SUB-NYQUIST SAMPLING OF MULTIBAND SIGNALS: PERFECT RECONSTRUCTION AND BOUNDS ON ALIASING ERROR

Raman Venkataramani and Yoram Bresler

Coordinated Science Laboratory
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign, Urbana, IL 61801

ABSTRACT

We consider the problem of periodic nonuniform sampling of a multiband signal and its reconstruction from the samples. We derive the conditions for exact reconstruction and find an explicit reconstruction formula. Key features of this method are that the sampling rate can be made arbitrarily close to the minimum (Landau) rate and that it can handle classes of multiband signals that are not packable. We compute various bounds on the aliasing error due to mismodeling the spectral support and examine the performance in the presence of additive white sample noise. Finally we provide optimal designs for the reconstruction system.

1. INTRODUCTION

The use of nonuniform sampling for efficiently representating a multiband signal has been well studied by various authors [1, 2, 3, 4]. The general case of interest, illustrated in fig. 1(a), is when the structure of $\mathcal{F}$ (the spectral support of the signal) is such that it is not packable, *i.e.*,

$$\inf\{\theta > 0 : \mathcal{F} \cap (n\theta \oplus \mathcal{F}) = \emptyset, \quad \forall n \neq 0\} = \lambda([\mathcal{F}])$$

where $\oplus$ is the translation operator defined as $\theta \oplus \mathcal{F} \triangleq \{\theta + f : f \in \mathcal{F}\}$. In other words, the Nyquist rate for sampling $x(t)$ bandlimited to $\mathcal{F}$ is equal to the total width of the spectral span of $\mathcal{F}$, and sampling uniformly at any lower rate would cause aliasing. One of the most important advantages of nonuniform sampling is that sampling $x(t)$ at an average rate arbitrarily close to the Landau minimum rate, will generally guarantee exact reconstruction of $x(t)$ from its samples [3, 4].

Our analysis explicitly addresses the schemes suggested in [4] (apparently first proposed by [2]) but applies to any of the multicoset periodic sampling schemes.

1.1. SOME DEFINITIONS

The class of continuous complex multiband signals of finite energy with spectral support $\mathcal{F}$ (consisting of a finite union of bounded intervals), is denoted by

$$\begin{aligned} \mathcal{B}(\mathcal{F}) &= \{x(t) \in L^2(R) \cap C(R) : X(f) = 0, \ f \notin \mathcal{F}\} \\ \mathcal{F} &= \bigcup_{i=1}^{n} [a_i, b_i), \quad a_1 < b_1 < \ldots < a_n < b_n \end{aligned} \tag{1}$$

where $X(f) = \int_{-\infty}^{\infty} x(t) \exp(\Leftrightarrow j2\pi ft)dt$ is the the Fourier transform of $x(t)$. We let $[\mathcal{F}] = [a_1, b_n)$, and denote the *spectral span* of $\mathcal{F}$. Here $\lambda(\cdot)$ denotes the Lebesgue measure. We will frequently use $\chi(f; \mathcal{H})$ to denote the indicator function of a set $\mathcal{H}$

$$\chi(f; \mathcal{H}) = \begin{cases} 1, & f \in \mathcal{H} \\ 0, & f \notin \mathcal{H} \end{cases}$$

Supported in part by the Joint Services Electronic Program Grant No. N00014-90-J-1270 and the National Science Foundation Grant No. MIP 91-57377

2. MULTICOSET SAMPLING

Let $x(t) \in \mathcal{B}(\mathcal{F})$. We assume with no loss of generality that $a_1 = 0$ and $b_n = \frac{1}{T}$, the Nyquist rate for $x(t)$. We first pick a suitable integer $L > 0$ and then sample the input signal *nonuniformly* at the instants $t = (nL + c_i)T$ for $1 \leq i \leq p$ and $n \in Z$ where $\{c_i\}$ are p distinct integers contained in the set:

$$\mathcal{L} \triangleq \{0, 1, \ldots, L \Leftrightarrow 1\}$$

For a given c_i, the set of sampling instants $t = (nL+c_i)T, \ n \in Z$ has uniform intersample spacing equal to LT. We call this the i-th *active coset.* We shall refer to the set $\mathcal{C} = \{c_i : 1 \leq i \leq p\}$ as the *sampling pattern* and the integer L as the *period of the pattern.*

