf{G}_{\text{MX}} = \arg \min_{\mathbf{G}} \max_{x \in \mathcal{W}, \|x\| \leq L} \{\epsilon(\mathbf{G}, x) - \epsilon(\mathbf{G}_{\text{CON}})\}, \quad (53)$$

where $\hat{\mathbf{x}}_0 = W \mathbf{G}_{\text{CON}} \mathbf{y}$ is the consistent reconstruction with $\mathbf{G}_{\text{CON}} = (S^* W)^{-1}$. This follows from the fact that, as we have seen in Section II, $\epsilon(\mathbf{G}_{\text{CON}})$ is independent of x . From Theorem 1, we have immediately that $\hat{\mathbf{x}}_{\text{MX}}$ is admissible, and strictly dominates $\hat{\mathbf{x}}_{\text{CON}}$ on \mathcal{U} . Theorem 3 can now be used to derive an explicit solution to (53), resulting in the following proposition.

Proposition 4: Consider the problem of Theorem 2. Let $\hat{\mathbf{x}}_{\text{MX}} = W \mathbf{G}_{\text{MX}} \mathbf{y}$ be the minimax MSE estimator with $\mathbf{G}_{\text{MX}} = \arg \min_{\mathbf{G}} \max_{x \in \mathcal{W}, \|x\| \leq L} \epsilon(\mathbf{G}, x)$. Then,

$$\mathbf{G}_{\text{MX}} = \frac{L^2}{L^2 + \text{Tr}((P_{\mathcal{W}} S \mathbf{C}^{-1} S^* P_{\mathcal{W}})^{\dagger})} (S^* W)^{-1}.$$

In addition, $\hat{\mathbf{x}}_{\text{MX}}$ strictly dominates the consistent reconstruction $\hat{\mathbf{x}}_{\text{CON}} = W (S^* W)^{-1} \mathbf{y}$ on \mathcal{U} .

Proof: With $\mathbf{G}_0 = \mathbf{G}_{\text{CON}}$, we have that $R_0 = W \mathbf{G}_0 S^* P_{\mathcal{W}} = P_{\mathcal{W}}$, where we used the fact that $P_{\mathcal{W}}$

$W(W^* W)^{-1} W^*$. Therefore, R_0 and $P_{\mathcal{W}} S \mathbf{C}^{-1} S^* P_{\mathcal{W}}$ are jointly diagonalizable, so that we can use Theorem 3 to find the minimax MSE reconstruction explicitly. Since the nonzero eigenvalues of $R_0 = P_{\mathcal{W}}$ are all equal 1, $\delta_i = 1$ and $\eta_i = \lambda$. Denoting $\gamma_0 = \sum_{i=1}^m 1/\sigma_i$ it follows that $\sqrt{\lambda} = \gamma_0 / (L^2 + \gamma_0)$ and

$$d_i = \frac{L^2}{L^2 + \gamma_0}, \quad 1 \leq i \leq m, \quad (54)$$

where $\gamma_0 = \text{Tr}((P_{\mathcal{W}} S \mathbf{C}^{-1} S^* P_{\mathcal{W}})^{\dagger})$. Thus,

$$\mathbf{G}_{\text{MX}} = \frac{L^2}{L^2 + \gamma_0} (S^* W)^{-1}, \quad (55)$$

completing the proof. \blacksquare

Note that when $L \rightarrow \infty$, we have $\mathbf{G}_{\text{MX}} \rightarrow \mathbf{G}_{\text{CON}}$. Therefore, when $x \in \mathcal{W}$ is not norm bounded, the minimax MSE and consistent reconstructions coincide.

D. Example

We now provide an example illustrating the performance of the consistent (least-squares), Tikhonov, and minimax MSE reconstructions.

Suppose we are given noisy local averages of a signal $\mathbf{x} \in \mathcal{W}$, where \mathcal{W} is the space of even signals of length 5, spanned by the columns of

$$\mathbf{W} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (56)$$

The samples (local averages) are obtained by taking the inner products with the columns of

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (57)$$

and are corrupted by zero-mean, independent Gaussian noise with variance σ^2 . To reconstruct the signal from the samples, we consider three choices of \mathbf{G} : \mathbf{G}_{CON} of (4), \mathbf{G}_{TIK} of (47) with $\lambda = (5/\|\mathbf{x}\|)^2$ and \mathbf{G}_{MX} of (55) with $L = \|\mathbf{x}\|$.

In Figs. 2 and 3 we plot the MSE as a function of the SNR defined by $-10 \log \sigma^2$, using each of the methods above, for two different choices of \mathbf{x} : $\mathbf{x} = [1 \ 1 \ 1 \ 1 \ 1]^*$ and $\mathbf{x} = [-1.7 \ 0.6 \ -2 \ 0.6 \ -1.7]^*$. As can be seen from the figures, the performance of $\hat{\mathbf{x}}_{\text{TIK}}$ and $\hat{\mathbf{x}}_{\text{MX}}$ depends on the particular choice of \mathbf{x} : for some values of \mathbf{x} , $\hat{\mathbf{x}}_{\text{MX}}$ performs better than $\hat{\mathbf{x}}_{\text{TIK}}$ while for other choices the reverse is true. However, it is evident from Fig. 2 that $\hat{\mathbf{x}}_{\text{TIK}}$ does not dominate $\hat{\mathbf{x}}_{\text{CON}}$. On the other hand, in both figures $\hat{\mathbf{x}}_{\text{MX}}$ dominates $\hat{\mathbf{x}}_{\text{CON}}$ for all SNR, as we expect from the theoretical results of Proposition 4. Note that the performance of $\hat{\mathbf{x}}_{\text{CON}}$ does not depend on \mathbf{x} .

V. MSE PERFORMANCE ON THE ENTIRE SPACE

We now treat the case in which x is an arbitrary vector in \mathcal{W} .

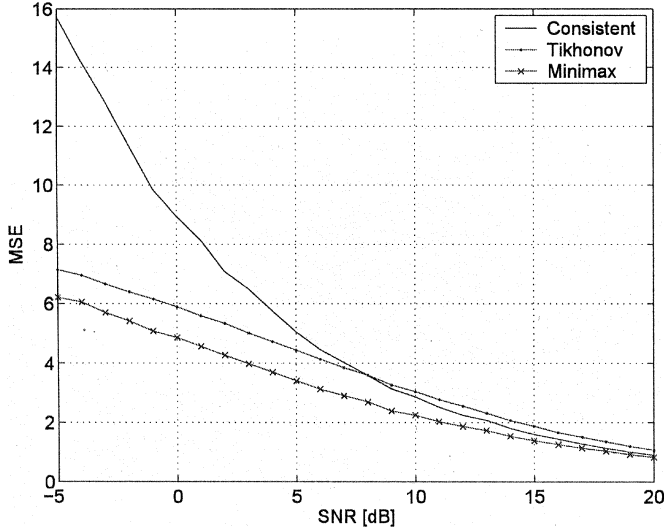


Fig. 2. MSE in reconstructing \mathbf{x} as a function of the SNR for $\mathbf{x} = [-1.7 \ 0.6 \ -2 \ 0.6 \ -1.7]^*$.

