Filter Design for MIMO Sampling and Reconstruction

Raman Venkataramani and Yoram Bresler*

Coordinated Science Laboratory
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
1308 W. Main Street, Urbana, IL 61801
Raman Venkataramani: Ph. (908) 582-8110 E-Mail raman@research.bell-labs.com
Y. Bresler: Ph. (217) 244 9660 E-Mail ybresler@uiuc.edu
FAX: (217) 244 1642

Abstract

We address the problem of FIR equalizer filter design for multiple-input multiple-output (MIMO) linear time-invariant channels with uniform sampling at the channel outputs. This scheme encompasses Papoulis' generalized sampling and several nonuniform sampling schemes as special cases. The input signals are modeled as either continuous-time or discrete-time multi-band input signals, with different band structures. We derive conditions on the channel and the sampling rate that allow perfect inversion of the channel. Additionally, we derive a stronger set of conditions under which the equalizer filters can be chosen to be continuous functions of the frequency variable. We also provide conditions for the existence of FIR perfect reconstruction filters, and when such do not exist, we address the optimal approximation of the ideal filters using FIR filters and a min-max l_2 end-to-end distortion criterion. The design problem is then reduced to a standard semi-infinite linear program. An example design of FIR equalizer filters is given.

Keywords: multiple-input multiple-output (MIMO) channel, multichannel deconvolution, MIMO equalization, multiple source separation, equalizer filter design, multiband sampling, multirate signal processing, signal reconstruction, min-max criterion, semi-infinite optimization.

EDICS: 2-SAMP, 1-MCHT, 1-IDSS, 2-FILB

* **Corresponding Author: Yoram Bresler**

I. INTRODUCTION

Multiple-input multiple-output (MIMO) deconvolution, or channel equalization, involves the recovery of the inputs to a MIMO channel whose outputs can be observed, and whose characteristics may either be known or unknown. The unknown inputs usually have overlapping spectra and hence share a common bandwidth. MIMO deconvolution is an important problem arising in numerous applications, including multiuser or multiaccess wireless communications and space-time coding with antenna arrays, or telephone digital subscriber loops [1–4], multisensor biomedical signals [5, 6], multi-track magnetic recording [7], multiple speaker (or other acoustic source) separation with microphone arrays [8, 9], geophysical data processing [10], and multichannel image restoration [11, 12].

In practice, digital processing is used to perform the channel inversion. Consequently, the channel outputs need to be sampled prior to processing, and the objective is to reconstruct the channel inputs from the sampled output signals. Therefore we restate the channel inversion problem as a problem in sampling theory, and call it *MIMO sampling*. To focus on the sampling and reconstruction issues, we restrict our attention to the scenario of a linear time-invariant MIMO channel with known frequency response matrix. As appropriate in many applications, the input signals to the channel are assumed to be multi-band signals, with possibly different band structures.

The continuous-time model for the MIMO channel and its reconstruction [13] is illustrated in Figure 1. However, because the processing is done digitally, it is convenient to use an equivalent discrete-time model for the continuous-time channel, with discrete-time sequences representing samples of the continuous-time counterparts at the Nyquist rate or higher. This is shown rigorously in the Appendix, where we arrive at the following models for the MIMO channel and the reconstruction system. Figure 2 depicts the block diagram of the discrete-time MIMO channel with R inputs and P outputs. The inputs to the channel are the sequences $x_r[k]$, $r = 1, \ldots, R$, and the outputs are $y_p[k]$, $p = 1, \ldots, P$. The channel outputs are then uniformly subsampled (downsampled) by an integer factor $L > 0$ to produce sequences $z_p[n]$. The reconstruction block, depicted in Figure 3, produces estimates $\tilde{x}_r[k]$ of the input signals from the quantities $z_p[n]$. The continuous-time inputs can finally be recovered from the discrete-time sequences $\tilde{x}_r[k]$ using a bank of conventional D/A converters.

We shall consider only uniform subsampling of the channel outputs (see [14, 15] for results on arbitrary nonuniform sampling). This scheme is fairly general, and subsumes periodic nonuniform subsampling of the MIMO outputs as a special case of uniform subsampling applied to a hypothetical channel with more rows [13]. Furthermore, several familiar sampling schemes can be viewed as special cases of MIMO sampling. For example, in Papoulis' generalized sampling [16], a single lowpass input signal is passed through a bank of M filters, and the outputs are sampled at $1/M$-th the Nyquist rate. This fits in our framework as a single-input multiple-output sampling problem, *i.e.*, $R = 1$. Additionally, if the channel filters are pure delays, we obtain multicoset or periodic nonuniform sampling of the input signal, which has been widely studied [17–29], as it allows to approach the Landau minimum sampling for multiband signals [30]. Seidner and Feder [31] provide a natural generalization of Papoulis' sampling expansions for a vector input with its components bandlimited to $[-B, B]$. Clearly, their sampling scheme is also a special case of MIMO sampling.

We studied the continuous-time MIMO sampling problem and presented necessary and sufficient conditions for perfect stable reconstruction of the channel inputs from uniform sampling of the outputs in [13]. Importantly, we demonstrated how to achieve stable sampling and reconstruction at rates lower than the Nyquist rate, and in some cases even at combined average rates lower than the Landau rate for individual channel inputs. This provides motivation for using the MIMO sampling theory to design and implement MIMO deconvolution, MIMO channel equalization, and source separation systems.

In this paper, we examine the related problem of finite impulse response (FIR) filter design for MIMO reconstruction filters. Whereas [13] only demonstrates the existence of ideal filters for stable perfect reconstruction subject to appropriate conditions on the channel and sampling rates, in this paper we address the practical problem of implementing the reconstruction system using FIR filters. We provide conditions for the existence of FIR perfect reconstruction filters, and when such do not exist, we address the optimal approximation of the ideal filters using FIR filters and a min-max l_2 reconstruction error criterion. We formulate the design problem as a semi-infinite linear program. Semi-infinite formulations have been successfully applied to other multirate filter design problems [32, 33] and solved using standard techniques [34]. Our FIR filter design formulation is fairly general and can be used to design the interpolation filters for those generalized sampling schemes discussed above.

The paper is organized as follows. Section II, contains some basic notation and definitions. In Section III we present discrete-time models for the channel and reconstruction block. The channel inputs are modeled as multiband signals. In Section IV, we present discrete-time versions of the results derived in [13]. In particular, we specify necessary and sufficient conditions for the existence of reconstruction filters that are continuous in the frequency domain. This property is important in the context of FIR filter design, as we elaborate later. Finally, in Section V we discuss the problem of FIR reconstruction filter design for the MIMO sampling problem. We formulate a cost function in terms of the filter coefficients. Minimizing the cost produces the optimal filter coefficients. The problem may be recast as a semi-infinite linear program. We present two design examples: one for multicoset sampling and another for MIMO sampling with two inputs.

II. PRELIMINARIES

We begin with a some basic definitions and notation. Denote the discrete-time Fourier transform of a $x[n] \in l^2$ by the periodic function:

$$X[\nu] = \sum_{k \in \mathbb{Z}} x[k] e^{-j2\pi\nu k}.$$

In general, we denote time signals (either scalar-valued or vector-valued) using lower-case letters, and their Fourier transforms by the corresponding upper-case letters. Let the class complex-valued, finite-energy discrete-time signals bandlimited to the set of frequencies $\mathcal{F} \subseteq [0, 1)$ by

$$\mathcal{B}_d(\mathcal{F}) = \{x[k] \in l^2 : X[\nu] = 0, \ \forall \nu \in [0, 1) \cap \mathcal{F}^c\}. \tag{1}$$

We denote the class of complex-valued matrices of size $M \times N$ by $\mathbb{C}^{M \times N}$, the conjugate-transpose of $\boldsymbol{A}$ by $\boldsymbol{A}^H$, its pseudo inverse by $\boldsymbol{A}^\dagger$, and its range space by $\mathcal{R}(\boldsymbol{A})$. For a given matrix $\boldsymbol{A}$, let $\boldsymbol{A}_{\mathcal{R},\mathcal{C}}$ denote the submatrix of $\boldsymbol{A}$ corresponding to rows indexed by the set $\mathcal{R}$ and columns by the set $\mathcal{C}$. The quantity $\boldsymbol{A}_{\bullet,\mathcal{C}}$ denotes a submatrix formed by keeping all rows of $\boldsymbol{A}$, but only columns indexed by $\mathcal{C}$, while $\boldsymbol{A}_{\mathcal{R},\bullet}$ denotes the submatrix formed by retaining rows indexed by $\mathcal{R}$ and all columns. We use a similar notation for vectors. Hence $\boldsymbol{X}_{\mathcal{R}}$ is the subvector of $\boldsymbol{X}$ corresponding to rows indexed by $\mathcal{R}$. We always apply the subscripts of a matrix before the superscript. So $\boldsymbol{A}^H_{\mathcal{R},\mathcal{C}}$ is the conjugate-transpose of $\boldsymbol{A}_{\mathcal{R},\mathcal{C}}$. When dealing with singleton index sets: $\mathcal{R} = \{r\}$ or $\mathcal{C} = \{c\}$, we omit the curly braces for readability. Therefore $\boldsymbol{A}_{r,\bullet}$ and $\boldsymbol{A}_{\bullet,c}$ are the r-th row and the c-th column of $\boldsymbol{A}$ respectively. For convenience, we always number the rows and columns of a finite-size matrix starting from 0. For infinite-size matrices, the row and column indices range over $\mathbb{Z}$. As a result of the above notation, we have the following straightforward proposition that is used later.

Proposition 1. *Suppose that $\boldsymbol{X} \in \mathbb{C}^L$ and $\boldsymbol{I}$ is the $L \times L$ identity matrix. Then $\boldsymbol{X}_{\mathcal{K}} = \boldsymbol{I}_{\mathcal{K},\bullet}\boldsymbol{X}$ for all $\mathcal{K} \subseteq \{0, \ldots, L-1\}$. Additionally, if $\boldsymbol{X}_{\mathcal{K}^c} = \boldsymbol{0}$, where $\mathcal{K}^c$ is the complement of $\mathcal{K}$, then $\boldsymbol{X} = \boldsymbol{I}_{\bullet,\mathcal{K}}\boldsymbol{X}_{\mathcal{K}}$.*

The identity matrix of size $N \times N$ is denoted by $\boldsymbol{I}_N$, and the zero matrix by $\boldsymbol{0}$. Finally suppose that $\mathcal{S}$ is a subset of $\mathbb{R}$ or $\mathbb{Z}$, and a is an element of $\mathbb{R}$ or $\mathbb{Z}$, then

$$\mathcal{S} \oplus a = \{s + a : s \in \mathcal{S}\}, \qquad \mathcal{S} \ominus a = \{s - a : s \in \mathcal{S}\},$$
$$a\mathcal{S} = \{as : s \in \mathcal{S}\}, \qquad \mathcal{S} \bmod a = \{s \bmod a : s \in \mathcal{S}\},$$

denote the positive and negative translations, scaling, and the modulus of $\mathcal{S}$ by a respectively.

III. Sampling and Reconstruction Models

Figure 2 depicts a MIMO channel whose inputs and outputs are discrete-time sequences $\{x_1[k], \ldots, x_R[k]\}$ and $\{y_1[k], \ldots, y_P[k]\}$ respectively. For convenience, let $\mathcal{R} = \{0, 1, \ldots, R-1\}$ and $\mathcal{P} = \{0, 1, \ldots, P-1\}$ denote index sets for the channel inputs and outputs. We model $x_r[k]$, $r \in \mathcal{R}$ as multiband signals: $x_r(t) \in \mathcal{B}_d(\mathcal{F}_r)$, where the *spectral support* $\mathcal{F}_r \subseteq [0, 1)$ is a finite union of disjoint intervals:

$$\mathcal{F}_r = \bigcup_{n=1}^{N_r} [a_{rn}, b_{rn}), \qquad a_{r1} < b_{r1} < a_{r2} < \cdots < a_{rN_r} < b_{rN_r}. \tag{2}$$

Let the channel inputs and outputs be expressed in vector form as

$$\boldsymbol{x}[k] = \begin{pmatrix} x_0[k] & x_1[k] & \cdots & x_{R-1}[k] \end{pmatrix}^T,$$
$$\boldsymbol{y}[k] = \begin{pmatrix} y_0[k] & y_1[k] & \cdots & y_{P-1}[k] \end{pmatrix}^T.$$

The MIMO channel is modeled as a linear shift-invariant system, thus enabling us to write

$$\boldsymbol{y}[k] = \boldsymbol{g} * \boldsymbol{x}[k] = \sum_{n \in \mathbb{Z}} \boldsymbol{g}[k - k']\boldsymbol{x}[k'],$$

where $*$ denotes convolution, and $\boldsymbol{g}[k] \in \mathbb{C}^{P \times R}$ is the MIMO channel impulse response matrix. Hence we have

$$\boldsymbol{Y}[\nu] = \boldsymbol{G}[\nu]\boldsymbol{X}[\nu], \tag{3}$$

where $\boldsymbol{X}[\nu]$, $\boldsymbol{Y}[\nu]$, and $\boldsymbol{G}[\nu]$ are the Fourier transforms of $\boldsymbol{x}[k]$, $\boldsymbol{y}[k]$, and $\boldsymbol{g}[k]$ respectively. We call $\boldsymbol{G}[\nu]$ the *channel transfer function matrix*. The channel outputs are uniformly subsampled by a factor of L, and the resulting sequences are denoted by $\boldsymbol{z}[n] = \boldsymbol{y}[nL]$, $n \in \mathbb{Z}$. Using (3), we now have

$$\boldsymbol{Z}[\nu] = \frac{1}{L}\sum_{l=0}^{L-1} \boldsymbol{Y}\left[\frac{\nu + l}{L}\right] = \frac{1}{L}\sum_{l=0}^{L-1} \boldsymbol{G}\left[\frac{\nu + l}{L}\right]\boldsymbol{X}\left[\frac{\nu + l}{L}\right], \quad \nu \in [0, 1). \tag{4}$$