Now consider the L discrete-time sequences defined below:

$$x_{[l]}(nT) \triangleq x(nT) \sum_{j=-\infty}^{\infty} \delta(n \Leftrightarrow (Lj + l)), \quad 0 \leq l \leq L \Leftrightarrow 1$$

It is clear that the sequence $x_{[c_i]}(n)$ contains the samples of the i-th active coset with samples separated by $L \Leftrightarrow 1$ interleaving zeros. It is straigthforward to compute the discrete-time Fourier transform of $x_{[l]}(n)$ using the Poisson summation formula.

$$\begin{aligned} X_{[l]}(e^{j2\pi fT}) &= \sum_{n=-\infty}^{\infty} x_{[l]}(nT) \exp(\Leftrightarrow j2\pi nfT) \\ &= \frac{1}{LT} \sum_{r=0}^{L-1} X\left(f + \frac{r}{LT}\right) \exp\left(\frac{j2\pi rl}{L}\right) \end{aligned} \tag{2}$$

It follows from (2) that for any $r \in Z$

$$X_{[l]}(e^{j2\pi(f+\frac{r}{LT})T}) = e^{(-j2\pi\frac{lr}{L})} X_{[l]}(e^{j2\pi fT}) \tag{3}$$

Therefore it suffices to examine $X_l(e^{j2\pi fT})$ for $f \in [0, \frac{1}{LT})$

$$X_l(e^{j2\pi fT}) = \frac{1}{LT} \sum_{r=0}^{L-1} X_r(f) \exp\left(\frac{j2\pi rl}{L}\right) \tag{4}$$

where

$$X_r(f) \triangleq X\left(f + \frac{r}{LT}\right) \chi(f; \mathcal{F}_0) \tag{5}$$

$$\mathcal{F}_0 \triangleq [0, \frac{1}{LT}) \tag{6}$$

In other words $X_r(f)$ is the r-th "spectral component" of the input (corresponding to the range of frequencies $\frac{r}{LT} \le f < \frac{r+1}{LT}$) shifted to the origin. Denoting the inverse Fourier transform of $X_r(f)$ by $x_r(t)$, we observe that

$$x(t) = \sum_{r=0}^{L-1} x_r(t) \exp\left(\frac{j2\pi rt}{LT}\right) \tag{7}$$

We will use (7) and (3) later in deriving the reconstruction equations. We shall now let $k = c_i,\ \ i = 1, 2, \ldots, p$ in (4) to get

$$X_{[c_i]}(e^{j2\pi fT}) = \frac{1}{LT} \sum_{r=0}^{L-1} \exp\left(\frac{j2\pi c_i r}{L}\right) X_r(f), \quad f \in \mathcal{F}_0 \tag{8}$$

This is the main equation relating the spectral components $X_r(f)$ to the information contained in the observed samples. Note that $X_{[c_i]}(e^{j2\pi fT})$ in the interval $\mathcal{F}_0$ contains relevant about the samples because $X_{[c_i]}(e^{j2\pi fT})e^{j2\pi fc_iT}$ is periodic with period $\frac{1}{LT}$. Reconstruction of the original signal $x(t)$ is achieved if we recover its spectral components $\{x_r(t)\}$.