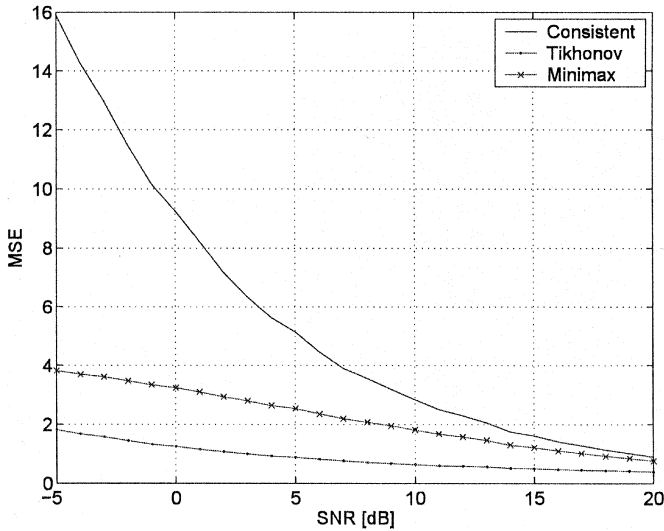


Fig. 3. MSE in reconstructing \mathbf{x} as a function of the SNR for $\mathbf{x} = [1 \ 1 \ 1 \ 1 \ 1]^*$.

A. Dominating Reconstructions on the Entire Space

From Theorem 1, it follows that an admissible reconstruction $\hat{\mathbf{x}}$ strictly dominating an inadmissible method $\hat{\mathbf{x}}_0$ on \mathcal{W} can be constructed as the solution to the problem

$$\min_{\mathbf{G}} \max_{x \in \mathcal{W}} \{\epsilon(\mathbf{G}, x) - \epsilon(\mathbf{G}_0, x)\}. \quad (58)$$

Necessary and sufficient optimality conditions on the solution to (58) are given in the following theorem.

Theorem 5: Consider the problem of Theorem 2. With $R = W\hat{\mathbf{G}}S^*P_{\mathcal{W}}$, $R_0 = W\mathbf{G}_0S^*P_{\mathcal{W}}$, denote the orthogonal projection onto $\mathcal{N}(P_{\mathcal{W}} - R_0)^\perp$ by $P_{\mathcal{R}}$ and define

$$X = (P_{\mathcal{W}}SC^{-1}S^*(P_{\mathcal{W}} - R))^\dagger \\ - P_{\mathcal{R}}W(W^*SC^{-1}S^*W)^{-1}W^*P_{\mathcal{R}}.$$

Then,

$$\hat{\mathbf{G}} = \min_{\mathbf{G}} \max_{x \in \mathcal{W}} \{\epsilon(\mathbf{G}, x) - \epsilon(\mathbf{G}_0, x)\}$$

if and only if $\hat{\mathbf{G}}$ satisfies the following conditions:

- 1) $\mathbf{C}^{-1}S^*W\hat{\mathbf{G}} = \hat{\mathbf{G}}^*W^*S\mathbf{C}^{-1}$;
- 2) $0 \preceq \mathbf{C}^{-1}S^*W\hat{\mathbf{G}} \preceq \mathbf{C}^{-1}$;
- 3) $(P_{\mathcal{W}} - R)P_{\mathcal{R}} = (P_{\mathcal{W}} - R)$;
- 4) $(P_{\mathcal{W}} - R)^*(P_{\mathcal{W}} - R) \preceq (P_{\mathcal{W}} - R_0)^*(P_{\mathcal{W}} - R_0)$;
- 5) $(P_{\mathcal{W}} - R)^*(P_{\mathcal{W}} - R)X = (P_{\mathcal{W}} - R_0)^*(P_{\mathcal{W}} - R_0)X$.

Proof: See Appendix C. ■

As we now show, the conditions of Theorem 5 can be used to derive explicit closed form solutions in the case in which $P_{\mathcal{W}}S\mathbf{C}^{-1}S^*P_{\mathcal{W}}$ and $W\mathbf{G}_0S^*P_{\mathcal{W}}$ have the same eigenvectors. For arbitrary choices of S , W , \mathbf{C}^{-1} and \mathbf{G}_0 the optimal solution can be obtained using an SDP formulation similar to that presented in Section III-B.

Theorem 6: Consider the problem of Theorem 5. Let $P_{\mathcal{W}}S\mathbf{C}^{-1}S^*P_{\mathcal{W}} = U\Sigma U^*$ where $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_m)$ with $\sigma_i > 0$, and suppose that $W\mathbf{G}_0S^*P_{\mathcal{W}} = U\Delta U^*$, where $\Delta = \text{diag}(\delta_1, \dots, \delta_m)$. Then

$$\hat{\mathbf{G}} = (W^*W)^{-1}W^*UDU^*W(S^*W)^{-1} \quad (59)$$

where $\mathbf{D} = \text{diag}(d_1, \dots, d_m)$ with

$$d_i = \begin{cases} 1 - |1 - \delta_i|, & |1 - \delta_i| < 1; \\ 0, & |1 - \delta_i| \geq 1. \end{cases} \quad (60)$$

Proof: See Appendix D. ■

B. Admissible Reconstructions on the Entire Space

Using Theorem 5 leads to the following necessary and sufficient conditions on $\hat{\mathbf{x}}$ to be admissible for all $x \in \mathcal{W}$.

Theorem 7: Consider the problem of Theorem 5. The reconstruction $\hat{\mathbf{x}} = W\hat{\mathbf{G}}\mathbf{y}$ is admissible on \mathcal{W} if and only if

- 1) $\mathbf{C}^{-1}S^*W\hat{\mathbf{G}} = \hat{\mathbf{G}}^*W^*S\mathbf{C}^{-1}$;
- 2) $0 \preceq \mathbf{C}^{-1}S^*W\hat{\mathbf{G}} \preceq \mathbf{C}^{-1}$.

Proof: The theorem follows immediately by noting that $\hat{\mathbf{G}}$ satisfies the conditions of Theorem 5 with $\mathbf{G}_0 = \hat{\mathbf{G}}$ if and only if it satisfies conditions 1 and 2. ■

Note from Theorems 4 and 7 that any reconstruction which is admissible on $x \in \mathcal{W}$, $\|x\| \leq L$ is also admissible on \mathcal{W} .

A general class of admissible reconstructions is given in the following proposition.

Proposition 5: Consider the problem of Theorem 7. Let $P_{\mathcal{W}}S\mathbf{C}^{-1}S^*P_{\mathcal{W}} = U\Sigma U^*$ where $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_m)$ with $\sigma_i > 0$. Then $\hat{\mathbf{x}} = W\hat{\mathbf{G}}\mathbf{y}$ with

$$\hat{\mathbf{G}} = (W^*W)^{-1}W^*UDU^*W(S^*W)^{-1} \quad (61)$$

is admissible on \mathcal{W} where $\mathbf{D} = \text{diag}(d_1, \dots, d_m)$ with $0 \leq d_i \leq 1$.