We model the reconstruction block as follows:

$$\tilde{\boldsymbol{x}}[k] = \sum_{n \in \mathbb{Z}} \boldsymbol{h}[k - nL]\boldsymbol{z}[n], \tag{5}$$

where $\boldsymbol{h}[k] \in \mathbb{C}^{R \times P}$ is the impulse response matrix of the reconstruction filter. From (5), it is obvious that the entire system consisting of the channel, subsampling, and reconstruction is invariant to time-shifts by any multiple of L, *i.e.*

$$\boldsymbol{x} \to \tilde{\boldsymbol{x}} \implies \boldsymbol{x}[\cdot - nL] \to \tilde{\boldsymbol{x}}[\cdot - nL], \quad \forall n \in \mathbb{Z}.$$

Conversely, (5) describes the most general linear transformation that allows this invariance. Taking the Fourier transform of (5), we obtain

$$\tilde{\boldsymbol{X}}[\nu] = \boldsymbol{H}[\nu]\boldsymbol{Z}[L\nu], \quad \nu \in [0,1), \tag{6}$$

where $\boldsymbol{H}[\nu]$, the Fourier transform of $\boldsymbol{h}[n]$, is called the *reconstruction filter matrix*. Since $\boldsymbol{Z}[\nu]$ is a periodic function, we can rewrite (6) as

$$\tilde{\boldsymbol{X}}\left[\nu + \frac{l'}{L}\right] = \boldsymbol{H}\left[\nu + \frac{l'}{L}\right]\boldsymbol{Z}[L\nu], \quad l' \in \mathbb{Z},\ \nu \in [0,1/L). \tag{7}$$

We can now write (4) and (7) compactly as

$$\boldsymbol{Z}[L\nu] = \boldsymbol{\mathcal{G}}[\nu]\boldsymbol{\mathcal{X}}[\nu], \tag{8}$$

$$\tilde{\boldsymbol{\mathcal{X}}}[\nu] = \boldsymbol{\mathcal{H}}[\nu]\boldsymbol{Z}[L\nu] \tag{9}$$

for $\nu \in [0,1/L)$, where $\boldsymbol{\mathcal{X}}[\nu] \in \mathbb{C}^{LR}$ is defined as

$$\mathcal{X}_{Rl+r}[\nu] = X_r\left[\nu + \frac{l}{L}\right], \quad (r,l) \in \mathcal{R} \times \mathcal{L}, \tag{10}$$

with an analogous definition for $\tilde{\boldsymbol{\mathcal{X}}}[\nu] \in \mathbb{C}^{LR}$, while $\boldsymbol{\mathcal{G}}[\nu] \in \mathbb{C}^{P\times RL}$ and $\boldsymbol{\mathcal{H}}[\nu] \in \mathbb{C}^{RL\times P}$ are the *modulated channel and reconstruction transfer function matrices* defined as

$$\mathcal{G}_{p,Rl+r}[\nu] = \frac{1}{L}G_{pr}\left[\nu + \frac{l}{L}\right], \quad (p,r,l) \in \mathcal{P} \times \mathcal{R} \times \mathcal{L}, \tag{11}$$

$$\mathcal{H}_{Rl+r,p}[\nu] = H_{rp}\left[\nu + \frac{l}{L}\right], \quad (p,r,l) \in \mathcal{P} \times \mathcal{R} \times \mathcal{L}. \tag{12}$$

In the next section, we provide precise conditions for stable reconstruction of the channel inputs from the subsampled output sequences. In particular, for FIR implementation reasons, we are interested in a reconstruction filter matrix $\boldsymbol{H}[\nu]$ whose entries are continuous in ν. Specifically, the continuity guarantees that the approximation error resulting from the FIR implementation can be made arbitrarily small by choosing sufficiently long FIR filters. This point will be elaborated later.

IV. Conditions for Perfect Reconstruction

In this section, we present the condition for perfect reconstruction from the MIMO channel outputs in the discrete-time setting. More specifically, we provide conditions on the channel transfer function matrix $\boldsymbol{G}[\nu]$ that guarantee stable reconstruction of the inputs *with*/*without* the continuity requirement on the reconstruction filter matrix $\boldsymbol{H}[\nu]$. These results are discrete-time versions of their continuous-time counterparts in [13].

Let $\mathbb{H} = \mathcal{B}_d(\mathcal{F}_1) \times \cdots \times \mathcal{B}_d(\mathcal{F}_R)$ denote the class of inputs to the MIMO channel. Then $\mathbb{H}$ is a Hilbert space equipped with the following inner product:

$$\langle \boldsymbol{x}, \boldsymbol{w} \rangle = \sum_{n \in \mathbb{Z}} \boldsymbol{w}^H[n]\boldsymbol{x}[n], \quad \boldsymbol{x}, \boldsymbol{w} \in \mathbb{H}.$$

Naturally, the norm on $\mathbb{H}$ is defined as $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$.

We first review an important notion called *stability* of MIMO sampling [13, 14].

Definition 1. *The MIMO sampling scheme is called* stable *if there exist constants $A, B > 0$ such that*

$$A\|\boldsymbol{x}\|^2 \leq \sum_{n \in \mathbb{Z}} \|\boldsymbol{z}[n]\|^2 \leq B\|\boldsymbol{x}\|^2, \tag{13}$$

for all $\boldsymbol{x} \in \mathbb{H}$.

The implication of (13) is that we can reconstruct the inputs from the output samples $\boldsymbol{z}[n]$ in a stable manner, in the sense that small perturbations in the inputs or the channel output samples, cannot cause large errors in the reconstructed outputs. The quantity $K = \sqrt{B/A} \geq 1$ is called the *condition number* of the sampling scheme, and it is a bound on the amplification of the normalized 2-norm of the error due to the reconstruction filters [13]. In particular the notion of stable sampling may be expressed as a frame-theoretic condition. We refer the reader to [35] for more about frames. The theory of frames provides a convenient tool to study sampling [36, 37].

Now, define the following index sets

$$\begin{aligned} \mathcal{K}_\nu &= \left\{ Rl + r : (r, l) \in \mathcal{R} \times \mathcal{L} \text{ and } \nu + l/L \in \mathcal{F}_r \right\}, \\ \mathcal{K}_\nu^c &= \mathcal{Z} \backslash \mathcal{K}_\nu \end{aligned} \tag{14}$$

for $\nu \in [0, 1/L]$, where $\mathcal{Z} = \{0, \ldots, RL-1\}$. Just as in the continuous-time case in [13], we can decompose the interval $[0, 1/L]$ into a union of intervals where $\mathcal{K}_\nu$ is piecewise constant.

Proposition 2. *Suppose that sets $\mathcal{F}_r$, $r \in \mathcal{R}$ have multiband structure as defined in (2). Then there exists a collection of disjoint intervals $\mathcal{I}_m$ and sets $\mathcal{K}_m$, $m = 1, \ldots, M$ such that*

$$\bigcup_{m=1}^{M} \mathcal{I}_m = \left[0, \frac{1}{L}\right) \quad \textit{and} \quad \mathcal{K}_\nu = \mathcal{K}_m, \quad \forall \nu \in \mathcal{I}_m.$$

This result is easily demonstrated by using an argument very similar to the one in [28]. We write $\mathcal{I}_m$ as

$$\mathcal{I}_m = [\gamma_m, \gamma_{m+1}), \quad m \in \mathcal{M},$$

with $\gamma_1 < \gamma_2 < \cdots < \gamma_{M+1}$, such that $\gamma_1 = 0$ and $\gamma_{M+1} = 1/L$. For convenience we also define

$$q_m \stackrel{\text{def}}{=} |\mathcal{K}_m|.$$

Equation (14) implies that all nonzero entries of $\boldsymbol{\mathcal{X}}[\nu]$ are captured in $\boldsymbol{\mathcal{X}}_{\mathcal{K}_\nu}[\nu]$. Hence, from (8) and (9) we conclude that for perfect reconstruction $\boldsymbol{\mathcal{H}}_{\mathcal{K}_\nu,\bullet}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_\nu}[\nu] = \boldsymbol{I}_{|\mathcal{K}_\nu|}$ and $\boldsymbol{\mathcal{H}}_{\mathcal{K}_\nu^c,\bullet}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_\nu}[\nu] = \mathbf{0}$ hold a.e. This characterization of $\boldsymbol{\mathcal{H}}$ that provides perfect reconstruction can be written compactly as

$$\boldsymbol{\mathcal{H}}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_\nu}[\nu] = \boldsymbol{I}_{\bullet,\mathcal{K}_\nu}, \quad \text{a.e.} \tag{15}$$

where $\boldsymbol{I}$ is the $RL \times RL$ identity matrix. Since $\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_\nu}[\nu] \in \mathbb{C}^{P\times|\mathcal{K}_\nu|}$, (15) requires that $\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_\nu}(\nu)$ have full column rank a.e.

As in the continuous-time case [13], it can be easily verified that $\boldsymbol{H}[\nu]$ is continuous if and only if $\boldsymbol{\mathcal{H}}[\nu]$ is continuous on $[0, 1/L]$, and the following "boundary conditions" hold:

$$\boldsymbol{\mathcal{H}}_{\mathcal{K},\bullet}\left[\frac{1}{L}\right] = \boldsymbol{\mathcal{H}}_{\mathcal{K}',\bullet}[0] \tag{16}$$

for all $\mathcal{K} \subseteq \mathcal{Z}$, where $\mathcal{K}' = (\mathcal{K} \oplus R) \bmod RL$. As we shall see later, in order to achieve continuity of $\boldsymbol{H}[\nu]$, it is convenient to impose continuity on $\boldsymbol{G}[\nu]$, and this produces a similar condition on $\boldsymbol{\mathcal{G}}[\nu]$. Specifically, if $G_{pr}[\nu]$ is continuous on $\overline{\mathcal{F}}_r$ (the closure of $\mathcal{F}_r$) then $\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_m}[\nu]$ is continuous on $\overline{\mathcal{I}}_m$, and

$$\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}}\left[\frac{1}{L}\right] = \boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}'}[0], \tag{17}$$

for all $\mathcal{K} \subseteq \mathcal{Z}$.

The following theorems, which are discrete-time versions of similar results presented in [13], provide precise conditions for stable and perfect reconstruction of the channel inputs. We do not prove them as they can be deduced in a manner very similar to their continuous-time counterparts. Before stating the results, we point out that ess inf and ess sup denote the essential infimum and supremum respectively, *i.e.*,

$$\operatorname*{ess\,inf}_t g(t) = \sup\{\gamma : g(t) \geq \gamma \ a.e.\},$$
$$\operatorname*{ess\,sup}_t g(t) = \inf\{\gamma : g(t) \leq \gamma \ a.e.\}.$$

for any real function g.

Theorem 1. *Suppose that $\boldsymbol{G}[\nu]$ is such that $G_{pr}[\nu]$ is continuous for $\nu \in \overline{\mathcal{F}}_r$, and $\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_m}[\nu]$ has full column rank for all $m \in \mathcal{M}$, $\nu \in \overline{\mathcal{I}}_m = [\gamma_m, \gamma_{m+1}]$, then the MIMO sampling scheme is stable, and the stability bounds are given by*

$$A = L \inf_{\nu \in [0,1/L]} \lambda_{\min}\big(\boldsymbol{\mathcal{G}}^H_{\bullet,\mathcal{K}_\nu}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_\nu}[\nu]\big), \tag{18}$$

$$B = L \sup_{\nu \in [0,1/L]} \lambda_{\max}\big(\boldsymbol{\mathcal{G}}^H_{\bullet,\mathcal{K}_\nu}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_\nu}[\nu]\big). \tag{19}$$

Theorem 2. *Suppose that channel transfer function matrix $\boldsymbol{G}[\nu]$ is such that $G_{pr}[\nu]$ is continuous for $\nu \in \overline{\mathcal{F}}_r$, then there exists a reconstruction filter matrix $\boldsymbol{H}[\nu]$ continuous in ν that achieves stable and perfect reconstruction of the MIMO channel inputs if and only if*

$$\operatorname{rank}\big(\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_m}[\nu]\big) = |\mathcal{K}_m|, \quad \forall \nu \in \operatorname{int}\mathcal{I}_m = (\gamma_m, \gamma_{m+1}), \tag{20}$$

$$\operatorname{rank}\big(\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{J}_m}(\gamma_m)\big) = |\mathcal{J}_m|, \quad m \in \mathcal{M}. \tag{21}$$

where

$$\begin{aligned} \mathcal{J}_m &= \mathcal{K}_m \cup \mathcal{K}_{m-1}, \quad m = 2, \ldots, M, \\ \mathcal{J}_1 &= \mathcal{K}_1 \cup ((\mathcal{K}_M \oplus R) \bmod RL). \end{aligned} \tag{22}$$

Theorem 1 guarantees stability of reconstruction, but not continuity of $\boldsymbol{H}[\nu]$, while Theorem 2 guarantees both stability and continuity of at least one solution $\boldsymbol{H}[\nu]$. The continuity requirement is desirable from the viewpoint of implementation, as we see in Section V. A simple necessary condition for perfect reconstruction using continuous reconstruction filters is that $P \geq \max_m |\mathcal{J}_m|$. In Theorem 2, the assumption that $\boldsymbol{G}[\nu]$ is continuous in ν is made for convenience; it is possible for continuous perfect reconstruction filter matrix $\boldsymbol{H}[\nu]$ to exist, despite the lack of continuity of $\boldsymbol{G}[\nu]$. However, this is rare and the conditions in the general case are cumbersome.