3. RECONSTRUCTION

Our primary objective is to invert the set of linear equations (8) to obtain $X_r(f)$. The recovery of $x(t)$ is then essentially an application of (7). Notice that with no further assumptions about the class of input signals, (8) cannot be solved because there are fewer equations (p) than unknown variables (L) for each $f \in \mathcal{F}_0$. For a given sampling pattern $\mathcal{C}$, a sufficient condition for solving (8) is that $\mathcal{C}$ be chosen wisely and that at any frequency $f \in \mathcal{F}_0$, no more than p of the quantities $\{X_r(f)\}$ for $r \in \mathcal{L}$ be nonzero. More specifically, we will show that $x(t) \in \mathcal{B}(\mathcal{F})$ can be reconstructed from the samples if the indicator function of $\mathcal{F}$ satisfies

$$\sum_{r=0}^{L-1} \chi\left(f + \frac{r}{LT}; \mathcal{F}\right) \le p, \quad f \in \mathcal{F}_0 \tag{9}$$

and if $\mathcal{C}$ is "universal" (as defined later). Our focus in this section will be the exact reconstruction of $x(t) \in \mathcal{B}(\mathcal{F})$ when $\mathcal{F}$ satisfies (9). The analogous problem for real signals has been considered by Herley and Wong [4] who did not, however, provide an explicit reconstruction formula. We shall derive the reconstruction equations formally.

Let $\mathcal{F}$ be as in (1) with $a_1 = 0$ and $b_n = \frac{1}{T}$. Consider the finite set

$$\Gamma \triangleq \bigcup_{i=1}^{n} \left\{ a_i \Leftrightarrow \frac{\lfloor LTa_i \rfloor}{LT} \right\} \cup \left\{ b_i \Leftrightarrow \frac{\lfloor LTb_i \rfloor}{LT} \right\}$$

where $\lfloor \cdot \rfloor$ is the floor function. Repeated values are listed only once in Γ. Let $\{\gamma_m\}$ be the $M \le 2n$ distinct elements of Γ arranged in increasing order. Then $\gamma_1 = 0$ since $a_1 = 0$. Additionally, we define $\gamma_{M+1} \triangleq \frac{1}{LT}$. Then the collection of intervals $\{\mathcal{G}_m\}$

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

clearly partitions the set $\mathcal{F}_0$. The importance of the above constructions is clarified in the following theorem.

Theorem 1 *For each $r \in \mathcal{L}$ and $m \in \{1, 2, \ldots, M\}$ the function $\chi(f + \frac{r}{LT}; \mathcal{F})$ is constant over the interval $\mathcal{G}_m$.*

The theorem states that each of the "subcells" $\frac{r}{LT} \oplus \mathcal{G}_m$ (*i.e.* translates of $\mathcal{G}_m$) is either fully contained in $\mathcal{F}$ of disjoint from it. This interpretation of the theorem motivates the following definition of sets $\mathcal{K}_m$ and their complements for $m \in \{1, 2, \ldots, M\}$.

$$\mathcal{K}_m \triangleq \left\{ r \in \mathcal{L} : \frac{r}{LT} \oplus \mathcal{G}_m \subset \mathcal{F} \right\} \quad \bar{\mathcal{K}}_m = \mathcal{L} \backslash \mathcal{K}_m \tag{10}$$

We can express these sets in terms of their elements as $\mathcal{K}_m = \{k_m(l) : 1 \le l \le q_m\}$ and $\bar{\mathcal{K}}_m = \{\bar{k}_m(l) : 1 \le l \le L \Leftrightarrow q_m\}$ where q_m is the number of elements in $\mathcal{K}_m$. In view of theorem 1, we see that the left-hand side of (9) equals q_m where m is the unique index (for each $f \in \mathcal{F}_0$) such that $f \in \mathcal{G}_m$. Therefore (9) now reduces to