An immediate consequence of the proposition is that the consistent reconstruction is admissible on \mathcal{W} , since it corresponds to $\hat{\mathbf{G}}$ of (61) with $\mathbf{D} = \mathbf{I}$.

VI. FIR RECONSTRUCTION

An interesting class of reconstruction methods is based on processing the samples by a digital correction filter. In this case

the noisy samples are convolved with an FIR filter of length $2M + 1$ with coefficients $\{g_i, -M \leq i \leq M\}$ producing the cleaned samples $d_i = y_i * g_i$. The corresponding matrix \mathbf{G} in Fig. 1 is a band Toeplitz convolution matrix with i th diagonal equal to g_i . For example, if $m = 5$ and $M = 1$, then

$$\mathbf{G} = \begin{bmatrix} g_0 & g_1 & 0 & 0 & 0 \\ g_{-1} & g_0 & g_1 & 0 & 0 \\ 0 & g_{-1} & g_0 & g_1 & 0 \\ 0 & 0 & g_{-1} & g_0 & g_1 \\ 0 & 0 & 0 & g_{-1} & g_0 \end{bmatrix}. \quad (62)$$

Reconstruction is then obtained as $\hat{x}(t) = \sum_{i=-N}^N d_i w_i(t)$ with $m = 2N + 1$. For simplicity, throughout this section we assume that $S^*W = I$ and the noise is white so that $\mathbf{C} = \sigma^2 \mathbf{I}$ for some $\sigma^2 > 0$. Under this model, our goal is to characterize all admissible FIR filters on \mathcal{W} .

From Theorem 7, the filter g_i is admissible if and only if the corresponding matrix \mathbf{G} satisfies

$$\mathbf{G} = \mathbf{G}^*; \quad (63)$$

$$0 \preceq \mathbf{G} \preceq \mathbf{I}. \quad (64)$$

Condition (63) implies that the filter must be symmetric, *i.e.*, $g_i = g_{-i}$, rendering \mathbf{G} to be a symmetric band Toeplitz matrix. Note that such a matrix is completely determined by its first row. To address the condition (64) we rely on the following proposition.

Lemma 1 [31]: Let \mathbf{A} be an $m \times m$ symmetric Toeplitz matrix with first row equal to $\{\tilde{a}_i, 0 \leq i \leq m-1\}$. Then $\mathbf{A} \succeq 0$ if and only if $A(\omega) \geq 0$ for all $\omega \in [0, 2\pi]$, where

$$A(\omega) = \sum_{n=-(m-1)}^{m-1} a_n e^{-j\omega n} \quad (65)$$

is the discrete-time Fourier transform (DTFT) of the sequence³

$$a_i = \begin{cases} \tilde{a}_i, & 0 \leq i \leq m-1; \\ \tilde{a}_{-i}, & -(m-1) \leq i \leq -1. \end{cases}$$

From Lemma 1 it follows that $\mathbf{G} \succeq 0$ if and only if $G(\omega) \geq 0$ where $G(\omega)$ is the DTFT of $\{g_i, -M \leq i \leq M\}$. Next we note that \mathbf{I} is a symmetric band Toeplitz matrix corresponding to the impulse response $h_i = \delta_i$. Therefore, $\mathbf{G} \preceq \mathbf{I}$, or equivalently, $\mathbf{I} - \mathbf{G} \succeq 0$ if and only if the DTFT $B(\omega)$ of the sequence $b_i = h_i - g_i$ is nonnegative. Since $B(\omega) = 1 - G(\omega)$, we conclude that (64) is equivalent to

$$0 \leq G(\omega) \leq 1, \quad \forall \omega \in [0, 2\pi]. \quad (66)$$

A sufficient condition for $G(\omega) \leq 1$ is $\sum_{i=-M}^M |g_i| \leq 1$. This follows from the fact that

$$G(\omega) \leq |G(\omega)| = \left| \sum_{n=-M}^M g_n e^{-j\omega n} \right| \leq \sum_{n=-M}^M |g_n|. \quad (67)$$

The admissibility conditions on FIR reconstruction are summarized in the following theorem.

³Note that since a_i is symmetric, $A(\omega)$ is real.

Theorem 8: Consider the problem of Theorem 5. Let $\hat{x}_0 = W\mathbf{d}$ be a linear reconstruction of x , where $d_i = g_i * y_i$, and suppose that $S^*W = I$ and $\mathbf{C} = \sigma^2 \mathbf{I}$ for some $\sigma^2 > 0$. Then \hat{x}_0 is admissible if and only if g_i is symmetric, and $0 \leq G(\omega) \leq 1$ where $G(\omega)$ is the DTFT of g_i .

A. Moving Average Reconstruction

In [14] the authors consider the problem of reconstructing a bandlimited signal from a finite set $m = 2N + 1$ of noisy measurements (see also [16] and reference therein). Their problem can be cast in our general framework by choosing

$$s_i(t) = w_i(t) = \text{sinc} \left(\frac{\pi(t - iT)}{T} \right), \quad -N \leq i \leq N. \quad (68)$$

To reconstruct the signal from the samples they proposed filtering the samples with a symmetric FIR filter with impulse response g_i satisfying $\sum_{i=-M}^M |g_i| \leq 1$. In particular, they considered in detail the case in which $g_i = 1/(2M + 1)$.

Using Theorem 8 we now show that their reconstruction method is inadmissible. We then propose an alternative reconstruction that strictly dominates their approach.

The choice of sampling and reconstruction vectors (68) results in $S^*W = W^*W = I$. It then follows from Theorem 8 that the reconstruction is admissible if and only if the DTFT of the filter, $G(\omega)$, satisfies $G(\omega) \geq 0$. Clearly this does not hold for $g_i = 1/(2M + 1)$.

Since this filter is inadmissible, there exists a transformation \mathbf{G} that strictly dominates it for all $x \in \mathcal{W}$. To find a strictly dominating \mathbf{G} we now use Theorem 6 where \mathbf{G}_0 is the matrix representing convolution with the filter $g_i = 1/(2M + 1)$, $-M \leq i \leq M$. With $w_i(t)$ and $s_i(t)$ given by (68),

$$P_{\mathcal{W}} S \mathbf{C}^{-1} S^* P_{\mathcal{W}} = \frac{1}{\sigma^2} W W^* = \frac{1}{\sigma^2} P_{\mathcal{W}}, \quad (69)$$

and

$$R_0 = W \mathbf{G}_0 S^* P_{\mathcal{W}} = W \mathbf{G}_0 W^*, \quad (70)$$

which clearly commutes with $(1/\sigma^2)P_{\mathcal{W}}$. Now, let \mathbf{G}_0 have an eigendecomposition $\mathbf{G}_0 = \mathbf{V} \Delta \mathbf{V}^*$ where \mathbf{V} is a unitary matrix and $\Delta = \text{diag}(\delta_1, \dots, \delta_m)$, and define $U = W \mathbf{V}$. Then $U^*U = I$ and $\mathcal{R}(U) = \mathcal{W}$. Therefore,

$$P_{\mathcal{W}} S \mathbf{C}^{-1} S^* P_{\mathcal{W}} = \frac{1}{\sigma^2} U U^* = U \Sigma U^*, \quad (71)$$

with $\Sigma = (1/\sigma^2)\mathbf{I}$, and

$$R_0 = U \Delta U^*. \quad (72)$$

From Theorem 6, it then follows that a matrix strictly dominating \mathbf{G}_0 is given by

$$\mathbf{G} = \mathbf{V} \mathbf{D} \mathbf{V}^*, \quad (73)$$

where $\mathbf{D} = \text{diag}(d_1, \dots, d_m)$ with

$$d_i = \begin{cases} 1 - |1 - \delta_i|, & |1 - \delta_i| < 1; \\ 0, & |1 - \delta_i| \geq 1. \end{cases} \quad (74)$$