Finally, under certain conditions, perfect reconstruction is realizable using FIR filters. We say that a filter if FIR if its impulse response has a finite number of nonzero terms. An FIR filter need not be causal, but can be made causal by adding a finite delay. Let

$$\boldsymbol{G}^\star[z] = \sum_{k \in \mathbb{Z}} \boldsymbol{g}[k] z^{-k}$$

denote the Z-transform of $\boldsymbol{g}[k]$. We use the superscript "$\star$" here to distinguish between the Z-transform and the discrete-time Fourier transform $\boldsymbol{G}[\nu]$. Then clearly

$$\boldsymbol{G}[\nu] = \boldsymbol{G}^\star[\exp(j2\pi\nu)], \quad \nu \in \mathbb{R}.$$

Let $\boldsymbol{H}^\star[z]$ denote the Z-transform of $\boldsymbol{h}[k]$. Finally, let $\boldsymbol{\mathcal{G}}^\star[z]$ and $\boldsymbol{\mathcal{H}}^\star[z]$ be the Z-domain analogues of the modulated matrices $\boldsymbol{\mathcal{G}}[\nu]$ and $\boldsymbol{\mathcal{H}}[\nu]$.

Theorem 3. *Suppose that the channel impulse response $\boldsymbol{g}[k]$ is a finite sequence of matrices, and let*

$$\mathcal{K} = \bigcup_{m=1}^{M} \mathcal{K}_m, \quad Q = |\mathcal{K}|.$$

Then perfect reconstruction using an FIR reconstruction filter matrix $\boldsymbol{H}^\star[z]$ is possible if and only if $P \geq Q$, and the $Q \times Q$ minors of $\boldsymbol{\mathcal{G}}^\star_{\bullet,\mathcal{K}}[z]$ have no zero common to all except $z = 0$ or $z = \infty$.

Proof. Suppose that $P \geq Q$ and the minors share no zeros except $z = 0$ or $z = \infty$. For every $\mathcal{J} \subseteq \mathcal{P}$ such that $|\mathcal{J}| = Q$, let

$$D_{\mathcal{J}}[z] = \det(\boldsymbol{\mathcal{G}}^\star_{\mathcal{J},\mathcal{K}}[z]).$$

Therefore, by Bezout's identity, there exist polynomials $A_{\mathcal{J}}[z]$ such that

$$\sum_{\mathcal{J} \subseteq \mathcal{P}: |\mathcal{J}| = Q} A_{\mathcal{J}}[z] D_{\mathcal{J}}[z] = z^d \tag{23}$$

for some $d \in \mathbb{Z}$. Let $\tilde{\boldsymbol{\mathcal{H}}}_{\mathcal{J}}[z]$ denote the adjoint of $\boldsymbol{\mathcal{G}}^\star_{\mathcal{J},\mathcal{K}}[z]$ so that

$$\tilde{\boldsymbol{\mathcal{H}}}_{\mathcal{J}}[z] \boldsymbol{\mathcal{G}}^\star_{\mathcal{J},\mathcal{K}}[z] = D_{\mathcal{J}}[z] \boldsymbol{I}, \tag{24}$$

where $\boldsymbol{I}$ is the $Q \times Q$ identity matrix. Define

$$\boldsymbol{\mathcal{H}}^\star_{\mathcal{K},\bullet}[z] = z^{-d} \sum_{\mathcal{J} \subseteq \mathcal{P}: |\mathcal{J}| = Q} A_{\mathcal{J}}[z] \tilde{\boldsymbol{\mathcal{H}}}_{\mathcal{J}}[z] \boldsymbol{I}_{\mathcal{J},\bullet} \quad \text{and} \quad \boldsymbol{\mathcal{H}}^\star_{\mathcal{K}^c,\bullet}[z] = \boldsymbol{0}.$$

where $\mathcal{K} = \bigcup_m \mathcal{K}_m$. Then $\boldsymbol{\mathcal{H}}^\star[z]$ clearly corresponds to an FIR reconstruction filter matrix $\boldsymbol{H}^\star[z]$. Therefore,

$$\begin{aligned} \boldsymbol{\mathcal{H}}^\star_{\mathcal{K},\bullet}[z] \boldsymbol{\mathcal{G}}^\star_{\bullet,\mathcal{K}}[z] &= z^{-d} \sum_{\mathcal{J}} A_{\mathcal{J}}[z] \tilde{\boldsymbol{\mathcal{H}}}_{\mathcal{J}}[z] \boldsymbol{I}_{\mathcal{J},\bullet} \boldsymbol{\mathcal{G}}^\star_{\bullet,\mathcal{K}}[z] \\ &= z^{-d} \sum_{\mathcal{J}} A_{\mathcal{J}}[z] \tilde{\boldsymbol{\mathcal{H}}}_{\mathcal{J}}[z] \boldsymbol{\mathcal{G}}^\star_{\mathcal{J},\mathcal{K}}[z] = \boldsymbol{I} \end{aligned}$$

where the last step follows from (23) and (24). We also obviously have

$$\boldsymbol{\mathcal{H}}^\star_{\mathcal{K}^c,\bullet}[z]\boldsymbol{\mathcal{G}}^\star_{\bullet,\mathcal{K}}[z] = \mathbf{0}.$$

Combining the last two results, we obtain $\boldsymbol{\mathcal{H}}^\star[z]\boldsymbol{\mathcal{G}}^\star_{\bullet,\mathcal{K}}[z] = \boldsymbol{I}_{\bullet,\mathcal{K}}$. From this it follows that

$$\boldsymbol{\mathcal{H}}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_m}[\nu] = \boldsymbol{I}_{\bullet,\mathcal{K}_m}, \quad \forall \nu \in \mathcal{I}_m. \tag{25}$$

which is essentially equivalent to (15). Thus, we have found an FIR realization of perfect reconstruction filters.

Conversely, suppose that $\boldsymbol{h}[k]$ is an FIR filter matrix achieving prefect reconstruction. Then $\boldsymbol{\mathcal{G}}[\nu]$ and $\boldsymbol{\mathcal{H}}[\nu]$ are infinitely differentiable since both $\boldsymbol{g}[k]$ and $\boldsymbol{h}[k]$ are finite sequences. Because, by (25), the entire function $\boldsymbol{\mathcal{H}}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_m}[\nu] - \boldsymbol{I}_{\bullet,\mathcal{K}_m}$ vanishes on an interval, it follows that it must vanish everywhere, and

$$\boldsymbol{\mathcal{H}}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}_m}[\nu] = \boldsymbol{I}_{\bullet,\mathcal{K}_m}$$

holds for all $\nu \in \mathbb{R}$, rather than just $\nu \in \mathcal{I}_m$. Therefore, we obtain $\boldsymbol{\mathcal{H}}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{K}}[\nu] = \boldsymbol{I}_{\bullet,\mathcal{K}}$ for all $\nu \in \mathbb{R}$, where $\mathcal{K} = \bigcup_m \mathcal{K}_m$. Equivalently, we have

$$\boldsymbol{\mathcal{H}}^\star[z]\boldsymbol{\mathcal{G}}^\star_{\bullet,\mathcal{K}}[z] = \boldsymbol{I}_{\bullet,\mathcal{K}}. \tag{26}$$

in the Z-domain. If $P < Q$ then $\operatorname{rank}(\boldsymbol{\mathcal{G}}^\star_{\bullet,\mathcal{K}}[z]) \leq Q-1$, implying that (26) fails to hold. Similarly, if all the $Q \times Q$ minors of $\boldsymbol{\mathcal{G}}^\star_{\bullet,\mathcal{K}}[z]$ share a common factor of the form $(z - z_0)$ where $z_0 \neq 0$, then $\boldsymbol{\mathcal{G}}^\star_{\bullet,\mathcal{K}}[z]$ loses rank at $z = z_0$, and this contradicts (26) because $\boldsymbol{\mathcal{H}}^\star[z]$ is FIR and can not cancel the $(z - z_0)$ factor. This proves the converse statement. □

The proof of Theorem 3 is partly based on similar results in [38, 39]. The import of this result is that perfect reconstruction is possible using FIR filters provided that the channel has a finite impulse response, and that the modulated channel transfer function matrix in the Z-domain $\boldsymbol{\mathcal{G}}^\star[z]$ is sufficiently “diverse” in the sense that its null space is empty for all $z \notin \{0, \infty\}$. Of course, we do not care about the cases $z = 0$ or $z = \infty$ because no causality requirement is imposed on the FIR filters. Suppose that $P = Q$, then $\boldsymbol{\mathcal{G}}^\star_{\bullet,\mathcal{K}}[z]$ is a $Q \times Q$ matrix and the necessary and sufficient condition for perfect reconstruction using FIR filters reduces to

$$\det \boldsymbol{\mathcal{G}}^\star_{\bullet,\mathcal{K}}[z] = K z^{-d}, \quad K \neq 0,\ d \in \mathbb{Z}.$$

This condition is similar to the perfect reconstruction condition for filter banks.

The problem in [39] deals with existence of FIR equalizer filters in the absence of downsampling of channel outputs, while the classical filter bank problem deals with a single-input multiple-output channel whose outputs are decimated. In the present problem the existence of an FIR reconstruction filter matrix depends not only on the channel transfer function matrix $\boldsymbol{G}[\nu]$ (as in [39]), but also on the decimation factor L (as in the filter bank problem), and band-structure of the inputs through $\mathcal{K}$. Thus, Theorem 3 generalizes and solves all these problems simultaneously.

V. Reconstruction Filter Design

A. Reconstruction Error

In this section, we study the problem of reconstruction filter design for a given MIMO sampling scheme. We have seen in Section IV that under certain conditions on the channel and the class of input signals, perfect reconstruction is possible. Unfortunately these ideal filters are not necessarily FIR filters. Conversely FIR filters do not generally guarantee perfect reconstruction of the channel inputs. Nevertheless, we can approximate the ideal reconstruction filters using FIR filters chosen judiciously so that an appropriate cost function, such as the end-to-end distortion, is minimized.

We model the input signals as discrete-time multiband functions $X_r[\nu] = 0$, $\nu \notin \mathcal{F}_r$ with $\boldsymbol{x} \in \mathcal{C}$, where $\mathcal{C}$ is the constraint set for the channel inputs:

$$\boldsymbol{x} \in \mathcal{C} = \{\boldsymbol{x} : \|x_r\| \leq \gamma_r\},$$

i.e., the input signal energies are upper bounded. The reconstruction filters are approximated by FIR filters, *i.e.*, we enforce the following parameterization on $\boldsymbol{H}[\nu]$:

$$H_{rp}[\nu] = \sum_{k \in \mathcal{Q}_{rp}} \alpha_{rpk} e^{-j2\pi\nu k}, \quad r \in \mathcal{R},\ p \in \mathcal{P}, \tag{27}$$

where $\mathcal{Q}_{rp}$ is a finite set representing the locations of the nonzero filter coefficients of $H_{rp}[\nu]$. We choose

$$\mathcal{Q}_{rp} = \left\{k : \kappa_{rp} \leq k \leq l_{rp} + \kappa_{rp} - 1\right\},$$

where l_{rp} is the length of the FIR reconstruction filter $H_{rp}[\nu]$ and κ_{rp} is the position of the first filter coefficient. This FIR parameterization no longer guarantees perfect reconstruction, and the objective is to minimize the norm of the resulting reconstruction error $\boldsymbol{e}[k] = \tilde{\boldsymbol{x}}[k] - \boldsymbol{x}[k]$. Define

$$\boldsymbol{\mathcal{E}}[\nu] = \tilde{\boldsymbol{\mathcal{X}}}[\nu] - \boldsymbol{\mathcal{X}}[\nu], \quad \nu \in \left[0, \frac{1}{L}\right]. \tag{28}$$

We shall now derive an expression for $\boldsymbol{\mathcal{E}}[\nu]$ as a function of the input signals, the channel and reconstruction filters alone. Define index sets

$$\mathcal{I}_r = (R\mathcal{L}) \oplus r = \{Rl + r : l \in \mathcal{L}\}, \tag{29}$$

$$\mathcal{K}_{r,\nu} = \mathcal{K}_\nu \cap \mathcal{I}_r = \left\{Rl + r : l \in \mathcal{L} \text{ and } \left(\nu + \frac{l}{L}\right) \in \mathcal{F}_r\right\} \tag{30}$$

for each $r \in \mathcal{R}$. It is clear from (10) and (29) that

$$\boldsymbol{\mathcal{X}}_{\mathcal{I}_r}[\nu] = \begin{pmatrix} X_r[\nu] & X_r[\nu + \frac{1}{L}] & \cdots & X_r[\nu + \frac{L-1}{L}] \end{pmatrix}^T,$$

with a similar expression for $\tilde{\boldsymbol{\mathcal{X}}}_{\mathcal{I}_r}[\nu]$, for each $r \in \mathcal{R}$, *i.e.*, these quantities are the length-L vectorized representations of $X_r[\nu]$ and $\tilde{X}_r[\nu]$ respectively. Hence, the energy of e_r can be expressed as a function of $\boldsymbol{\mathcal{E}}[\nu]$ using Parseval's theorem:

$$\|e_r\|^2 = \int_{[0,1]} |E_r[\nu]|^2 d\nu = \int_{[0,\frac{1}{L}]} \|\boldsymbol{\mathcal{E}}_{\mathcal{I}_r}[\nu]\|^2 d\nu. \tag{31}$$