$$q_m \le p, \quad 1 \le m \le M \tag{11}$$

Fig. 1 illustrates the entire process of construction of all the rel-

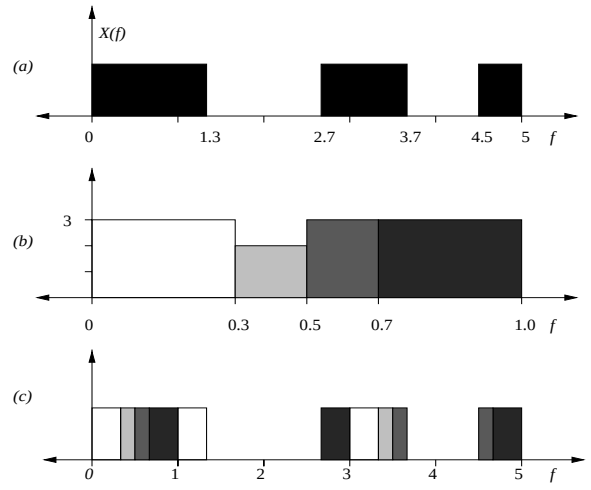

Figure 1: (a) Indicator function of the spectral support of span $\frac{1}{T} = 5$ (b) Number of overlaps (for $L = 5$) which is constant on each of the four sets $\mathcal{G}_m$ (c) The active spectral subcells. For each m, the translates of $\mathcal{G}_m$ are shown in the same color.

evant sets. Next we relate these spectral components to the data samples. We define the $p \times 1$ data matrix $\mathbf{y}(f)$ as

$$[\mathbf{y}(f)]_i = T\sqrt{L} X_{[c_i]}(e^{j2\pi fT}) \tag{12}$$

This matrix clearly contains all the information present in the samples taken nonuniformly. For each m, we define matrices $\mathbf{A}_m$, $\mathbf{B}_m$, $\mathbf{x}_m^+(f)$ and $\mathbf{x}_m^-(f)$ (of sizes $p \times q_m$, $p \times (L \Leftrightarrow q_m)$, $q_m \times 1$, $(L \Leftrightarrow q_m) \times 1$ respectively) as follows:

$$\begin{aligned} [\mathbf{A}_m]_{il} &= \tfrac{1}{\sqrt{L}} \exp\left(\tfrac{j2\pi c_i k_m(l)}{L}\right) \\ [\mathbf{B}_m]_{il} &= \tfrac{1}{\sqrt{L}} \exp\left(\tfrac{j2\pi c_i \bar{k}_m(l)}{L}\right) \\ [\mathbf{x}_m^+(f)]_l &= X_{k_m(l)}(f)\chi(f; \mathcal{G}_m) \\ [\mathbf{x}_m^-(f)]_l &= X_{\bar{k}_m(l)}(f)\chi(f; \mathcal{G}_m) \end{aligned} \tag{13}$$

$\mathbf{A}_m$ and $\mathbf{B}_m$ are submatrices of the $L \times L$ DFT matrix and the quantities $\mathbf{x}_m^+(f)$ and $\mathbf{x}_m^-(f)$ determine the "in-band" and "out-of-band" portions of the spectrum $X(f)$. Evidently $x_m^-(f) = \mathbf{0}$ whenever $x(t) \in \mathcal{B}(\mathcal{F})$. However, when $x(t) \notin \mathcal{B}(\mathcal{F})$, $x_m^-(f) \neq \mathbf{0}$, leading to aliasing. Therefore (8), (12) and (13) yield

$$\mathbf{y}(f) = \mathbf{A}_m \mathbf{x}_m^+(f) + \mathbf{B}_m \mathbf{x}_m^-(f), \quad \forall f \in \mathcal{G}_m \tag{14}$$

for each m. Suppose $x(t) \in \mathcal{B}(\mathcal{F})$. Then it is clear that $\mathbf{x}_m^+(f)$ can be recovered from $\mathbf{y}(f)$ if and only if $\mathbf{A}_m$ has full rank. This imposes a mild "universality" condition on the sampling pattern [3].

Definition 1 *A pattern $\mathcal{C} = \{c_1, \ldots, c_p\}$ is said to be* (p, q)-universal *if every selection of $q \leq p$ distinct columns of the matrix*

$$\left[\frac{1}{\sqrt{L}} \exp\left(\frac{j2\pi c_i l}{L}\right)\right]_{il} \tag{15}$$

is a linearly independent set. A pattern is simply universal *if it is (p, p)-universal.*