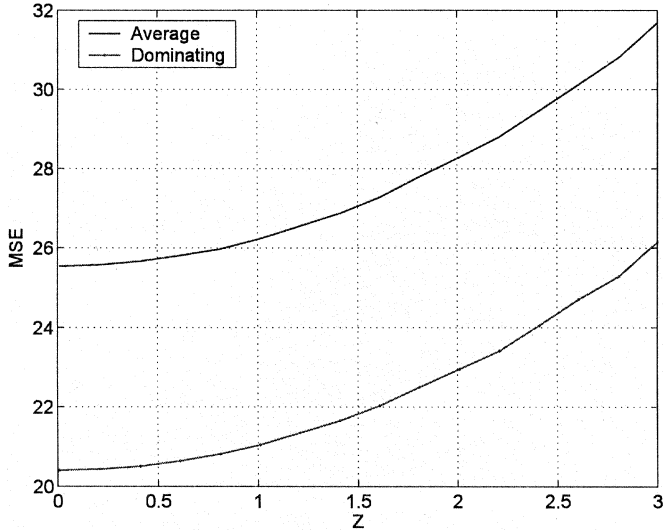


Fig. 4. MSE in estimating $x(t)$ as a function of Z using the moving average method and the estimator of (73) for $x(t)$ given by (76).

Since $\delta_i \leq 1$,

$$d_i = \begin{cases} \delta_i, & 0 \leq \delta_i \leq 1; \\ 0, & \delta_i \leq 0. \end{cases} \quad (75)$$

Thus, the matrix \mathbf{G} is constructed from the matrix \mathbf{G}_0 by replacing each negative eigenvalue by 0.

To illustrate the performance advantage of \mathbf{G}_0 , suppose that we are given $m = 2N + 1$ noisy samples of the bandlimited signal

$$x(t) = 4\pi\sqrt{2\pi} \frac{\cos^2\left(\frac{t}{2}\right)}{\pi^2 - t^2}, \quad (76)$$

at times $t = iT$, $-N \leq i \leq N$, where T is the sampling period. The noise is independent and uniform on $[-Z/2, Z/2]$. The sampling and reconstruction functions are given by (68).

In Fig. 4, we plot the MSE as a function of Z with $m = 75$, $M = 5$, and $T = 1.5$ using the moving average estimator of [14] and the estimator of (73). As can be seen from the figure, the estimator of (73) dominates the moving average estimator for all values of Z . In Fig. 5, we repeat the simulations for a random, bandlimited input signal.

Although the reconstruction (73) dominates the moving average estimator, it no longer corresponds to digital filtering. An important problem therefore that we are currently investigating is to design an FIR filter that leads to a reconstruction strictly dominating a given inadmissible FIR-based method.

VII. CONCLUSION

In this paper, we treated the problem of reconstructing a signal x in a finite-dimensional subspace \mathcal{W} of an arbitrary Hilbert space \mathcal{H} from its nonideal, noisy samples. The samples are modeled as the inner products of x with a set of reconstruction vectors, which span a sampling space \mathcal{S} , and recovery is

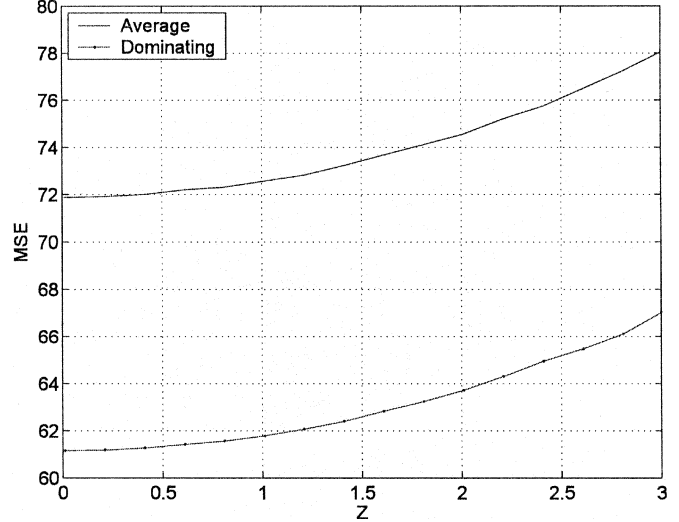


Fig. 5. MSE in estimating $x(t)$ as a function of Z using the moving average method and the estimator of (73) for a random choice of $x(t)$.

performed in the reconstruction space \mathcal{W} , where in general \mathcal{W} and \mathcal{S} can be different but satisfy the condition (3).

To reconstruct the signal from the samples, we suggested formulating the reconstruction problem within a statistical framework and characterizing all reconstructions that are MSE admissible. We also developed explicit methods for dominating an inadmissible reconstruction and proposed several specific admissible strategies, by adapting known estimation approaches to our context. In particular, we considered solutions based on digital filtering and showed that admissibility can be determined by examining the frequency response of the filter.

One of the main contributions of this paper is introducing a framework for systematic design of reconstruction strategies. There are still various aspects that need to be further studied. However, it is our hope that this viewpoint of the reconstruction problem will serve as a catalyst to further investigations of reconstruction from noisy samples.

APPENDIX A PROOF OF THEOREM 2

To prove the theorem, we begin by considering the maximization problem in (22):

$$\max_{\|x\| \leq L, x \in \mathcal{W}} \langle x, Zx \rangle. \quad (77)$$

Let W have a singular value decomposition (SVD)

$$W = U\Sigma V^*, \quad (78)$$

where $U: \mathcal{C}^m \rightarrow \mathcal{H}$ is an orthonormal set transformation satisfying $U^*U = I$, $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_m)$ with $\sigma_i > 0$, and V is an $m \times m$ unitary matrix. Since $x \in \mathcal{W}$, it follows that $x = U\mathbf{d}$ for some vector $\mathbf{d} \in \mathcal{C}^m$. Furthermore, $\|x\| = \|\mathbf{d}\|$ because $U^*U = I$. Therefore,

$$\max_{\|x\| \leq L, x \in \mathcal{W}} \langle x, Zx \rangle = \max_{\|\mathbf{d}\| \leq L} \mathbf{d}^* ((\mathbf{I} - \mathbf{M}\mathbf{H})^* (\mathbf{I} - \mathbf{M}\mathbf{H}) - \mathbf{A}) \mathbf{d}, \quad (79)$$

where we defined $\mathbf{M} = U^*W\mathbf{G}$, $\mathbf{H} = S^*U$, and

$$\mathbf{A} = U^*(I - W\mathbf{G}_0S^*)^*(I - W\mathbf{G}_0S^*)U. \quad (80)$$

Note that since $\mathcal{R}(U) = \mathcal{W}$ it follows from Proposition 1 that \mathbf{H} is invertible. Next, using the fact that

$$P_{\mathcal{W}} = UU^*, \quad (81)$$

we have

$$\begin{aligned} \text{Tr}(W^*W\mathbf{G}\mathbf{C}\mathbf{G}^*) &= \text{Tr}(U^*W\mathbf{G}\mathbf{C}\mathbf{G}^*W^*U) \\ &= \text{Tr}(\mathbf{M}\mathbf{C}\mathbf{M}^*). \end{aligned} \quad (82)$$