Similar relations hold for x_r and other signals in terms of the vectorized version of their Fourier transforms. Now for each $r \in \mathcal{R}$ and $\nu \in [0, 1/L]$, (8) and its analog for $\tilde{\boldsymbol{\mathcal{X}}}_{\mathcal{I}_r}[\nu]$ yield

$$\begin{aligned}\tilde{\boldsymbol{\mathcal{X}}}_{\mathcal{I}_r}[\nu] &= \boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu]\boldsymbol{\mathcal{G}}[\nu]\boldsymbol{\mathcal{X}}[\nu] \\ &= \sum_{s\in\mathcal{R}} \boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{I}_s}[\nu]\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu],\end{aligned} \tag{32}$$

where the second step holds because the sets $\{\mathcal{I}_r\}$ partition $\{0, 1, \ldots, RL-1\}$. Therefore (28) and (32) give us

$$\begin{aligned}\boldsymbol{\mathcal{E}}_{\mathcal{I}_r}[\nu] &= \tilde{\boldsymbol{\mathcal{X}}}_{\mathcal{I}_r}[\nu] - \boldsymbol{\mathcal{X}}[\nu] \\ &= \sum_{s\in\mathcal{R}} \boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{I}_s}[\nu]\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu] - \boldsymbol{\mathcal{X}}[\nu] \\ &= \sum_{s\in\mathcal{R}} \Big(\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{I}_s}[\nu] - \delta_{rs}\boldsymbol{I}_L\Big)\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu],\end{aligned} \tag{33}$$

where δ_{rs} is the Kronecker delta function and $\boldsymbol{I}_L$ is the identity matrix of size $L \times L$. We know from Proposition 1 that

$$\boldsymbol{\mathcal{X}}_{\mathcal{K}_{s,\nu}}[\nu] = \boldsymbol{I}_{\mathcal{K}_{s,\nu},\mathcal{I}_s}\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu], \tag{34}$$

where $\boldsymbol{I}$ is the identity matrix of size $LR \times RL$. Since $\boldsymbol{\mathcal{X}}_{\mathcal{K}_{s,\nu}}[\nu]$ captures all the nonzero components of

$\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu]$, we can invoke Proposition 1 again to write

$$\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu] = \boldsymbol{I}_{\mathcal{I}_s,\mathcal{K}_{s,\nu}}\boldsymbol{\mathcal{X}}_{\mathcal{K}_{s,\nu}}[\nu]. \tag{35}$$

Combining (34) and (35) we obtain

$$\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu] = \boldsymbol{E}_{s,\nu}\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu], \tag{36}$$

where

$$\boldsymbol{E}_{s,\nu} \stackrel{\text{def}}{=} \boldsymbol{I}_{\mathcal{I}_s,\mathcal{K}_{s,\nu}}\boldsymbol{I}_{\mathcal{K}_{s,\nu},\mathcal{I}_s} \tag{37}$$

is a diagonal matrix with zeros or ones on the diagonal. Hence (33) and (36) yield

$$\boldsymbol{\mathcal{E}}_{\mathcal{I}_r}[\nu] = \sum_{s\in\mathcal{R}} \boldsymbol{T}^{rs}[\nu]\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu], \tag{38}$$

$$\boldsymbol{T}^{rs}[\nu] = \Big(\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{I}_s}[\nu] - \delta_{rs}\boldsymbol{I}_L\Big)\boldsymbol{E}_{s,\nu}. \tag{39}$$

We point out that if $\boldsymbol{H}[\nu]$ is a perfect reconstruction filter matrix, then using (15), it is easily shown that

$$\boldsymbol{T}^{rs}[\nu] = \Big(\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{I}_s}[\nu] - \delta_{rs}\boldsymbol{I}_L\Big) = \boldsymbol{0}. \tag{40}$$

For simplicity we rewrite (38) as

$$e_r = \sum_{s\in\mathcal{R}} T^{rs}x_s, \tag{41}$$

where $T^{rs} : \mathcal{B}_d(\mathcal{F}_s) \to l_2$ is the linear operator equivalent of $\boldsymbol{T}^{rs}[\nu]$ acting on x_s.

The norm of the operator $\boldsymbol{T}^{rs}[\nu]$, which is needed later, can be computed as follows

$$\|T^{rs}\|^2 = \sup_{\substack{\|x_s\|\le 1\\ x_s\in\mathcal{B}_d(\mathcal{F}_s)}} \|T^{rs}x_s\|^2 \tag{42}$$

$$= \sup \int_{\nu\in[0,\frac{1}{L}]} \|\boldsymbol{T}^{rs}[\nu]\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu]\|^2 d\nu \quad \text{s.t.} \quad \int_{\nu} \|\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu]\|^2 d\nu \le 1, \quad x_s \in \mathcal{B}_d(\mathcal{F}_s) \tag{43}$$

Note that the condition $x_s \in \mathcal{B}_d(\mathcal{F}_s)$ implies that some entries of $\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu]$ necessarily vanish due to (36) because $\boldsymbol{E}_{s,\nu}$ is a diagonal matrix with zeros or ones on the diagonal. In other words, $\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu]$ is not an arbitrary vector in $C^{L\times 1}$. From (39) and (37), it is clear that if the k-th component of $\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu]$ vanishes for some k, then the k-th column of $\boldsymbol{T}^{rs}[\nu]$ also vanishes. Hence the range space of $(\boldsymbol{T}^{rs}[\nu])^H$ equals the signal space for input s, namely $X_s[\nu]$. Now it follows immediately that we can drop the constraint $x_s \in \mathcal{B}_d(\mathcal{F}_s)$ in (43) to obtain

$$\|T^{rs}\|^2 = \sup_{\nu\in[0,\frac{1}{L}]} \|\boldsymbol{T}^{rs}[\nu]\|^2 = \sup_{\nu\in[0,\frac{1}{L}]} \sigma_{\max}(\boldsymbol{T}^{rs}[\nu]), \tag{44}$$

where for each ν $\|\boldsymbol{T}^{rs}[\nu]\|$ is the spectral norm of matrix $\boldsymbol{T}^{rs}[\nu]$, and $\sigma_{\max}(\cdot)$ is the largest singular value.

B. *The cost function*

Our problem is to design an FIR reconstruction filter matrix $\boldsymbol{H}[\nu]$ such that a measure of the reconstruction error $\boldsymbol{e}$ is minimized. Using (12) we see that

$$\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu] = \begin{pmatrix} \boldsymbol{H}_{r,\bullet}[\nu] \\ \boldsymbol{H}_{r,\bullet}[\nu + \frac{1}{L}] \\ \vdots \\ \boldsymbol{H}_{r,\bullet}[\nu + \frac{L-1}{L}] \end{pmatrix}$$

depends only on the r-th row of $\boldsymbol{H}[\nu]$, namely $\boldsymbol{H}_{r,\bullet}[\nu]$. We also see from (27) that $\boldsymbol{H}_{r,\bullet}[\nu]$ is a linear combination parameterized by the coefficients

$$\boldsymbol{\alpha}^r = \{\alpha_{rpk} : \ p \in \mathcal{P}, \ k \in \mathcal{Q}_{rp}\}.$$

which is a subset of the entire set of filter coefficients. Therefore, $\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu]$ depends only on $\boldsymbol{\alpha}^r$. In view of (39), $\boldsymbol{T}^{rs}[\nu]$ depends only on $\boldsymbol{\alpha}^r$ and the channel transfer function matrix $\boldsymbol{G}[\nu]$. It follows from (31) and (38) that

$$\|e_r\|^2 = \int_{[0,\frac{1}{L}]} \Big\| \sum_{s\in\mathcal{R}} \boldsymbol{T}^{rs}[\nu]\boldsymbol{\mathcal{X}}_{\mathcal{I}_s}[\nu], \Big\|^2 d\nu. \tag{45}$$

which is completely parameterized by $\boldsymbol{\alpha}^r$. Consequently, for each $r \in \mathcal{R}$, the set of coefficients $\boldsymbol{\alpha}^r$ (or equivalently the r-th row of $\boldsymbol{H}[\nu]$) can be optimized independently of the others by minimizing a cost that measures the fidelity of reconstruction of the r-th input. Our choice of the cost function is the norm of the error in the worst case when $\boldsymbol{x} \in \mathcal{C}$, *i.e.*,

$$C_r(\boldsymbol{\alpha}^r) = \sup \|e_r\| \quad \text{s.t.} \quad \|x_s\| \le \gamma_s, \quad x_s \in \mathcal{B}_d(\mathcal{F}_s), \quad s \in \mathcal{R}. \tag{46}$$

It turns out that (46) is difficult to minimize directly, so we look for an alternate expression for the cost such as a bound on $C_r(\boldsymbol{\alpha}^r)$. The following proposition provides upper and lower bounds on the cost function.

Proposition 3. *The cost function $C_r(\boldsymbol{\alpha}^r)$ can be bounded as*

$$\frac{1}{\sqrt{R}}\bar{C}_r(\boldsymbol{\alpha}^r) \le C_r(\boldsymbol{\alpha}^r) \le \bar{C}_r(\boldsymbol{\alpha}^r),$$

where

$$\bar{C}_r(\boldsymbol{\alpha}^r) \stackrel{\text{def}}{=} \sum_{s\in\mathcal{R}} \gamma_s \|T^{rs}\| = \sum_{s\in\mathcal{R}} \gamma_s \sup_{\nu\in[0,\frac{1}{L}]} \|\boldsymbol{T}^{rs}[\nu]\|.$$

Proof. From (41) and (46) we obtain an upper bound on the cost function:

$$\begin{aligned} C_r(\boldsymbol{\alpha}^r) &= \sup \Big\| \sum_{s\in\mathcal{R}} T^{rs} x_s \Big\| \quad \text{s.t.} \quad \|x_s\| \le \gamma_s, \quad x_s \in \mathcal{B}_d(\mathcal{F}_s), \quad s\in\mathcal{R} \\ &\le \sup \sum_{s\in\mathcal{R}} \big\| T^{rs} x_s \big\| \quad \text{s.t.} \quad \|x_s\| \le \gamma_s, \quad x_s \in \mathcal{B}_d(\mathcal{F}_s), \quad s\in\mathcal{R} \\ &= \sum_{s\in\mathcal{R}} \gamma_s \|T^{rs}\| = \sum_{s\in\mathcal{R}} \gamma_s \sup_{\nu\in[0,\frac{1}{L}]} \|\boldsymbol{T}^{rs}[\nu]\|, \end{aligned}$$

where the last step follows from (44). Hence $C_r(\boldsymbol{\alpha}^r) \le \bar{C}_r(\boldsymbol{\alpha}^r)$.