Note that a (p, q)-universal pattern $\mathcal{C}$ is (p, r)-universal for each $r \leq q$. The "bunched" sampling pattern $\mathcal{C} = \{0, 1, \ldots, p \Leftrightarrow 1\}$ is an example of a universal pattern (for arbitrary L) because any set of p columns of (15) would form a Vandermonde matrix. Equation (11) and (p, q')-universality of $\mathcal{C}$ (where $q' = \max_m q_m$) are sufficient for $\mathbf{A}_m$ to have full column rank for each m. we will assume throughout that this is the case. We can now reconstruct the spectral components of the input, $\tilde{\mathbf{x}}_m^{\pm}(f)$ defined similarly as $\mathbf{x}_m^{\pm}(f)$, using the following formula

$$\begin{bmatrix} \tilde{\mathbf{x}}_m^+(f) \\ \tilde{\mathbf{x}}_m^-(f) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_m^{\text{inv}} \\ \mathbf{C}_m \end{bmatrix} \mathbf{y}(f)\chi(f; \mathcal{G}_m) \tag{16}$$

for $1 \leq m \leq M$, where $\mathbf{A}_m^{\text{inv}}$ and $\mathbf{C}_m$ (sizes $q_m \times p$ and $(L \Leftrightarrow q_m) \times p$ respectively) are (nonunique) matrices satisfying

$$\mathbf{A}_m^{\text{inv}} \mathbf{A}_m = \mathbf{I} \quad \text{and} \quad \mathbf{C}_m \mathbf{A}_m = \mathbf{0} \tag{17}$$

for each m. $\mathbf{A}_m^{\text{inv}}$ is guaranteed to exist by choice of a good $\mathcal{C}$ while $\mathbf{C}_m$ exists trivially. Equations (14), (16) and (17) yield

$$\begin{bmatrix} \tilde{\mathbf{x}}_m^+(f) \\ \tilde{\mathbf{x}}_m^-(f) \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{A}_m^{\text{inv}} \mathbf{B}_m \\ \mathbf{0} & \mathbf{C}_m \mathbf{B}_m \end{bmatrix} \begin{bmatrix} \mathbf{x}_m^+(f) \\ \mathbf{x}_m^-(f) \end{bmatrix} \tag{18}$$

We have perfect reconstruction when $\mathbf{x}_m^-(f) = \mathbf{0} \Leftrightarrow x(t) \in \mathcal{B}(\mathcal{F})$. One of the problems we will consider is picking "good" matrices $\mathbf{A}_m^{\text{inv}}$ and $\mathbf{C}_m$ that make the transformation matrix in (18) as close to the identity matrix as possible. This is important because we want to minimize aliasing errors when $x(t) \notin \mathcal{B}(\mathcal{F})$. These equations specify all the information required to reconstruct the spectrum $X(f)$ on all its spectral subcells. We now present the interpolation formula for $x(t) \in \mathcal{B}(\mathcal{F})$ which is obtained by converting (16) to the time domain

$$x(t) = \sum_{i=1}^{p} \sum_{j \in Z} x((c_i + Lj)T)\phi_i(t \Leftrightarrow (c_i + Lj)T) \tag{19}$$

where $\phi_i(t)$ is the inverse Fourier transform of $\Phi_i(f)$,

$$\Phi_i(f) = \begin{cases} T\sqrt{L}[\mathbf{A}_m^{\text{inv}}]_{li} e^{j2\pi \frac{c_i k_m(l)}{L}}, & f \in \frac{k_m(l)}{LT} \oplus \mathcal{G}_m \\ T\sqrt{L}[\mathbf{C}_m]_{li} e^{j2\pi \frac{c_i \bar{k}_m(l)}{L}}, & f \in \frac{\bar{k}_m(l)}{LT} \oplus \mathcal{G}_m \end{cases}$$

Each of the filters $\Phi_i(f)$, $1 \leq i \leq p$ has a piecewise constant frequency response.

The reconstruction scheme is illustrated in fig. 2. $\Psi_i(z)$ is a digital filter whose impulse resonse is $\{\phi(nT)\}$. The filters used are ideal. We ignore here the issues involved in their approximate realization and concentrate instead on the aliasing errors due to signal mismodelling.