Therefore, (22) is equivalent to

$$\min_{\mathbf{M}} \left\{ \text{Tr}(\mathbf{M}\mathbf{C}\mathbf{M}^*) + \max_{\|\mathbf{d}\| \leq L} \mathbf{d}^* ((\mathbf{I} - \mathbf{M}\mathbf{H})^*(\mathbf{I} - \mathbf{M}\mathbf{H}) - \mathbf{A}) \mathbf{d} \right\}. \quad (83)$$

The problem (83) was considered in [20] in which it was shown that \mathbf{M} is optimal if and only if $\mathbf{M} = \mathbf{B}\mathbf{Q}^{-1}\mathbf{H}^*\mathbf{C}^{-1}$ where $\mathbf{Q} = \mathbf{H}^*\mathbf{C}^{-1}\mathbf{H}$ and \mathbf{B} satisfies the conditions below for some $\lambda \geq 0$:

- 1) $\mathbf{Q}\mathbf{B} = \mathbf{B}^*\mathbf{Q}$;
- 2) $0 \preceq \mathbf{Q}\mathbf{B} \prec \mathbf{Q}$;
- 3) $(\mathbf{I} - \mathbf{B})^*(\mathbf{I} - \mathbf{B}) \preceq \lambda\mathbf{I} + \mathbf{A}$;
- 4) $(\mathbf{I} - \mathbf{B})^*(\mathbf{I} - \mathbf{B})\mathbf{B} = (\lambda\mathbf{I} + \mathbf{A})\mathbf{B}$;
- 5) $\text{Tr}(\mathbf{B}(\mathbf{I} - \mathbf{B})^{-1}\mathbf{Q}^{-1}) \leq L^2$;
- 6) $\lambda \left(\text{Tr}(\mathbf{B}(\mathbf{I} - \mathbf{B})^{-1}\mathbf{Q}^{-1}) - L^2 \right) = 0$.

In our case, \mathbf{H} is invertible, so that $\mathbf{Q}^{-1}\mathbf{H}^*\mathbf{C}^{-1} = \mathbf{H}^{-1}$. Therefore, the optimal \mathbf{M} has the form $\mathbf{M} = \mathbf{B}\mathbf{H}^{-1}$, where

$$\mathbf{B} = \mathbf{M}\mathbf{H} = U^*W\mathbf{G}S^*U. \quad (84)$$

Using (84) and

$$\mathbf{Q} = U^*S\mathbf{C}^{-1}S^*U, \quad (85)$$

the first condition becomes

$$\mathbf{Q}\mathbf{B} = U^*S\mathbf{X}S^*U = U^*S\mathbf{X}^*S^*U, \quad (86)$$

where $\mathbf{X} = \mathbf{C}^{-1}S^*W\mathbf{G}$. Since S^*U is invertible, (86) is equivalent to $\mathbf{X} = \mathbf{X}^*$, which proves the first condition.

We now consider the second condition, which using (86) can be written as

$$0 \preceq U^*S\mathbf{X}S^*U \prec U^*S\mathbf{C}^{-1}S^*U. \quad (87)$$

Multiplying on the right by $(S^*U)^{-1}$ and on the left by $(U^*S)^{-1}$, (87) is equivalent to $0 \preceq \mathbf{X} \prec \mathbf{C}^{-1}$, proving the second condition.

To establish the third condition, we note that from (84),

$$\mathbf{I} - \mathbf{B} = U^*(I - W\mathbf{G}S^*)U = U^*(P_{\mathcal{W}} - R)U, \quad (88)$$

and

$$(\mathbf{I} - \mathbf{B})^*(\mathbf{I} - \mathbf{B}) = U^*(P_{\mathcal{W}} - R)^*(P_{\mathcal{W}} - R)U. \quad (89)$$

Therefore, the third condition becomes

$$U^*(I - R)^*(I - R)U \preceq \lambda\mathbf{I} + \mathbf{A}. \quad (90)$$

Noting that for any two matrices \mathbf{T} and \mathbf{Y} we have that $\mathbf{T} \preceq \mathbf{Y}$ if and only if $U\mathbf{T}U^* \preceq U\mathbf{Y}U^*$, (90) becomes

$$(P_{\mathcal{W}} - R)^*(P_{\mathcal{W}} - R) \preceq \lambda P_{\mathcal{W}} + U\mathbf{A}U^*. \quad (91)$$

The third condition then follows from the fact that

$$U\mathbf{A}U^* = (P_{\mathcal{W}} - R_0)^*(P_{\mathcal{W}} - R_0). \quad (92)$$

The fourth condition is proven in a similar way.

To prove the fifth and sixth conditions, we need to find an expression for $\mathbf{B}(\mathbf{I} - \mathbf{B})^{-1}\mathbf{Q}^{-1}$. Using (88) and (86),

$$\begin{aligned} \mathbf{B}(\mathbf{I} - \mathbf{B})^{-1}\mathbf{Q}^{-1} &= \mathbf{B}(\mathbf{Q} - \mathbf{Q}\mathbf{B})^{-1} \\ &= U^*W\mathbf{G}(\mathbf{C}^{-1} - \mathbf{C}^{-1}S^*W\mathbf{G})^{-1}(U^*S)^{-1}, \end{aligned} \quad (93)$$

where we used the fact that from condition 2, $\mathbf{C}^{-1} - \mathbf{C}^{-1}S^*W\mathbf{G}$ is invertible. Therefore,

$$\begin{aligned} \text{Tr}(\mathbf{B}(\mathbf{I} - \mathbf{B})^{-1}\mathbf{Q}^{-1}) &= \\ &= \text{Tr}(W\mathbf{G}(\mathbf{C}^{-1} - \mathbf{C}^{-1}S^*W\mathbf{G})^{-1}(U^*S)^{-1}U^*). \end{aligned} \quad (94)$$

Noting that

$$\begin{aligned} (U^*S)^{-1}U^* &= (\mathbf{V}\Sigma U^*S)^{-1}\mathbf{V}\Sigma U^* \\ &= (W^*S)^{-1}W^*, \end{aligned} \quad (95)$$

equation (94) becomes

$$\begin{aligned} \text{Tr}(\mathbf{B}(\mathbf{I} - \mathbf{B})^{-1}\mathbf{Q}^{-1}) &= \\ &= \text{Tr}(W^*W\mathbf{G}(\mathbf{C} - \mathbf{C}S^*W\mathbf{G})^{-1}(W^*S)^{-1}), \end{aligned} \quad (96)$$

which completes the proof.