To prove the other inequality, we start by choosing a set of signals $x'_s \in \mathcal{B}_d(\mathcal{F}_s)$, $s\in\mathcal{R}$ such that $\|x'_s\| = \gamma_s$ and $\|T^{rs}x'_s\| = \gamma_s\|T^{rs}\| - \epsilon/R$, where $\epsilon > 0$. In view of (42), such x'_s exist for any $\epsilon > 0$. First let $\bar{x}_0 = x'_0$. Then, for $s = 1, 2, \ldots, R-1$, let $\bar{x}_s$ be either x'_s or $-x'_s$ such that

$$\Big\langle T^{rs}\bar{x}_s, \sum_{\sigma=0}^{s-1} T^{r\sigma}\bar{x}_\sigma \Big\rangle \ge 0. \tag{47}$$

Hence we have $\|\bar{x}_s\| = \gamma_s$ and $\|T^{rs}\bar{x}_s\| = \gamma_s\|T^{rs}\| - \epsilon/R$. Now, (47) implies that for any $s\in\mathcal{R}$

$$\begin{aligned} \Big\| \sum_{\sigma=0}^{s} T^{r\sigma}\bar{x}_\sigma \Big\|^2 &= \|T^{rs}\bar{x}_s\|^2 + 2\Big\langle T^{rs}\bar{x}_s, \sum_{\sigma=0}^{s-1} T^{r\sigma}\bar{x}_\sigma \Big\rangle + \Big\|^2 \sum_{\sigma=0}^{s-1} T^{r\sigma}\bar{x}_\sigma \Big\| \\ &\ge \|T^{rs}\bar{x}_s\|^2 + \Big\| \sum_{\sigma=0}^{s-1} T^{r\sigma}\bar{x}_\sigma \Big\|^2. \end{aligned} \tag{48}$$

Therefore,

$$C_r(\boldsymbol{\alpha}^r)^2 \overset{(a)}{\ge} \Big\| \sum_{\sigma=0}^{R-1} T^{r\sigma}\bar{x}_\sigma \Big\|^2 \overset{(b)}{\ge} \sum_{s\in\mathcal{R}} \|T^{rs}\bar{x}_s\|^2 \overset{(c)}{=} \sum_{s\in\mathcal{R}} \gamma_s^2 \|T^{rs}\|^2 - \epsilon \tag{49}$$

where (a) follows from the definition of $C_r(\boldsymbol{\alpha}^r)$, (b) by recursively applying (48) starting with $s = R-1$, and (c) by the choice of $\bar{x}_s$. Now, using the Cauchy-Schwarz inequality, we have

$$\sum_{s\in\mathcal{R}} \gamma_s^2 \|T^{rs}\|^2 \ge \frac{1}{R}\Big(\sum_{s\in\mathcal{R}} \gamma_s \|T^{rs}\| \Big)^2 = \frac{1}{R}\bar{C}_r(\boldsymbol{\alpha}^r)^2. \tag{50}$$

Finally, from (49) and (50), and because the x'_s can be chosen to make ϵ arbitrarily small, we obtain the other desired inequality: $C_r(\boldsymbol{\alpha}^r) \geq \bar{C}_r(\boldsymbol{\alpha}^r)/\sqrt{R}$. $\square$

Instead of minimizing $C_r(\boldsymbol{\alpha}^r)$ to compute the optimal filter coefficients $\boldsymbol{\alpha}^r$, we minimize $\bar{C}_r(\boldsymbol{\alpha}^r)$ as it produces a considerably simpler algorithm to implement. Therefore the approximate optimal filter coefficients are given by

$$\boldsymbol{\alpha}^r_\circ = \arg\min_{\boldsymbol{\alpha}^r} \bar{C}_r(\boldsymbol{\alpha}^r), \quad \bar{C}_r(\boldsymbol{\alpha}^r) = \sum_{s\in\mathcal{R}} \gamma_s \sup_{\nu\in[0,\frac{1}{L}]} \|\boldsymbol{T}^{rs}[\nu]\|. \tag{51}$$

Owing to Proposition 3, the approximate solution will produce a cost that is greater by a factor of not more than $\sqrt{R}$ times the true minimum, *i.e.*,

$$\min_{\boldsymbol{\alpha}^r} C_r(\boldsymbol{\alpha}^r) \leq \bar{C}_r(\boldsymbol{\alpha}^r_\circ) \leq \sqrt{R}\min_{\boldsymbol{\alpha}^r} C_r(\boldsymbol{\alpha}^r). \tag{52}$$

An important question pertaining to the FIR design is whether the resulting approximation error goes to zero when the filter lengths go to infinity. Under some conditions, we can answer affirmatively, as the following theorem shows.

Theorem 4. *Suppose that $\boldsymbol{G}[\nu]$ is continuous and that $\boldsymbol{H}[\nu]$ has an FIR parameterization described in (27). If there exists a perfect reconstruction filter matrix continuous in ν, then $\min_{\boldsymbol{\alpha}^r} C_r(\boldsymbol{\alpha}^r) \to 0$ as $\kappa_{rp} \to -\infty$ and $\kappa_{rp} + l_{rp} \to \infty$.*

Proof. In view of Proposition 3, it suffices to prove that $\lim_{\tau\to\infty} \min_{\boldsymbol{\alpha}^r} \|T^{rs}\| \to 0$ for all $r, s \in \mathcal{R}$, where

$$\tau = \min\left(\{-\kappa_{rp} : r \in \mathcal{R},\ p \in \mathcal{P}\} \cup \{\kappa_{rp} + l_{rp} : r \in \mathcal{R},\ p \in \mathcal{P}\}\right).$$

Suppose that $\boldsymbol{H}^\circ[\nu]$ is continuous in ν and achieves perfect reconstruction. Also let $\boldsymbol{\mathcal{H}}^\circ[\nu]$ be the modulated reconstruction matrix corresponding to $\boldsymbol{H}^\circ[\nu]$. From (40) we conclude that

$$\left(\boldsymbol{\mathcal{H}}^\circ_{\mathcal{I}_r,\bullet}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{I}_s}[\nu] - \delta_{rs}\boldsymbol{I}_L\right)\boldsymbol{E}_{s,\nu} = \boldsymbol{0} \tag{53}$$

for all $r, s \in \mathcal{R}$ because we would be guaranteed perfect reconstruction if we chose $\boldsymbol{\mathcal{H}}[\nu] = \boldsymbol{\mathcal{H}}^\circ[\nu]$. Then combining (39) and (53) we obtain

$$\boldsymbol{T}^{rs}[\nu] = \left(\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu] - \boldsymbol{\mathcal{H}}^\circ_{\mathcal{I}_r,\bullet}[\nu]\right)\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{I}_s}[\nu]\boldsymbol{E}_{s,\nu}$$

for any reconstruction matrix $\boldsymbol{\mathcal{H}}[\nu]$. Therefore,

$$\sup_{\nu\in[0,\frac{1}{L}]} \left\|\boldsymbol{T}^{rs}[\nu]\right\| \leq \sup_{\nu\in[0,\frac{1}{L}]} \left\|\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu] - \boldsymbol{\mathcal{H}}^{\circ}_{\mathcal{I}_r,\bullet}[\nu]\right\| \cdot \left\|\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{I}_s}[\nu]\boldsymbol{E}_{s,\nu}\right\|$$
$$\leq C \sup_{\nu\in[0,\frac{1}{L}]} \left\|\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu] - \boldsymbol{\mathcal{H}}^{\circ}_{\mathcal{I}_r,\bullet}[\nu]\right\|,$$

where

$$C = \sup_{\nu\in[0,1/L]} \left\|\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{I}_s}[\nu]\boldsymbol{E}_{s,\nu}\right\|$$

is finite because $\boldsymbol{\mathcal{G}}[\nu]$ is a continuous function on the compact set $[0, 1/L]$, and $\boldsymbol{E}_{s,\nu}$ is a constant on each $\mathcal{I}_m$. Using (12) and (29) we obtain

$$\sup_{\nu\in[0,\frac{1}{L}]} \left\|\boldsymbol{T}^{rs}[\nu]\right\| \leq C\sqrt{L} \sup_{\nu\in[0,1]} \|\boldsymbol{H}_{r,\bullet}[\nu] - \boldsymbol{H}^{\circ}_{r,\bullet}[\nu]\|$$
$$\leq C\sqrt{LP} \max_{p} \sup_{\nu\in[0,1]} \|H_{rp}[\nu] - H^{\circ}_{rp}[\nu]\|. \tag{54}$$

Now suppose that $\boldsymbol{H}[\nu]$ has an FIR parameterization as in (27), *i.e.*,

$$H_{rp}[\nu] = \sum_{k\in\mathcal{Q}_{rp}} \alpha_{rpk} e^{-j2\pi\nu k},$$

where $\mathcal{Q}_{rp} = \{\kappa_{rp}, \ldots, \kappa_{rp} + l_{rp}\}$. Then clearly each entry of $\boldsymbol{H}[\nu]$ can be expressed as a trigonometric polynomial of degree at least τ. Moreover, the coefficients of the polynomial can be individually controlled by changing the parameters $\boldsymbol{\alpha}^r$. Equivalently, we can reparameterize so that the new parameters are the coefficients of the trigonometric polynomials (rather than the filter coefficients). Now, by the Stone-Weierstrass theorem [40], we obtain

$$\min_{\boldsymbol{\alpha}^r} \sup_{\nu\in[0,1]} \|H_{rp}[\nu] - H^{\circ}_{rp}[\nu]\| \leq \epsilon \tag{55}$$

for any $\epsilon > 0$ if τ is sufficiently large. Combining (54) and (55) we obtain the desired result:

$$\lim_{\tau\to\infty} \min_{\boldsymbol{\alpha}^r} \sup_{\nu\in[0,\frac{1}{L}]} \left\|\boldsymbol{T}^{rs}[\nu]\right\| = 0.$$

Incidentally this also proves that $\lim_{\tau\to\infty} \min_{\boldsymbol{\alpha}^r} \bar{C}_r(\boldsymbol{\alpha}^r) = 0$ by Proposition 3 □

Theorem 4 guarantees the existence of continuous FIR solutions $\boldsymbol{H}[\nu]$ that can get arbitrarily close to perfect reconstruction, provided that there exists a continuous $\boldsymbol{H}^{\circ}[\nu]$ with the perfect reconstruction property. Furthermore, empirical evidence (cf. V-D) suggests that for the optimal FIR approximation, the worst-case reconstruction error falls off exponentially fast with increasing filter length. Combined with Proposition 3 and (52) that follows from it, this suggest that the impact of using the surrogate cost function $\bar{C}(\boldsymbol{\alpha})$ instead of $C(\boldsymbol{\alpha})$ has a minimal impact: at most a small $O(\log R)$ increase in filter length required to meet a a fixed error tolerance.

C. *Semi-infinite linear program formulation*

Next, we present an algorithm to compute the optimal solution to the problem in (51). We show that this problem can be reduced to a *semi-infinite linear program* which can then be solved by a standard method.

We begin by expressing the matrices $\boldsymbol{T}^{rs}[\nu]$ as functions of the filter coefficients $\boldsymbol{\alpha}^r$.

Proposition 4. *The quantity $\boldsymbol{T}^{rs}[\nu]$ defined in (39) can be written as*

$$\boldsymbol{T}^{rs}[\nu] = \boldsymbol{F}_0^{rs}[\nu] + \sum_{p\in\mathcal{P}} \sum_{k\in\mathcal{Q}_{rp}} \alpha_{rpk} \boldsymbol{F}_{pk}^{rs}[\nu] \tag{56}$$

for appropriate matrices $\boldsymbol{F}_0^{rs}[\nu]$ and $\boldsymbol{F}_{pk}^{rs}[\nu]$.

Proof. Observe from (12) and (27) that the (l, p) entry of $\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu]$ is given by

$$[\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu]]_{lp} = H_{rp}\left[\nu + \frac{l}{L}\right] = \sum_{k\in\mathcal{Q}_{rp}} \alpha_{rpk} e^{-j2\pi(\nu+l/L)k}.$$

In other words $\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu]$ can be written as the following linear combination

$$\boldsymbol{\mathcal{H}}_{\mathcal{I}_r,\bullet}[\nu] = \sum_{p\in\mathcal{P}} \sum_{k\in\mathcal{Q}_{rp}} \alpha_{rpk} \boldsymbol{K}^{rpk}[\nu], \tag{57}$$

where $\boldsymbol{K}^{rpk}[\nu]$ are matrices whose entries are

$$[\boldsymbol{K}^{rpk}[\nu]]_{lp'} = \delta_{pp'} e^{-j2\pi(\nu+l/L)k}.$$

Combining (39) and (57), we obtain the desired affine form in (56), where

$$\boldsymbol{F}_0^{rs}[\nu] = -\delta_{rs}\boldsymbol{E}_{s,\nu} \quad \text{and} \quad \boldsymbol{F}_{pk}^{rs}[\nu] = \boldsymbol{K}^{rpk}[\nu]\boldsymbol{\mathcal{G}}_{\bullet,\mathcal{I}_s}[\nu]\boldsymbol{E}_{s,\nu}.$$

are the explicit expressions for the matrices involved, provided for the sake of completeness. ☐

Proposition 4 shows that $\boldsymbol{T}^{rs}[\nu]$ has an affine form in terms of the filter coefficients $\boldsymbol{\alpha}^r$. Next, for a fixed index r, we rewrite the optimization in (51) as

$$\min \sum_{s\in\mathcal{R}} \gamma_s\delta_s \quad \text{s.t.} \quad \delta_s \geq \Re\Big(\boldsymbol{y}^H\boldsymbol{T}^{rs}[\nu]\boldsymbol{x}\Big), \ \ \forall s \in \mathcal{R}, \forall\nu \in \left[0, \frac{1}{L}\right], \text{and } \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{B}_\circ, \tag{58}$$

where $\mathcal{B}_\circ = \{\boldsymbol{v} \in \mathbb{C}^L : \|\boldsymbol{v}\| \leq 1\}$ is the unit ball for length-L vectors. For convenience, we treat $\boldsymbol{\alpha}^r$ as a row-vector (with any ordering of coefficients), *i.e.*,

$$\boldsymbol{\alpha}^r_{j(p,k)} = \alpha_{rpk},$$

where $j(p,k)$ is an invertible mapping that takes the pair of indices $p \in \mathcal{P}$ and $k \in \mathcal{Q}_{rp}$ to a single index j in the set $\mathcal{J}_r$ defined as

$$\mathcal{J}_r = \{0, \ldots, J_r - 1\}, \quad J_r \stackrel{\text{def}}{=} \sum_{p \in \mathcal{P}} |\mathcal{Q}_{rp}| = \sum_{p \in \mathcal{P}} l_{rp}.$$