4. ERROR BOUNDS

Suppose the input $x(t) \notin \mathcal{B}(\mathcal{F})$ then the signal $\tilde{x}(t)$ reconstructed using (19) would be in error because the transformation matrix in (18) will not equal the identity matrix in general. For example, this would happen if, in the system design, we choose to ignore certain frequencies that contain negligible signal energy in favor of having a smaller spectral support $\mathcal{F}$. In this section we obtain bounds on the aliasing error $e(t) = \tilde{x}(t) \Leftrightarrow x(t)$ resulting from an underestimation of the spectral support.

In the following analysis we will assume, for simplicity, that $x(t) \in \mathcal{B}([\mathcal{F}])$. In other words we assume that the spectral span $[\mathcal{F}]$ is correctly specified, but the multiband structure to which $x(t)$ is band-limited within $[\mathcal{F}]$ may be misspecified.

4.1. RECONSTRUCTION ERROR

We denote the aliasing error by $e(t) = \tilde{x}(t) \Leftrightarrow x(t)$. Then its spectral components $\mathbf{e}_m^{\pm}(t) = \tilde{\mathbf{x}}_m^{\pm}(t) \Leftrightarrow \mathbf{x}_m^{\pm}(t)$ can be obtained using (18)

$$\begin{bmatrix} \mathbf{e}_m^+(f) \\ \mathbf{e}_m^-(f) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_m^{\text{inv}} \mathbf{B}_m \\ \mathbf{C}_m \mathbf{B}_m \Leftrightarrow \mathbf{I} \end{bmatrix} \mathbf{x}_m^-(f) \tag{20}$$

We define matrices $\mathbf{S}_m$, $\mathbf{D}_m$ and $\mathbf{F}_m$ for each m as follows

$$\mathbf{S}_m = \begin{bmatrix} \mathbf{D}_m \\ \mathbf{F}_m \end{bmatrix} = \begin{bmatrix} \mathbf{A}_m^{\text{inv}} \mathbf{B}_m \\ \mathbf{C}_m \mathbf{B}_m \Leftrightarrow \mathbf{I} \end{bmatrix} \tag{21}$$

We provide the relevant bounds on the sup-norm and the 2-norm of the aliasing error $e(t)$ without going into the details of the proof.

Theorem 2 *Suppose $x(t) \in \mathcal{B}([\mathcal{F}])$ is sampled nonuniformly according to a sampling pattern $\mathcal{C}$ designed for the class of signals $\mathcal{B}(\mathcal{F})$, the resulting aliasing error $e(t)$ satisfies*

$$\max_t |e(t)| \leq \max_m \|\mathbf{S}_m\|_1 \int_{[\mathcal{F}] \setminus \mathcal{F}} |X(f)| df \tag{22}$$

$$[\max_m \left\|\mathbf{S}_m^{\dagger}\right\|_2]^{-1} \sqrt{\mathcal{E}_{\text{out}}} \leq \|e\|_2 \leq \max_m \|\mathbf{S}_m\|_2 \sqrt{\mathcal{E}_{\text{out}}} \tag{23}$$

where $\|\cdot\|_1$ denotes the maximum column sum norm and

$$\mathcal{E}_{\text{out}} = \int_{[\mathcal{F}] \setminus \mathcal{F}} |X(f)|^2 df. \tag{24}$$

The bounds in (23) are sharp but (22) may not be sharp.

4.2. PERFORMANCE IN THE PRESENCE OF NOISE

Finally we consider the effect of additive white sample noise representing, for example, quantization error. The sampled signal can be modelled as $\bar{x}(nT) = x(nT) + w(nT)$ where $w(n)$ is a noise process with $E[w(mT)w(nT)] = \sigma^2 \delta(m \Leftrightarrow n)$ and $x(t)$ is the actual signal we would like to be sampling. The corresponding output noise power for the system is easy to calculate. We find that the average power of the (nonstationary) output noise is

$$\langle E|\tilde{w}(t)|^2 \rangle_t = \sigma^2 T \sum_{m=1}^{M} \lambda(\mathcal{G}_m)(\|\mathbf{A}_m^{\text{inv}}\|_F^2 + \|\mathbf{C}_m\|_F^2) \tag{25}$$

where $\|\cdot\|_F$ denotes the Frobenius norm and $\lambda(\mathcal{G}_m) = (\gamma_{m+1} \Leftrightarrow \gamma_m)$ is the length of $\mathcal{G}_m$.