APPENDIX B PROOF OF THEOREM 3

To prove Theorem 3, we need to show that $\widehat{\mathbf{G}}$ with d_i given by (26) satisfies the optimality conditions of Theorem 2.

We first note that $\mathcal{R}(P_{\mathcal{W}}S\mathbf{C}^{-1}S^*P_{\mathcal{W}}) = \mathcal{R}(P_{\mathcal{W}}S\mathbf{C}^{-1}) = \mathcal{R}(P_{\mathcal{W}})$. Therefore,

$$P_{\mathcal{W}} = UU^*. \quad (97)$$

Using (97) together with the fact that $W^* = W^*P_{\mathcal{W}}$, we have that

$$\begin{aligned} \mathbf{C}^{-1}S^*W\widehat{\mathbf{G}} &= (W^*S)^{-1}W^*P_{\mathcal{W}}S\mathbf{C}^{-1}S^*W\widehat{\mathbf{G}} \\ &= (W^*S)^{-1}W^*U\Sigma DU^*W(S^*W)^{-1}. \end{aligned} \quad (98)$$

From (98) the first condition is equivalent to $d_i = d_i^*$. Since $\eta_i \geq 0$, d_i is real and the first condition is satisfied.

Since W^*U is invertible, the second condition becomes

$$0 \preceq \Sigma \mathbf{D} \prec (W^*U)^{-1}W^*S\mathbf{C}^{-1}S^*W(U^*W)^{-1}, \quad (99)$$

or,

$$0 \preceq \Sigma \mathbf{D} \prec \Sigma. \quad (100)$$

Using the fact that $\Sigma \succ 0$, (100) reduces to $0 \leq d_i < 1$. By definition of d_i , we have that $d_i \geq 0$. To show that $d_i < 1$ we need to show that $\eta_i > 0$. Since $\lambda \geq 0$, $\eta_i \geq 0$. Furthermore, $\eta_i = 0$ if and only if $\delta_i = 0$ and $\lambda = 0$. But, if $\eta_i = 0$, then $\mathcal{T}(0) \rightarrow \infty$ so that λ cannot be equal to 0. Therefore, $\eta_i > 0$ and $d_i < 1$.

We now consider the third condition. With our definition of $\widehat{\mathbf{G}}$,

$$W\widehat{\mathbf{G}}S^* = P_{\mathcal{W}}UDU^*W(S^*W)^{-1}S^* = UDU^*E_{\mathcal{W}S^\perp}. \quad (101)$$

Since $E_{\mathcal{W}S^\perp}P_{\mathcal{W}} = P_{\mathcal{W}}$,

$$(I - W\widehat{\mathbf{G}}S^*)P_{\mathcal{W}} = U(I - \mathbf{D})U^*, \quad (102)$$

and

$$(I - W\widehat{\mathbf{G}}_0S^*)P_{\mathcal{W}} = U(I - \Delta)U^*. \quad (103)$$

Therefore, conditions 3 and 4 become $(1 - d_i)^2 \leq \lambda + |1 - \delta_i|^2$ and $d_i(1 - d_i)^2 = d_i(\lambda + |1 - \delta_i|^2)$, which can be written as

$$(1 - d_i)^2 = \lambda + |1 - \delta_i|^2, \quad i : d_i \neq 0, \quad (104)$$

and

$$1 \leq \lambda + |1 - \delta_i|^2, \quad i : d_i = 0. \quad (105)$$

Since $\eta_i = \lambda + |1 - \delta_i|^2$, both conditions are satisfied.

To address conditions 5 and 6 we now find an expression for $\text{Tr}(W^*W\widehat{\mathbf{G}}(\mathbf{C}^{-1} - \mathbf{C}^{-1}S^*W\widehat{\mathbf{G}})^{-1}(W^*S)^{-1})$. To this end we note that

$$\begin{aligned} & (\mathbf{C}^{-1} - \mathbf{C}^{-1}S^*W\widehat{\mathbf{G}})^{-1}(W^*S)^{-1} \\ &= (W^*S\mathbf{C}^{-1} - W^*U\Sigma DU^*W(S^*W)^{-1})^{-1} \\ &= (W^*U\Sigma(\mathbf{I} - \mathbf{D})U^*W(S^*W)^{-1})^{-1} \\ &= S^*W(U^*W)^{-1}(\mathbf{I} - \mathbf{D})^{-1}\Sigma^{-1}(W^*U)^{-1}. \end{aligned} \quad (106)$$

Thus,

$$\begin{aligned} & \text{Tr}(W^*W\widehat{\mathbf{G}}(\mathbf{C}^{-1} - \mathbf{C}^{-1}S^*W\widehat{\mathbf{G}})^{-1}(W^*S)^{-1}) - L^2 \\ &= \text{Tr}((\mathbf{I} - \mathbf{D})^{-1}\Sigma^{-1}) - L^2 \\ &= \sum_{i=1}^m \frac{d_i}{(1 - d_i)\sigma_i} - L^2 \\ &= \mathcal{T}(\lambda). \end{aligned} \quad (107)$$

Finally, the last two conditions are satisfied by our definition of $\mathcal{T}(\lambda)$. It remains to show that if $\mathcal{T}(0) > 0$, then there is a unique value of $0 < \lambda < \alpha$ such that $\mathcal{T}(\lambda) = 0$.

Since η_i is monotonically increasing in $\lambda > 0$, and each term in the sum in the definition of $\mathcal{T}(\lambda)$ is positive, $\mathcal{T}(\lambda)$ is monotonically decreasing in λ as long as there exists at least on i for which $\eta_i < 1$, or equivalently, as long as $\lambda < \alpha$. In addition, $\mathcal{T}(\lambda)$ is continuous. Now, we are assuming that $\mathcal{T}(0) > 0$. Furthermore, $\mathcal{T}(\lambda) = -L^2$ for $\lambda \geq \alpha$. Therefore, there is a unique $0 < \lambda < \alpha$ such that $\mathcal{T}(\lambda) = 0$.