Recall that $\mathcal{Q}_{rp} = \{k : \kappa_{rp} \leq k \leq l_{rp} + \kappa_{rp} - 1\}$. Hence, an example of one such mapping is

$$j(p,k) = \sum_{p'=0}^{p-1} l_{rp'} + (k - \kappa_{rp}).$$

Define a row-vector $\boldsymbol{\delta} = \begin{bmatrix}\delta_0 & \cdots & \delta_{R-1}\end{bmatrix}$. Using the affine representation in (56), we can rewrite (58) as

$$\min \sum_{s \in \mathcal{R}} \gamma_s \delta_s \quad \text{s.t.} \quad \Re\Big(\delta_s - \sum_{p \in \mathcal{P}} \sum_{k \in \mathcal{Q}_{rp}} \alpha_{rpk} \big(\boldsymbol{y}^H \boldsymbol{F}^{rs}_{pk}[\nu] \boldsymbol{x}\big)\Big) \geq \Re\Big(\boldsymbol{y}^H \boldsymbol{F}^{rs}_0[\nu] \boldsymbol{x}\Big), \;\; \forall (s, \nu, \boldsymbol{x}, \boldsymbol{y}) \in \mathcal{U},$$

where $\mathcal{U} = \mathcal{R} \times [0, \frac{1}{L}] \times \mathcal{B}_\circ \times \mathcal{B}_\circ$. This problem can be recast as a semi-infinite linear program:

$$\min \Re(\boldsymbol{c}\boldsymbol{\xi}) \quad \text{s.t.} \quad \Re\big(\boldsymbol{a}(\boldsymbol{u})\boldsymbol{\xi}\big) \geq \Re\big(b(\boldsymbol{u})\big), \;\; \forall \boldsymbol{u} \in \mathcal{U}, \tag{59}$$

where $\boldsymbol{\xi} = \begin{bmatrix}\boldsymbol{\delta} & \boldsymbol{\alpha}^r\end{bmatrix}$ is the set of program variables, $\boldsymbol{u} = (s, \nu, \boldsymbol{x}, \boldsymbol{y}) \in \mathcal{U}$ parameterizes the constraints, and $b(\boldsymbol{u})$ is a complex-scalar. The quantities $\boldsymbol{a}(\boldsymbol{u})$ and $\boldsymbol{c}$ are row-vectors of length $R + J_r$ whose first R entries are real, and the remaining J_r entries are complex-valued:

$$\boldsymbol{a}_n(\boldsymbol{u}) = \begin{cases} 1, & \text{if } n = s, \\ 0, & \text{if } n \in \mathcal{R}, \;\; l \neq s, \\ -\boldsymbol{y}^H \boldsymbol{F}^{rs}_{pk}[\nu] \boldsymbol{x}, & \text{if } l = R + j(p,k), \end{cases}$$

$$b(\boldsymbol{u}) = \boldsymbol{y}^H \boldsymbol{F}^{rs}_0[\nu] \boldsymbol{x},$$

$$\boldsymbol{c}_n = \begin{cases} \gamma_s, & \text{if } n = s, \\ 0, & \text{otherwise.} \end{cases}$$

The semi-infinite program in (59) is in a nonstandard form, since it contains a mixture of real and complex variables. Nevertheless, it can be converted to the standard real form by decomposing all complex variables into their real and imaginary parts. Finding the dual of this real program, and reconverting to the complex form produces the following dual program:

$$\max \int_{\mathcal{U}} b(\boldsymbol{u}) dw(\boldsymbol{u}) \quad \text{s.t.} \quad \boldsymbol{c} = \int_{\mathcal{U}} \boldsymbol{a}(\boldsymbol{u}) dw(\boldsymbol{u}), \quad w \geq 0, \tag{60}$$

where w is a real and positive measure on $\mathcal{U}$. The dual problem can be solved using a simplex-type algorithm for semi-infinite programs [34].

Recall that whenever the technical conditions in Theorem 1 are satisfied, the set

$$\mathcal{S}_H = \{\boldsymbol{H}[\nu] : \text{perfect reconstruction is achieved.}\}$$

is nonempty. However $\mathcal{S}_H$ need not be a singleton set, because the perfect reconstruction filter matrices are not necessarily unique. The optimization always produces the FIR filter matrix that is "closest" to the set of reconstruction filter matrices $\mathcal{S}_H$. If the conditions in Theorem 2 are satisfied, then $\mathcal{S}_H$ contains a continuous $\boldsymbol{H}[\nu]$, and guarantees, by Theorem 4, that the approximation error can be made arbitrarily small by using sufficiently long FIR filters.

D. Design Examples

In this section, we consider two FIR filter design examples. In the first example, we design reconstruction filters for the multicoset sampling scheme which is a special case of MIMO sampling [28, 29]. In the second example, we consider MIMO sampling using a channel having two inputs and five outputs.

Example 1. *In this example, we design FIR reconstruction filters for multicoset sampling. Let* $\mathcal{F} = [0, 0.2) \cup [0.55, 0.75)$*, as illustrated in Figure 4, be the spectral support for the class of signals to be subsampled. The Landau lower bound on the downsampling rate for this spectral support is* 0.4 *(the total measure of* $\mathcal{F}$*). However, the minimum downsampling rate that can be achieved for this spectral support by uniform sampling is* 0.75*, because translates of* $\mathcal{F}$ *do not pack efficiently. (Recall that fractional downsampling is meaningful in the context of the continuous-time model underlying the discrete-time model – see the appendix).*

Consider instead the nonuniform multicoset subsampling set

$$\Lambda = \bigcup_{n \in \mathbb{Z}} \{4n, 4n+1\}$$

for $\mathcal{B}_d(\mathcal{F})$, i.e., *the sampling pattern is* $\mathcal{C} = \{0, 1\}$*. This corresponds to nonuniform downsampling by a factor of two, or a downsampling rate of* 0.5 *– just slightly higher than the Landau rate, and* $2/3$ *of the best* *uniform subsampling rate. This sampling scheme can be recast as a uniform MIMO sampling scheme with* $L = 4$ *and*

$$\boldsymbol{G}[\nu] = \begin{pmatrix} 1 \\ e^{-j2\pi\nu} \end{pmatrix}.$$

This is a single-input double-output channel, and we seek the optimal FIR reconstruction filter matrix $\boldsymbol{H}[\nu] = \begin{bmatrix} H_1[\nu] & H_2[\nu] \end{bmatrix}$ *of length* 21 *centered at the origin,* i.e. $\mathcal{Q}_{1p} = \{-10, \ldots, 10\}$, $p = 1, 2$. *Since* $R = 1$, *we can take* $\gamma_1 = 1$ *without loss of generality. Applying the semi-infinite algorithm, we obtain the optimal FIR filters* $H_1[\nu]$ *and* $H_2[\nu]$ *shown in Figure 5. The resulting maximum approximation error* $\|\boldsymbol{T}^{rs}[\nu]\|$ *at optimality is shown for* $\nu \in [0, 1/4)$ *in Figure 6. The equal-ripple nature of this plot is due to the minimax criterion:*

$$\delta = \min_{\boldsymbol{\alpha}^1} C_1(\boldsymbol{\alpha}^1) = \min_{\boldsymbol{\alpha}^1} \sup_{\nu} \|\boldsymbol{T}^{1s}[\nu]\|$$

The optimal cost is $\delta = 0.1348$.

Example 2. *Consider a MIMO channel with* $R = 2$ *inputs and* $P = 5$ *outputs. Suppose that the inputs* $x_1[k]$ *and* $x_2[k]$ *have spectral supports illustrated in Figure 7, namely*

$$\mathcal{F}_1 = [0, 0.4) \cup [0.75, 1.0) \;\; \textit{and} \;\; \mathcal{F}_2 = [0.25, 0.5).$$

Let $\gamma_1 = \gamma_2 = 0.5$ *be the weights and* $L = 4$ *the subsampling factor. For this choice we have* $M = 2$, $\mathcal{I}_1 = [0, 0.15)$ *and* $\mathcal{I}_1 = [0.15, 0.25)$. *Using (14) and (22), it is easy to check that*

$$\begin{aligned}
\mathcal{K}_1 &= \{0, 2, 3, 6\}, \\
\mathcal{K}_2 &= \{0, 3, 6\}, \\
\mathcal{J}_1 &= \mathcal{K}_1 \cup ((\mathcal{K}_2 \oplus 2) \bmod 8) = \{0, 2, 3, 5, 6\}, \\
\mathcal{J}_2 &= \mathcal{K}_2 \cup \mathcal{K}_1 = \{0, 2, 3, 6\}.
\end{aligned}$$

Hence by Theorem 1, $P \geq \max_m |\mathcal{K}_m| = 4$ *is required for the existence of a reconstruction filter matrix* $\boldsymbol{H}[\nu]$ *that achieves perfect reconstruction. If we also require that the filters be continuous, Theorem 2 states that* $P \geq \max_n |\mathcal{J}_n| = 5$ *is necessary. Let* $\boldsymbol{G}[\nu]$ *denote the following continuous channel transfer function matrix with* $P = 5$ *outputs:*

$$\boldsymbol{G}[\nu] = \begin{pmatrix} 1 & 1 \\ 1 & 1 + z^{-1} \\ z^{-1} & 0.25 + z^{-2} \\ 1 + 0.5z^{-1} & 1 + z^{-2} \\ 0.25 + z^{-2} & z^{-1} \end{pmatrix}, \;\; \textit{where} \;\; z = e^{j2\pi\nu}. \tag{61}$$

It can be verified numerically that (20) and (21) in Theorem 2 are satisfied. Hence, we are guaranteed the existence of a perfect reconstruction filter matrix $\boldsymbol{H}[\nu]$ that is continuous in ν. As a consequence of Theorem 4, the approximation error approaches zero as the filter lengths are increased. Using a semi-infinite algorithm, we design six sets of reconstruction filters of varying filter lengths, indexed by $\tau \in \{1, 2, \ldots, 6\}$, having the following specifications:

$$l_{rp} = 2\tau + 1$$

$$\kappa_{rp} = \begin{cases} 0 - \tau & \text{if } p \in \{0, 1\}, \\ 1 - \tau & \text{if } p \in \{2, 3, 4\}. \end{cases}$$

In other words, all the FIR reconstruction filters for a given τ have equal lengths ($2\tau + 1$). Furthermore the filters $g_{pr}[k]$ are centered at $k = 0$ for $p = 0, 1$, and at $k = 1$ for $p = 2, 3, 4$. Table I and Figure 8 show the cost functions for the two outputs and the six design cases. Observe that the cost falls off quickly as the filter lengths increase. Therefore, as discussed earlier, the use of the surrogate cost function $\bar{C}(\boldsymbol{\alpha})$ instead of $C(\boldsymbol{\alpha})$ leads to at most a slight increase in the lengths of filter required to meet a fixed error tolerance. In this example $\sqrt{R} = \sqrt{2}$, so that by (52)*, the required increase in length is two or less in most cases listed in Table I. Finally, in this example, the costs would converge to zero as $\tau \to \infty$, since the conditions required in Theorem 2 are satisfied.*

TABLE I

COST FUNCTIONS $\bar{C}_0(\boldsymbol{\alpha}_\circ^0)$ AND $\bar{C}_1(\boldsymbol{\alpha}_\circ^1)$ AT OPTIMALITY FOR FIR RECONSTRUCTION FILTERS OF LENGTH $2\tau + 1$, $1 \leq \tau \leq 6$.

$2\tau+1$	3	5	7	9	11	13
$\bar{C}_0(\boldsymbol{\alpha}_\circ^0)$	0.4835	0.3643	0.1093	0.0836	0.0716	0.0329
$\bar{C}_1(\boldsymbol{\alpha}_\circ^1)$	0.3554	0.1690	0.0637	0.0124	0.0076	0.0034

VI. CONCLUSION

We examined the problem of FIR reconstruction filter design for uniform MIMO sampling of multiband signals with different band structures. The analysis is facilitated by the conversion to an equivalent hypothetical discrete-time system. We presented necessary and sufficient conditions for perfect reconstruction of the channel inputs with and without a continuity requirement on the transfer functions of the reconstruction filters. We also presented necessary and sufficient conditions for the existence of FIR perfect reconstruction filters when the channel itself is FIR. These conditions, which depend on the channel, input multiband structures, and downsampling rate, generalize previous results on multichannel deconvolution and filter banks. In general, perfect reconstruction FIR filters do not exist for the MIMO sampling problem. Therefore, from an implementation viewpoint, we considered the problem of FIR approximation to the reconstruction system. The continuity property was shown to be important in this context, as it allows to make the signal reconstruction error arbitrarily small by designing sufficiently long filters in the FIR approximation.

Finally, we formulated the reconstruction filter design problem as a minimax optimization, which was recast as a standard semi-infinite linear program and solved efficiently by computer. The generality of the MIMO setting allows this algorithm to be used for various other sampling schemes that fit into the MIMO framework as special cases.

APPENDIX

FROM CONTINUOUS-TIME TO DISCRETE-TIME MODELS

The purpose of this appendage is to justify the discrete-time models for the MIMO channel and the reconstruction system in Section III. More specifically, we show that the continuous-time MIMO channel with sampling can be replaced by a hypothetical discrete-time channel with downsampling. The reconstruction system, which is implemented by digital signal processing, also has a discrete-time model, and the continuous-time inputs to the original continuous-time channel can eventually be reconstructed by using D/A converters at the output of the digital processing stage.