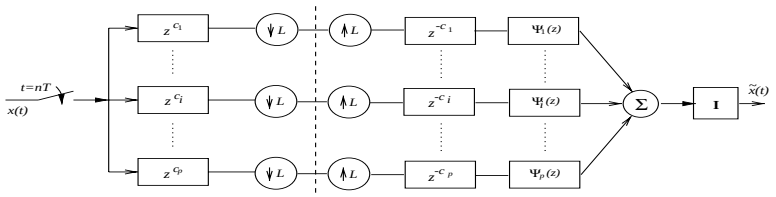

Figure 2: Multicoset sampling and reconstruction. The block "I" is an ideal sinc interpolator.

5. OPTIMIZING THE RECONSTRUCTION PROCESS

We will now examine the problem of optimial reconstruction. For each fixed index m, the multiplying constants in the upper bounds (22) and (23) are related to the 1- and 2-norms, respectively, of the matrix $\mathbf{S}_m$ in (21). We will fix the index m and drop it everywhere from now on, for readability. Given sampling pattern $\mathcal{C}$ and a spectral index set $\mathcal{K}$, our objective is to find matrices $\mathbf{A}^{\text{inv}}$ and $\mathbf{C}$ satisfying (17) that minimize the norm of $\mathbf{S}$ defined in (21), where "norm" means either $\|\cdot\|_2$ or $\|\cdot\|_1$. The other possibility is to minimize the output noise power. Equation (25) reveals that we need to minimize the quantity

$$\|\mathbf{A}^{\text{inv}}\|_F^2 + \|\mathbf{C}\|_F^2 \tag{26}$$

over all valid matrices $\mathbf{A}^{\text{inv}}$ and $\mathbf{C}$.

Note that if $p = q$, the matrix $\mathbf{A}$ is square and hence $\mathbf{A}^{\text{inv}} = \mathbf{A}^{-1}$ and $\mathbf{C} = \mathbf{0}$ are the only valid matrices. Therefore, the reconstruction system only needs to be optimized when $p > q$. Let this be the case in the rest of the section. The other point to note is that the optimization needs to be carried out for each value of the index m (the subscript we have chosen to omit).

5.1. MINIMIZING THE ALIASING ERROR ENERGY

The problem of minimizing the spectral norm of $\mathbf{S}$ defined in (21) admits a simple analytic solution:

Theorem 3 *Let $[\mathbf{A} \mid \mathbf{B}]$ be a $p \times L$ submatrix of the $L \times L$ DFT matrix $\mathbf{W}$, $[\mathbf{W}]_{mn} = \frac{1}{\sqrt{L}} \exp(\frac{j2\pi}{L}mn)$, with possible rearrangements of columns. Suppose $\mathbf{A}$ has full column rank $q \leq p$. Then the minimization*

$$\min \|\mathbf{S}\|_2 \equiv \min \left\| \left[\frac{\mathbf{A}^{\text{inv}}\mathbf{B}}{\mathbf{CB} \Leftrightarrow \mathbf{I}} \right] \right\|_2 \tag{27}$$

performed over all matrices $\mathbf{A}^{\text{inv}}$ and $\mathbf{C}$ that satisfy $\mathbf{A}^{\text{inv}}\mathbf{A} = \mathbf{I}$ and $\mathbf{CA} = \mathbf{0}$, has the solution

$$\mathbf{A}_\circ^{\text{inv}} = \mathbf{A}^\dagger = (\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^* \tag{28}$$

$$\mathbf{C}_\circ = \mathbf{B}^*(\mathbf{I} \Leftrightarrow \mathbf{A}\mathbf{A}^\dagger) \tag{29}$$

and the corresponding minumum value is

$$\|\mathbf{S}_\circ\|_2 = \begin{cases} \sqrt{\lambda_{\max}((\mathbf{A}^*\mathbf{A})^{-1})}, & p < L \\ 0, & p = L \end{cases} \tag{30}$$