APPENDIX C PROOF OF THEOREM 5

We consider the problem

$$\min_{\mathbf{G}} \left\{ \text{Tr}(W^*W\mathbf{G}\mathbf{C}\mathbf{G}^*) + \sup_{x \in \mathcal{W}} \langle x, Zx \rangle \right\} \quad (108)$$

where Z is defined in (23). Using the SVD (78) of W ,

$$\sup_{x \in \mathcal{W}} \langle x, Zx \rangle = \sup_{\mathbf{d}} \mathbf{d}^* ((\mathbf{I} - \mathbf{M}\mathbf{H})^*(\mathbf{I} - \mathbf{M}\mathbf{H}) - \mathbf{A}) \mathbf{d}, \quad (109)$$

where $\mathbf{M} = U^*W\mathbf{G}$, $\mathbf{H} = S^*U$, and \mathbf{A} is defined in (80). Now, for any matrix \mathbf{Z} ,

$$\sup_{\mathbf{x}} \langle \mathbf{x}, \mathbf{Z}\mathbf{x} \rangle = \begin{cases} 0, & \mathbf{Z} \preceq 0; \\ \infty, & \text{otherwise.} \end{cases} \quad (110)$$

Therefore, (108) is equivalent to

$$\min_{\mathbf{G}} \text{Tr}(\mathbf{M}^*\mathbf{C}\mathbf{M}) \quad (111)$$

subject to

$$(\mathbf{I} - \mathbf{M}\mathbf{H})^*(\mathbf{I} - \mathbf{M}\mathbf{H}) \preceq \mathbf{A}. \quad (112)$$

The problem of minimizing (111) subject to (112) was considered in [20], in which it was shown that \mathbf{M} is optimal if and only if $\mathbf{M} = \mathbf{B}\mathbf{Q}^{-1}\mathbf{H}^*\mathbf{C}^{-1}$ where $\mathbf{Q} = \mathbf{H}^*\mathbf{C}^{-1}\mathbf{H}$ and \mathbf{B} satisfies the following conditions:

- 1) $\mathbf{Q}\mathbf{B} = \mathbf{B}^*\mathbf{Q}$;
- 2) $0 \preceq \mathbf{Q}\mathbf{B} \preceq \mathbf{Q}$;
- 3) $(\mathbf{I} - \mathbf{B})\mathbf{P}_{\mathcal{R}(\mathbf{A})} = \mathbf{I} - \mathbf{B}$;
- 4) $(\mathbf{I} - \mathbf{B})^*(\mathbf{I} - \mathbf{B}) \preceq \mathbf{A}$;
- 5) $(\mathbf{I} - \mathbf{B})^*(\mathbf{I} - \mathbf{B})\mathbf{X} = \mathbf{A}\mathbf{X}$.

Here, $\mathbf{X} = (\mathbf{Q}(\mathbf{I} - \mathbf{B}))^\dagger - \mathbf{P}_{\mathcal{R}(\mathbf{A})}\mathbf{Q}^{-1}\mathbf{P}_{\mathcal{R}(\mathbf{A})}$ and $\mathbf{P}_{\mathcal{R}(\mathbf{A})}$ denotes the orthogonal projection onto $\mathcal{R}(\mathbf{A})$. As we now show, adapting these conditions to our problem results in the conditions of Theorem 5.

The proof of the first two conditions follows as in the proof of Theorem 2. Condition 4 follows in a similar way as condition 3 in Theorem 2 with $\lambda = 0$.

To prove the third condition, we note that from (84),

$$\mathbf{I} - \mathbf{B} = U^*(I - W\mathbf{G}S^*)U = U^*(P_{\mathcal{W}} - R)U. \quad (113)$$

In addition, since $\mathbf{A} = U^*(P_{\mathcal{W}} - R_0)^*(P_{\mathcal{W}} - R_0)U$,

$$\mathbf{P}_{\mathcal{R}(\mathbf{A})} = U^*P_{\mathcal{R}}U. \quad (114)$$

Therefore, the third condition becomes

$$U^*(P_{\mathcal{W}} - R)P_{\mathcal{R}}U = U^*(P_{\mathcal{W}} - R)U. \quad (115)$$

Since $\mathcal{N}(U) = \{0\}$, $\mathcal{N}(U^*) = \mathcal{W}^\perp$ and $\mathcal{R}(P_{\mathcal{W}} - R) \subseteq \mathcal{W}$, (115) is equivalent to $(P_{\mathcal{W}} - R)P_{\mathcal{R}} = P_{\mathcal{W}} - R$, proving the third condition.

We now prove the 5th condition. To this end we note that

$$\mathbf{X} = (U^*SC^{-1}S^*(P_{\mathcal{W}} - R)U)^\dagger - U^*P_{\mathcal{R}}U(U^*SC^{-1}S^*U)^{-1}U^*P_{\mathcal{R}}U. \quad (116)$$

Using the fact that $U^*S = U^*P_{\mathcal{W}}S$,

$$\begin{aligned} & (U^*SC^{-1}S^*(P_{\mathcal{W}} - R)U)^\dagger \\ &= (U^*P_{\mathcal{W}}SC^{-1}S^*(P_{\mathcal{W}} - R)U)^\dagger \\ &= U^*(P_{\mathcal{W}}SC^{-1}S^*(P_{\mathcal{W}} - R))^\dagger U. \end{aligned} \quad (117)$$

The last equality follows from the fact that $(U^*AU)^\dagger = U^*A^\dagger U$ for any A with $\mathcal{R}(A) \subseteq \mathcal{W}$ and $\mathcal{N}(A) \supseteq \mathcal{W}^\perp$. Combining (116), (117) and $U(U^*SC^{-1}S^*U)^{-1}U^* = W(W^*SC^{-1}S^*W)^{-1}W^*$, we have that $\mathbf{X} = U^*XU$. Thus,

$$(\mathbf{I} - \mathbf{B})^*(\mathbf{I} - \mathbf{B})\mathbf{X} = U^*(P_{\mathcal{W}} - R)^*(P_{\mathcal{W}} - R)XU, \quad (118)$$

and

$$\mathbf{A}\mathbf{X} = U^*(P_{\mathcal{W}} - R_0)^*(P_{\mathcal{W}} - R_0)XU, \quad (119)$$

which proves the fifth condition.

APPENDIX D PROOF OF THEOREM 6

To prove Theorem 6, we need to show that $\hat{\mathbf{G}}$ of (59) with d_i given by (60) satisfies the optimality conditions of Theorem 5. Following similar steps as in the proof of Theorem 3, the first two conditions become $d_i = d_i^*$ and $0 \leq d_i \leq 1$ and the fourth condition becomes $(1 - d_i)^2 \leq |1 - \delta_i|^2$, which are all satisfied.

We now consider the third condition. Since

$$P_{\mathcal{W}} - R_0 = U(\mathbf{I} - \Delta)U^*, \quad (120)$$

we have that $P_{\mathcal{R}} = U\tilde{\mathbf{I}}U^*$, where $\tilde{\mathbf{I}} = \text{diag}(e_1, \dots, e_m)$ and

$$e_i = \begin{cases} 1, & \delta_i \neq 1; \\ 0, & \delta_i = 1. \end{cases} \quad (121)$$

Noting that $R = UDU^*$, the third condition becomes

$$U(\mathbf{I} - \mathbf{D})\tilde{\mathbf{I}}U^* = U(\mathbf{I} - \mathbf{D})U^*. \quad (122)$$

Since from our definition of d_i , $d_i = 1$ for all i such that $\delta_i = 1$, (122) is satisfied.