Channel Model

Suppose that the MIMO channel and its inputs have continuous-time models as described in [13] and illustrated in Figure 1. The input-output relation for the channel is

$$\boldsymbol{Y}(f) = \boldsymbol{G}(f)\boldsymbol{X}(f), \tag{A.1}$$

Now, each component of the input $\boldsymbol{x}(t)$ is a multiband signal whose Fourier transform $\boldsymbol{X}(f)$ has a bounded support $\mathcal{F}_r$. Hence, by the classical sampling theorem, $\boldsymbol{x}(t)$ is representable in terms of its samples taken uniformly at a sufficiently high rate (the Nyquist rate). Let

$$\mathcal{F} = \bigcup_{r\in\mathcal{R}} \mathcal{F}_r, \quad f_{\max} = \sup_{f\in\mathcal{F}} f, \quad f_{\min} = \inf_{f\in\mathcal{F}} f,$$

and suppose that $L \in \mathbb{Z}$ such that $(f_{\max} - f_{\min}) \leq L/T$. Hence it follows that

$$|f - f'| \geq \frac{L}{T} \implies \boldsymbol{J}(f)\boldsymbol{J}(f') = 0, \tag{A.2}$$

where $\boldsymbol{J}(f)$ is the following diagonal matrix:

$$\boldsymbol{J}(f) = \operatorname{diag}\big(\chi(f \in \mathcal{F}_1), \ldots, \chi(f \in \mathcal{F}_R)\big). \tag{A.3}$$

Observe that

$$\boldsymbol{J}(f)\boldsymbol{X}(f) = \boldsymbol{X}(f) \tag{A.4}$$

because $X_r(f) = 0$ for $f \notin \mathcal{F}_r$. Next, using (A.2) and (A.4), we have

$$|f' - f| \geq L/T \implies \boldsymbol{J}(f)\boldsymbol{X}(f') = \boldsymbol{J}(f)\boldsymbol{J}(f')\boldsymbol{X}(f') = 0. \tag{A.5}$$

So let us define sequences $\boldsymbol{x}[k]$, $\boldsymbol{y}[k]$, and a discrete-time transfer function $\boldsymbol{G}[\nu]$ as follows:

$$\begin{aligned}
\boldsymbol{x}[k] = \boldsymbol{x}\Big(\frac{kT}{L}\Big) &\iff \boldsymbol{X}[\nu] = \frac{L}{T}\sum_{l\in\mathbb{Z}} \boldsymbol{X}\Big(\frac{L(\nu+l)}{T}\Big) \\
\boldsymbol{y}[k] = \boldsymbol{y}\Big(\frac{kT}{L}\Big) &\iff \boldsymbol{Y}[\nu] = \frac{L}{T}\sum_{l\in\mathbb{Z}} \boldsymbol{Y}\Big(\frac{L(\nu+l)}{T}\Big) \\
\boldsymbol{G}[\nu] = \frac{T}{L}\sum_{l\in\mathbb{Z}} \boldsymbol{G}\Big(\frac{L(\nu+l)}{T}\Big)\boldsymbol{J}\Big(\frac{L(\nu+l)}{T}\Big)
\end{aligned} \tag{A.6}$$

Notice that discrete-time sequences (and their Fourier transforms) use square brackets, while continuous-time quantities use round brackets. The quantity $\boldsymbol{X}_r[\nu]$ is clearly supported on the set $\mathcal{F}_r^d = (T/L)\mathcal{F}$. We can assume, without loss of generality, that $\mathcal{F}_r^d \subseteq [0, 1)$. Taking the Fourier transform of $\boldsymbol{y}[k]$, we obtain

$$\begin{aligned}
\boldsymbol{Y}[\nu] &= \frac{T}{L}\sum_{l\in\mathbb{Z}} \boldsymbol{Y}\Big(\frac{L(\nu+l)}{T}\Big) \\
&\overset{(a)}{=} \frac{T}{L}\sum_{l\in\mathbb{Z}} \boldsymbol{G}\Big(\frac{L(\nu+l)}{T}\Big)\boldsymbol{J}\Big(\frac{L(\nu+l)}{T}\Big)\boldsymbol{X}\Big(\frac{L(\nu+l)}{T}\Big) \\
&\overset{(b)}{=} \frac{T}{L}\sum_{l\in\mathbb{Z}} \boldsymbol{G}\Big(\frac{L(\nu+l)}{T}\Big)\boldsymbol{J}\Big(\frac{L(\nu+l)}{T}\Big)\sum_{l'\in\mathbb{Z}} \boldsymbol{X}\Big(\frac{L(\nu+l')}{T}\Big) \\
&= \boldsymbol{G}[\nu]\boldsymbol{X}[\nu], \qquad \nu \in [0, 1],
\end{aligned} \tag{A.7}$$

where (a) follows from (A.1) and (A.4), while (b) holds because

$$l \neq l' \implies \boldsymbol{J}\Big(\frac{L(\nu+l)}{T}\Big)\boldsymbol{X}\Big(\frac{L(\nu+l')}{T}\Big) = 0,$$

which itself is a consequence of (A.5). Finally note that the actual observations from the channel output samples $\boldsymbol{z}[n]$ can be obtained by down-sampling $\boldsymbol{y}[k]$ by a factor of L, *i.e.*, $\boldsymbol{z}[n] = \boldsymbol{y}(nT) = \boldsymbol{y}[nL]$. Hence,

$$\boldsymbol{Z}[\nu] = \frac{1}{L}\sum_{l\in\mathcal{L}} \boldsymbol{Y}\Big[\frac{\nu+l}{L}\Big], \qquad \nu \in [0, 1], \tag{A.8}$$

where $\mathcal{L} = \{0, 1, \ldots, L-1\}$. In other words, everything to the left of the dashed line in Figure 1 can be represented by the discrete-time system shown in Figure 2. The block represents a discrete linear shift-invariant system with transfer function $\boldsymbol{G}[\nu]$.

Finally, observe that only one term of the summation in (A.6) is nonzero for $\nu \in [0, 1)$. because the terms do not overlap at any frequency ν due to (A.2). Therefore, if $G_{pr}(f)$ is continuous on the set $\overline{\mathcal{F}}_r$ (the closure of $\mathcal{F}_r$), then clearly $G_{pr}[\nu]$ is continuous on $\overline{\mathcal{F}^d}_r$. Hence the continuity property, which is desirable for implementation reasons, is preserved.

Reconstruction Model

We model the continuous-time reconstruction block as

$$\tilde{\boldsymbol{x}}(t) = \sum_{n \in \mathbb{Z}} \boldsymbol{h}(t - nT)\boldsymbol{z}[n], \tag{A.9}$$

where the matrix $\boldsymbol{h}(t) \in \mathbb{C}^{R \times P}$ is the impulse response of the reconstruction filter, and $\tilde{\boldsymbol{x}}(t)$ is the continuous-time reconstructed output. We assume that $\tilde{\boldsymbol{x}}(t)$ lies in $\mathbb{H}$. This is a reasonable assumption because, after all, $\tilde{\boldsymbol{x}}(t)$ is an estimate of $\boldsymbol{x}(t) \in \mathbb{H}$. Therefore, the sampling theorem implies that $\tilde{\boldsymbol{x}}(t)$ is fully represented by the sampled sequence $\tilde{\boldsymbol{x}}(kT/L)$. So, we model our reconstruction system as a discrete-time system producing an output $\tilde{\boldsymbol{x}}[k] = \tilde{\boldsymbol{x}}(kT/L)$, because it suffices to reconstruct $\tilde{\boldsymbol{x}}(t)$. Note that $\tilde{\boldsymbol{x}}(t) \in \mathbb{H}$ if each column of $\boldsymbol{h}(t)$ lies in $\mathbb{H}$. From (A.9) we see that

$$\tilde{\boldsymbol{x}}[k] = \sum_{n \in \mathbb{Z}} \boldsymbol{h}\Big(\frac{(k - nL)T}{L}\Big)\boldsymbol{z}[n] \equiv \sum_{k \in \mathbb{Z}} \boldsymbol{h}[k - Lk]\boldsymbol{z}[k], \tag{A.10}$$

where $\boldsymbol{h}[k] = \boldsymbol{h}(kT/L)$. We can rewrite (A.10) in the frequency domain as

$$\tilde{\boldsymbol{X}}[\nu] = \boldsymbol{H}[\nu]\boldsymbol{Z}[L\nu], \tag{A.11}$$

where

$$\boldsymbol{H}[\nu] = \frac{T}{L} \sum_{l \in \mathbb{Z}} \boldsymbol{H}\Big(\frac{L(\nu + l)}{T}\Big). \tag{A.12}$$

Therefore, the reconstruction block, shown to the right of the dashed line in Figure 1, can be replaced by the discrete-time system illustrated in Figure 3. Equation (A.6) describes the hypothetical discrete-time channel that replaces the continuous-time model, while (A.12) describes the real reconstruction system. Finally, since the discrete-time sequences $\tilde{\boldsymbol{x}}_r[k]$ represent samples of $\tilde{\boldsymbol{x}}_r(t)$ at a rate higher than the Nyquist rate, we can reconstruct $\tilde{\boldsymbol{x}}_r(t)$ by using D/A converters.

If the channel has a discrete-time model as shown in Figure 2, then the reconstruction block in Figure 3 is the most general structure with the property that the entire MIMO system (consisting of the channel, the down- and up-samplers, and the reconstruction filters) is invariant to shifts by multiples of L samples:

$$\boldsymbol{x}(k) \to \tilde{\boldsymbol{x}}(k) \implies \boldsymbol{x}(k-nL) \to \tilde{\boldsymbol{x}}(k-nL), \quad \forall k, n \in \mathbb{Z}.$$

The models presented here are applicable whether the discrete-time inputs represent underlying continuous-time function, or whether they are genuinely discrete-time by nature.

Finally, we point out that the sets $\mathcal{F}_r$ are not used in the discrete-time setting, thereby obviating the need for the superscript in $\mathcal{F}^d$, *i.e.*, we denote the spectral support of $x_r[k]$ by $\mathcal{F}_r$.

REFERENCES

[1] B. R. Petersen and D. D. Falconer, “Supression of adjacent-channel, co-channel, and intersimbol interference by equalizers and linear combiners,” *IEEE Trans. Comm.*, vol. 42, pp. 3109–3118, December 1994.

[2] J. Yang and S. Roy, “On joint receiver and transmitted optimization for multiple-input-multiple-output (MIMO) transmission systems,” *IEEE Trans. Comm.*, vol. 42, pp. 3221–3231, December 1994.

[3] G. G. Raleigh and J. M. Cioffi, “Spatio-temporal coding for wireless communication,” *IEEE Trans. Comm.*, vol. 46, no. 3, pp. 357–366, March 1998.

[4] L. Ye and K. J. R. Liu, “Adaptive blind source separation and equalization for multiple-input/multiple-output systems,” *IEEE Trans. Info. Theory*, vol. 44, no. 7, pp. 2864–2876, November 1998.

[5] W. Zishun and J. D. Z. chen, “Blind separation of slow waves and spikes from gastrointestinal myoelectrical recordings,” *IEEE Trans. Info. Tech. Biomed.*, vol. 5, no. 2, pp. 133–137, June 2001.

[6] V. Zarzoso and A. K. Nandi, “Noninvasive fetal electrocardiogram extraction: blind separation versus adaptive noise cancellation,” *IEEE Trans. Biomed. Eng.*, vol. 48, no. 1, pp. 12–18, January 2001.

[7] P. a. Voois and J. M. Cioffi, “Multichannel signal processing for multiple-head digital magnetic recording,” *IEEE Trans. Magnetics*, vol. 30, no. 6, pp. 5100–5114, November 1994.

[8] K.-C. Yen and Y. Zhao, “Adaptive co-channel speech separation and recognition,” *IEEE Trans. Speech Audio Process.*, vol. 7, no. 2, pp. 138–151, March 1999.

[9] A. González and J. J. Lopéz, “Fast transversal filters for deconvolution in multichannel sound reproduction,” *IEEE Trans. Speech Audio Process.*, vol. 9, no. 4, pp. 429–440, May 2001.

[10] J. Idier and Y. Goussard, “Multichannel seismic deconvolution,” *IEEE Trans. Geoscience Remote Sensing*, vol. 31, no. 5, pp. 961–979, September 1993.

[11] G. Harikumar and Y. Bresler, “Exact image deconvolution from multiple FIR blurs,” *IEEE Trans. Image Process.*, vol. 8, no. 6, pp. 846–862, June 1999.

[12] G. Harikumar and Y. Bresler, “Blind restoration of images blurred by multiple filters: theory and efficient algorithms,” *IEEE Trans. Image Processing*, vol. 8, pp. 202–219, Feb. 1999.

[13] R. Venkataramani and Y. Bresler, “Sampling theorems for uniform and periodic nonuniform MIMO sampling of multiband signals,” submitted to *IEEE Trans. Sig. Process.*.

[14] R. Venkataramani and Y. Bresler, “MIMO sampling: Necessary density conditions and stability issues,” submitted to *IEEE Trans. Info. Theory*.

[15] R. Venkataramani, *Sub-Nyquist Multicoset and MIMO Sampling: Perfect Reconstruction, Performance Analysis, and Necessary Density Conditions*, Ph.D. thesis, University of Illinois, Urbana-Champaign, IL, November 2001.

[16] R. J. Papoulis, “Generalized sampling expansions,” *IEEE Trans. Circuits Syst.*, vol. CAS-24, pp. 652–654, November 1977.

[17] R. E. Kahn and B. Liu, “Sampling representations and the optimum reconstruction of signals,” *IEEE Trans. Info. Theory*, vol. 11, no. 3, pp. 339–347, July 1965.

[18] J. L. Brown Jr., “Sampling expansions for multi-band signals,” *IEEE Trans. Acoust. Speech Signal Process.*, vol. 33, pp. 312–315, February 1985.