5.2. MINIMIZING THE OUTPUT NOISE POWER

We seek the optimal matrices $\mathbf{A}^{\text{inv}}$ and $\mathbf{C}$ that satisfy $\mathbf{A}^{\text{inv}}\mathbf{A} = \mathbf{I}$ and $\mathbf{CA} = \mathbf{0}$, in order to minimize the quantity (26)

$$\min_{\mathbf{A}^{\text{inv}},\mathbf{C}} (\|\mathbf{A}^{\text{inv}}\|_F^2 + \|\mathbf{C}\|_F^2) \tag{31}$$

The solution $\mathbf{C}_\circ = \mathbf{0}$ and $\mathbf{A}_\circ^{\text{inv}} = \mathbf{A}^\dagger$ is straightfowrard to verify. Further the minimum value of the objective function (31) is

$$\|\mathbf{A}^\dagger\|_F = \sqrt{\text{tr}((\mathbf{A}^*\mathbf{A})^{-1})} \tag{32}$$

5.3. MINIMIZING THE PEAK ALIASING ERROR

The relevant quantity to minimize in order to obtain the best bound in (22) is the 1-norm of the matrix $\mathbf{S}$ defined in (21). The problem of choosing valid matrices $\mathbf{A}^{\text{inv}}$ and $\mathbf{C}$ in order to minimize $\|\mathbf{S}\|_1$, unlike the spectral or Frobenius norms of $\mathbf{S}$, cannot be solved analytically. We resort to numerical methods instead.

$$\min_{\mathbf{A}^{\text{inv}},\mathbf{C}} \left\| \left[\frac{\mathbf{A}^{\text{inv}}\mathbf{B}}{\mathbf{CB} \Leftrightarrow \mathbf{I}} \right] \right\|_1 \tag{33}$$

We eliminate the constraint (17) by parametrizing $\mathbf{A}^{\text{inv}}$ and $\mathbf{C}$ as $\mathbf{A}^{\text{inv}} = \mathbf{A}^\dagger + \mathbf{X}_1\mathbf{P}$ and $\mathbf{C} = \mathbf{X}_2\mathbf{P}$ where $\mathbf{P} = (\mathbf{I} \Leftrightarrow \mathbf{A}\mathbf{A}^\dagger)$ is the least-squares projection operator onto the null space of $\mathbf{A}^*$. The matrix $\mathbf{X}_1$ is a $q \times p$ and $\mathbf{X}_2$, $(L \Leftrightarrow q) \times p$. We can now rewrite (33) in an unconstrained form as

$$\min_{\mathbf{X}} \|\mathbf{S}_0 + \mathbf{XF}\|_1 \tag{34}$$

where $\mathbf{S}_0 = \left[\frac{\mathbf{A}^\dagger\mathbf{B}}{-\mathbf{I}} \right]$, $\mathbf{F} = \mathbf{PB}$ and $\mathbf{X} = \left[\frac{\mathbf{X}_1}{\mathbf{X}_2} \right]$.

It turns out that the optimization problem (34) is hard to solve even numerically, even though it involves the minimization of a convex function. Solving it directly is difficult because the objective function is nondifferentiable. However it *does* lend itself to an approximate linear program formulation. In practice, an approximate solution is generally sufficient.

6. REFERENCES

[1] K. Cheung and R. Marks, "Image sampling below the Nyquist density without aliasing," *J. Opt. Soc. Am. A*, vol. 7, pp. 92–105, 1990.

[2] R. E. Kahn and B. Liu, "Sampling representations and the optimum reconstruction of signals," *IEEE Trans. Info. Theory*, vol. 11, pp. 339–347, 1965.

[3] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proc. IEEE Int. Conf. Acoust. Speech, Sig. Proc.*, (Atlanta, GA), IEEE, May 1996.

[4] C. Herley and P. W. Wong, "Minimum rate sampling of signals with arbitrary frequency support," in *Proc. IEEE Int. Conf. Image Proc.*, (Lausanne, Switzerland), IEEE, Sept. 1996