To prove that the fifth condition is also satisfied, we note that under the assumptions of the theorem,

$$X = U(\Sigma(\mathbf{I} - \mathbf{D}))^\dagger U^* - U\tilde{\mathbf{I}}\Sigma^{-1}\tilde{\mathbf{I}}U^*$$

$$= U\Sigma^{-1} \left((\mathbf{I} - \mathbf{D})^\dagger - \tilde{\mathbf{I}} \right) U^*. \quad (123)$$

With X given by (123), the fifth condition becomes

$$\begin{aligned} U(\mathbf{I} - \mathbf{D})^*(\mathbf{I} - \mathbf{D})\Sigma^{-1} \left((\mathbf{I} - \mathbf{D})^\dagger - \tilde{\mathbf{I}} \right) U^* &= \\ = U(\mathbf{I} - \Delta)^*(\mathbf{I} - \Delta)\Sigma^{-1} \left((\mathbf{I} - \mathbf{D})^\dagger - \tilde{\mathbf{I}} \right) U^*, \end{aligned} \quad (124)$$

which is equivalent to

$$|1 - d_i|^2 = |1 - \delta_i|^2, \quad \forall i: (1 - d_i)^\dagger - e_i \neq 0. \quad (125)$$

Now, from our definition of d_i we have that $|1 - d_i| \neq |1 - \delta_i|$ only if $\delta_i < 0$. In this case $d_i = 0$ and $e_i = 1$, so that $(1 - d_i)^\dagger = e_i$. Therefore, (125) is also satisfied, completing the proof.

REFERENCES

- [1] E. Margolis and Y. C. Eldar, "Filterbank reconstruction of periodic signals and sampling in polar coordinates," presented at the 2003 Workshop Sampling Theory Applications (SampTA'OS), May 2003.
- [2] —, "Reconstruction of nonuniformly sampled periodic signals: Algorithms and stability analysis," in *Proc. 11th IEEE Int. Conf. Electronics, Circuits, Systems (ICECS)*, Dec. 2004, pp. 555–558.
- [3] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 2915–2925, Nov. 1994.
- [4] Y. C. Eldar, "Sampling and reconstruction in arbitrary spaces and oblique dual frame vectors," *J. Fourier Anal. Appl.*, vol. 1, no. 9, pp. 77–96, Jan. 2003.
- [5] —, "Sampling without input constraints: Consistent reconstruction in arbitrary spaces," in *Sampling, Wavelets and Tomography*, A. I. Zayed and J. J. Benedetto, Eds. Boston, MA: Birkhauser, 2004, pp. 33–60.
- [6] Y. C. Eldar and T. Werther, "General framework for consistent sampling in Hilbert spaces," *Int. J. Wavelets, Multiresolution, Inf. Process.*, vol. 3, pp. 347–359, Sep. 2005.
- [7] Y. C. Eldar and T. Dvorkind, "A minimum squared-error framework for generalized sampling," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2155–2167, Jun. 2006.
- [8] W. Chen, B. Han, and R. Q. Jia, "On simple oversampled A/D conversion in shift invariant spaces," *IEEE Trans. Inf. Theory*, vol. 51, pp. 648–657, Feb. 2005.
- [9] K. F. Cheung and R. J. Marks, II, "Ill-posed sampling theorems," *IEEE Trans. Circuits Syst.*, vol. CAS-32, pp. 481–484, May 1985.
- [10] D. S. Chen and J. P. Allebach, "Analysis of error in reconstruction of two-dimensional signals from irregularly spaced samples," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, pp. 173–180, 1987.
- [11] G. Wan and W. Han, "Minimum error bound of signal reconstruction," *IEEE Signal Process. Lett.*, vol. 6, pp. 309–311, Dec. 1999.
- [12] H. Johansson and P. Lowenborg, "Reconstruction of nonuniformly sampled bandlimited signals by means of digital fractional delay filters," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2757–2767, Nov. 2002.
- [13] E. J. Diethorn and D. C. Munson Jr., "A linear, time-varying system framework for noniterative discrete-time band-limited signal extrapolation," *IEEE Trans. Signal Process.*, vol. 39, no. 1, pp. 55–68, Jan. 1991.
- [14] A. Krzyzak, E. Rafajłowicz, and M. Pawlak, "Moving average restoration of bandlimited signals from noisy observations," *IEEE Trans. Signal Process.*, vol. 45, no. 12, pp. 2967–2976, Dec. 1997.
- [15] A. Aldroubi and K. Grochenig, "Nonuniform sampling and reconstruction in shift invariant spaces," *SIAM Rev.*, vol. 43, pp. 585–620, 2001.
- [16] M. Pawlak, E. Rafajłowicz, and A. Krzyzak, "Postfiltering versus prefiltering for signal recovery from noisy samples," *IEEE Trans. Inform. Theory*, vol. 49, pp. 3195–3212, Dec. 2003.

- [17] Y. C. Eldar and M. Unser, "Nonideal sampling and interpolation from noisy observations in shift-invariant spaces," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2636–2651, Jul. 2006.
- [18] A. Cohen, "All admissible linear estimates of the normal mean," *Ann. Math. Statist.*, vol. 37, pp. 458–463, 1966.
- [19] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, second ed. New York: Springer-Verlag, 1998.
- [20] Y. C. Eldar, "Comparing between estimation approaches: Admissible and dominating linear estimators," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1689–1702, May 2006.
- [21] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 168–181, Jan. 2005.
- [22] —, "Linear minimax regret estimation of deterministic parameters with bounded data uncertainties," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2177–2188, Aug. 2004.
- [23] A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems*. Washington, DC: V. H. Winston, 1977.
- [24] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
- [25] D. P. Bertsekas, *Nonlinear Programming*, second ed. Belmont, MA: Athena Scientific, 1999.
- [26] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 40–95, Mar. 1996.
- [27] Y. Nesterov and A. Nemirovski, *Interior-Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA: SIAM, 1994.
- [28] F. Alizadeh, "Combinatorial optimization with interior point methods and semi-definite matrices," Ph.D. dissertation, Univ. of Minnesota, Minneapolis, MN, 1991.
- [29] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
- [30] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, pp. 55–67, Feb. 1970.

- [31] L. D. Grossman and Y. C. Eldar, "On positive semidefinite Toeplitz matrices," to be published.



Yonina C. Eldar (S'98–M'02) received the B.Sc. degree in physics and the B.Sc. degree in electrical engineering both from Tel-Aviv University (TAU), Tel-Aviv, Israel, in 1995 and 1996, respectively, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 2001.

From January 2002 to July 2002, she was a Post-doctoral Fellow with the Digital Signal Processing Group, MIT. She is currently an Associate Professor with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa. She is also a Research Affiliate with the Research Laboratory of Electronics at MIT. Her current research interests are in the general areas of signal processing, statistical signal processing, and computational biology.

Dr. Eldar was in the program for outstanding students at TAU from 1992 to 1996. In 1998, she held the Rosenblith Fellowship for study in Electrical Engineering at MIT, and in 2000, she held an IBM Research Fellowship. She is currently a Horev Fellow of the Leaders in Science and Technology program at the Technion and an Alon Fellow. In 2004, she was awarded the Wolf Foundation Krill Prize for Excellence in Scientific Research, and in 2005 the Andre and Bella Meyer Lectureship. She is a member of the IEEE Signal Processing Theory and Methods Technical Committee, an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the *EURASIP Journal on Signal Processing*, and on the Editorial Board of *Foundations and Trends in Signal Processing*.