[19] K. Cheung and R. Marks, “Image sampling below the Nyquist density without aliasing,” *J. Opt. Soc. Am. A*, vol. 7, no. 1, pp. 92–105, January 1990.

[20] P. Vaidyanathan and V. Liu, “Efficient reconstruction of band-limited sequences from non-uniformly decimated versions by use of polyphase filter banks,” *IEEE Trans. Acoust., Speech Signal Process.*, vol. 38, pp. 1927–1936, November 1990.

[21] S. C. Scoluar and W. J. Fitzgerald, “Periodic nonuniform sampling of multi-band signals,” *Signal Process.*, vol. 28, no. 2, pp. 195–200, August 1992.

[22] K. Cheung, “A multidimensional extension of Papoulis’ generalized sampling expansion with the application in minimum density sampling,” in *Advanced Topics in Shannon Sampling and Interpolation Theory*, R. J. Marks II, Ed., pp. 85–119. Springer-Verlag, New York, 1993.

[23] R. G. Shenoy, “Nonuniform sampling of signals and applications,” in *Int. Symposium on Circuits and Systems*, London, May 1994, vol. 2, pp. 181–184.

[24] P. Feng and Y. Bresler, “Spectrum-blind minimum-rate sampling and reconstruction of multi-band signals,” in *Proc. IEEE Int. Conf. Acoust. Speech, Sig. Process.*, Atlanta, GA, May 1996.

[25] Y. Bresler and P. Feng, “Spectrum-blind minimum-rate sampling and reconstruction of 2-D multiband signals,” in *Proc. 3rd IEEE Int. Conf. on Image Processing, ICIP’96*, Lausanne, Switzerland, September 1996, vol. I, pp. 701–704.

[26] B. Foster and C. Herley, “Exact reconstruction from periodic nonuniform sampling of signals with arbitrary frequency support,” in *Proc. IEEE Int. Conf. Acoust. Speech Sig. Process.*, Detroit, MI, May 1998.

[27] C. Herley and P. W. Wong, “Minimum rate sampling of signals with arbitrary frequency support,” *IEEE Trans. Info Theory*, pp. 1555-6, vol.45, no.5, July 1999.

[28] R. Venkataramani and Y. Bresler, “Perfect reconstruction formulae and bounds on aliasing error in sub-Nyquist nonuniform sampling of multiband signals,” *IEEE Trans. Info. Theory*, vol. 46, no. 6, pp. 2173–2183, September 2000.

[29] R. Venkataramani and Y. Bresler, “Optimal sub-Nyquist nonuniform sampling and reconstruction of multiband signals,” *IEEE Trans. Sig. Process.*, vol. 49, no. 10, pp. 2301–2313, October 2001.

[30] H. Landau, “Necessary density conditions for sampling and interpolation of certain entire functions,” *Acta Math.*, vol. 117, pp. 37–52, 1967.

[31] D. Seidner and M. Feder, “Vector sampling expansions,” *IEEE Trans. Sig. Process.*, vol. 48, no. 5, pp. 1401–1416, May 2000.

[32] R. G. Shenoy, D. Burnside, and T. W. Parks, “Linear periodic systems and multirate filter design,” *IEEE Trans. Sig. Process.*, vol. 42, no. 9, pp. 2242–2255, September 1994.

[33] P. Moulin, M. Anitescu, K. O. Kortanek, and F. A. Potra, “The role of linear semi-infinite programming in signal-adapted QMF bank design,” *IEEE Trans. Sig. Process.*, vol. 45, no. 9, pp. 2160–2174, September 1997.

[34] E. J. Anderson and P. Nash, *Linear Programming in Infinite-Dimensional Spaces*, John Wiley & Sons, Chichester, 1987.

[35] R. M. Young, *An Introduction to Nonharmonic Fourier Analysis*, Academic Press, New York, 2001.

[36] J. R. Higgins, *Sampling Theory in Fourier and Signals Analysis Foundations*, Oxford Science Pub., New York, 1996.

[37] J. Benedetto and P. J. S. G. Ferreira, Eds., *Modern Sampling Theory: Mathematics and Applications*, Birkhäuser, Boston, 2001.

[38] C. A. Berenstein and D. C. Struppa, “On explicit solutions to the Bezout equation,” *Systems and Control Letters*, vol. 4, no. 1, pp. 33–39, February 1984.

[39] R. Rajagopal, "Exact FIR inverses of FIR filters," M.S. thesis, Ohio state University, 2000.

[40] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, New York, 1976.

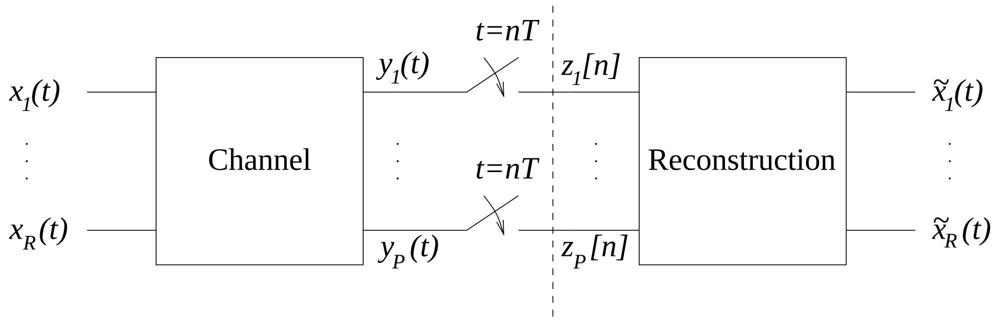

Fig. 1. Models for MIMO sampling and reconstruction. Only the sampled channel outputs $z_i[n]$ are observed, and the goal is to reconstruct the continuous-time channel inputs $x_i(t)$.

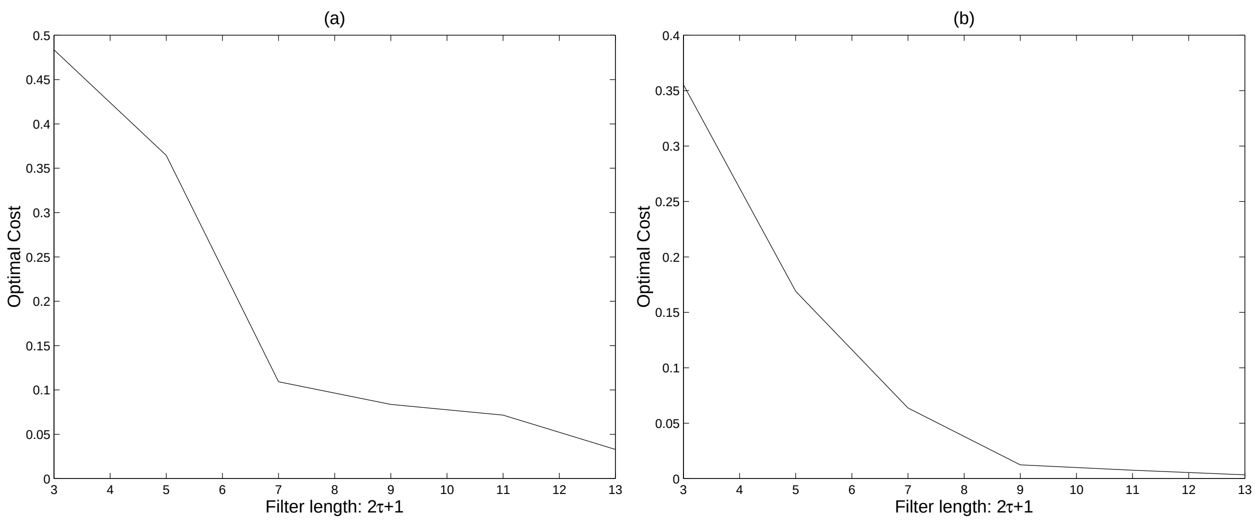

Fig. 8. Optimal costs (a) $\bar{C}_0(\boldsymbol{\alpha}_\circ^0)$ and (b) $\bar{C}_1(\boldsymbol{\alpha}_\circ^1)$ for FIR reconstruction filters of length $2\tau + 1$, $1 \leq \tau \leq 6$.

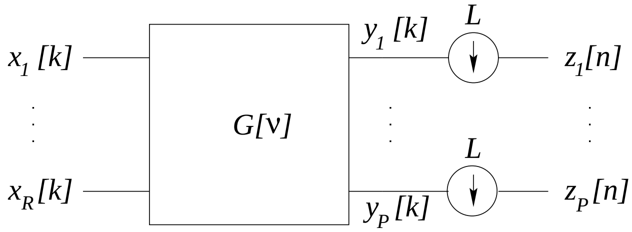

Fig. 2. Discrete-time model for the MIMO channel.

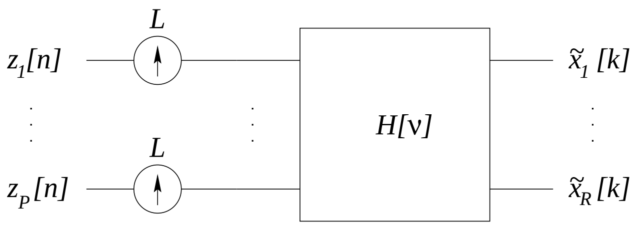

Fig. 3. Discrete-time model for MIMO reconstruction.

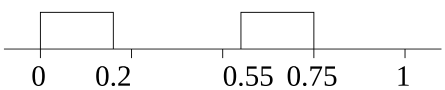

Fig. 4. Indicator function of the spectral support $\mathcal{F}$ for Example 1

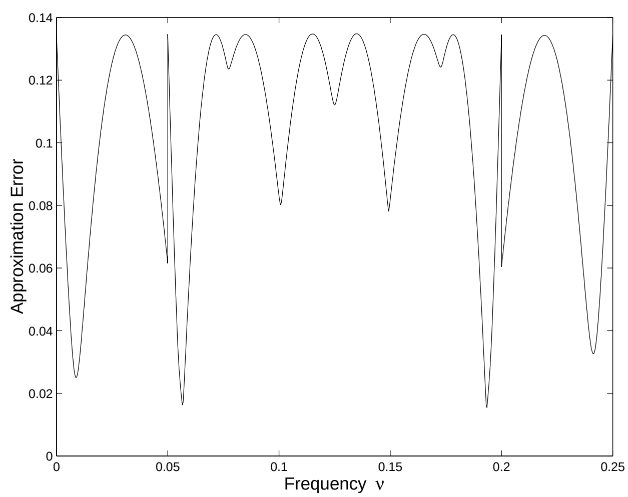

Fig. 6. Approximation error $||\boldsymbol{T}^{rs}[\nu]||$ at optimality.

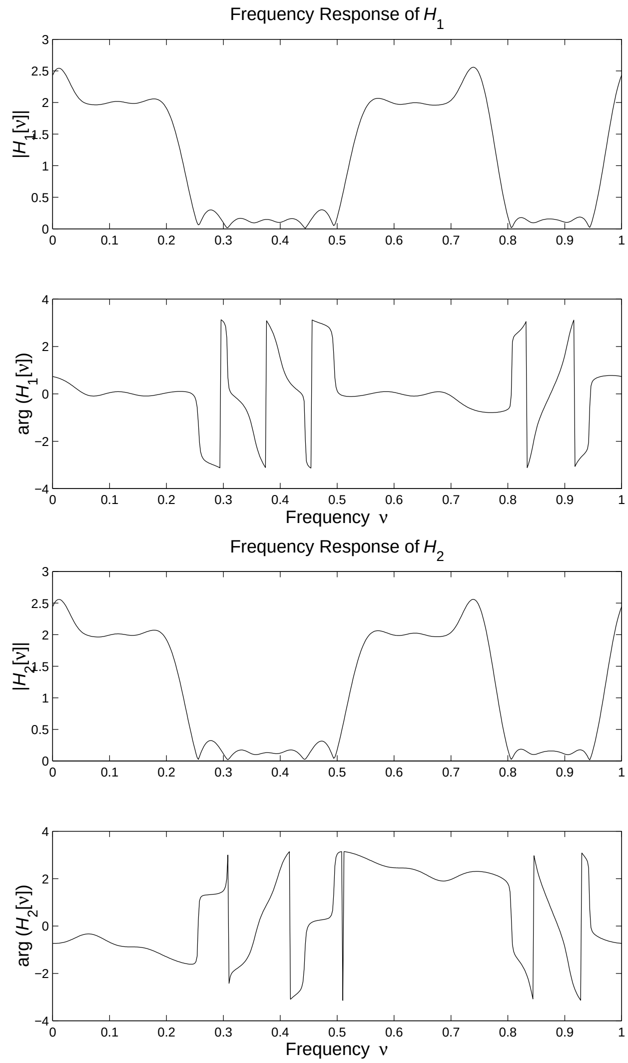

Fig. 5. Magnitude and phase responses of the optimal FIR filters $H_1[\nu]$ and $H_2[\nu]$.

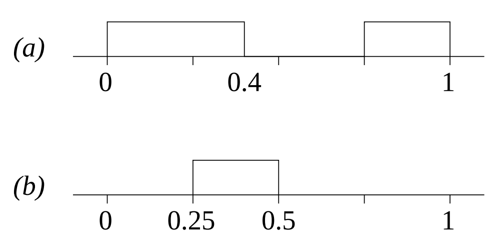

Fig. 7. (a) Spectral support of (a) $X_1[\nu]$, and (b) $X_2[\nu]$, for Example 2.