The 35th Asilomar Conference on Signals, Systems, and Computers
Pacific Grove, CA

Computability Constraints in Space-Time Delta-Sigma Arrays

Dan P. Scholnik scholnik@nrl.navy.mil
Jeffrey O. Coleman jeffc@radar.nrl.navy.mil
Naval Research Laboratory, Washington, DC

ABSTRACT

Modern array-based RF systems require linear amplification to preserve the spectral integrity of simultaneously transmitted signals. Recently delta-sigma modulation jointly in space and time has been proposed to provide this linearity when paired with an emerging class of high-power switches. Since delta-sigma modulation involves a feedback loop, computability requirements restrict the class of possible loop filters. We examine several different computability restrictions and their effect on system performance in terms of quantization noise suppression and parallelization.

1. INTRODUCTION

Recently, joint $\Delta\Sigma$ modulation in space and time has been proposed for antenna arrays and halftoned movies [1]. In the array application joint $\Delta\Sigma$ modulation promises high linearity when coupled with a new class of high-powered RF switches [2]. Although the resulting data streams are binary, the inevitable quantization noise is pushed out of the spatio-temporal frequency band of interest, where it is removed by analog filters and the natural spatial-lowpass nature of far-field propagation [3].

For real time implementations, it is often imperative that the modulator computations be parallelized to minimize the required clock rate. This has serious implications for the loop filtering in the modulator, as the feedback loop represents the critical delay in the system. Specifically, the support of the impulse response of the loop filter is subject to some *computability constraint* so that the overall system is consistent. There are several possible constraints, some of which parallelize well. The constraint also affects the possible filter transfer functions, which in turn determine the modulator performance. Unsurprisingly, the constraints that result in the easiest parallelization have the poorest performance, and vice-versa. In this paper we examine several possible classes of filters and their impact on modulator performance and implementation.

2. SYSTEM ANALYSIS

Figure 1 shows a simplified version of the vector $\Delta\Sigma$ modulator introduced in [1]. The scalar input $s(n)$ feeds

This work was supported by the Office of Naval Research.

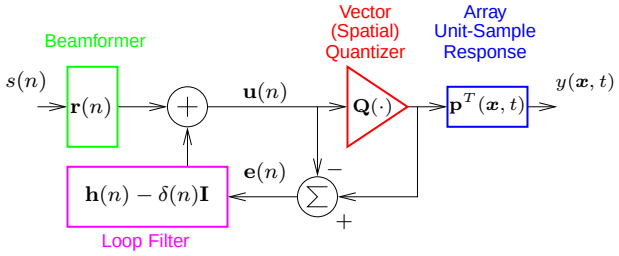


Figure 1: A space-time vector $\Delta\Sigma$-modulated array.

a $K \times 1$ beamformer vector $\mathbf{r}(n)$ that computes the vector modulator input. The modulator output represents a quantized (usually binary-valued) vector of inputs to the elements of an antenna array. One polarization component of the array's effective unit-sample response is given by vector $\mathbf{p}(\boldsymbol{x},t)$. The array output $y(\boldsymbol{x},t)$ is the antenna current density as a function of space and time. The error $\mathbf{e}(n)$ between vector quantizer input and output is fed back through the loop filter, which has a $K \times K$ matrix impulse response. The basic input/output relationship of the array before and after Fourier transforming on t is

$$\begin{aligned} y(\boldsymbol{x},t) &= \big(\mathbf{p}^T \star \mathbf{r} * s\big)(\boldsymbol{x},t) \\ &\quad + \big(\mathbf{p}^T \star \mathbf{h} * \mathbf{e}\big)(\boldsymbol{x},t) \end{aligned} \tag{1}$$

$$\begin{aligned} Y(\boldsymbol{x},f) &= \mathbf{P}^T(\boldsymbol{x},f)T\,\mathbf{R}(fT)\,S(fT) \\ &\quad + \mathbf{P}^T(\boldsymbol{x},f)T\,\mathbf{H}(fT)\,\mathbf{E}(fT), \end{aligned} \tag{2}$$

where $*$ indicates discrete-time convolution and defining a temporal convolution in continuous/discrete-time by

$$(\mathbf{p} \star \mathbf{w})(\boldsymbol{x},t) \triangleq T \sum_k \mathbf{p}(\boldsymbol{x}, t-kT)\,\mathbf{w}(k).$$

On the right in (1) and (2) the first and second terms are the desired beamformed signal and the shaped quantization error respectively, each shaped by the antenna impulse response. Beamformer design is the subject of [3] and will not be considered further here, since it is not in the feedback loop and does not pose as severe a computational problem. The shaped noise term is of interest, and so we now consider the far-field response of the array.

2.1 Far-Field Propagation

The spatial and temporal output of the array is the current density $y(\boldsymbol{x},t)$. We are interested in the far-field array

response, which can be found by solving Maxwell's equations. We use the method of [4], first solving for one polarization component $A(\boldsymbol{x}, f)$ of the vector potential, and then finding the electric field $E(\boldsymbol{x}, f)$ in the far field. The vector potential component represents a spherical wave propagating away from each point on the antenna at a speed c:

$$A(\boldsymbol{x}, f) = \frac{\mu}{4\pi} \int_{\mathcal{V}} Y(\boldsymbol{x}', f) \frac{e^{-j2\pi\|\boldsymbol{x}'-\boldsymbol{x}\|f/c}}{\|\boldsymbol{x}'-\boldsymbol{x}\|}\, d\boldsymbol{x}', \quad (3)$$

where $\mathcal{V}$ is the spatial support of the current density $Y(\boldsymbol{x}, f)$. In the far field, defined here simply as $\|\boldsymbol{x}\| \to \infty$, the electric field can be found as

$$E(\boldsymbol{x}, f) \approx -j2\pi f A(\boldsymbol{x}, f) \quad (4)$$

Combining (3) and (4) and letting direction vector $\boldsymbol{u} = \boldsymbol{x}/\|\boldsymbol{x}\|$, the electric field can be reduced to

$$E(\boldsymbol{x}, f) \approx \frac{-j\mu}{2\|\boldsymbol{x}\|} e^{-j2\pi\|\boldsymbol{x}\|f/c} f \int_{\mathcal{V}} Y(\boldsymbol{x}', f) e^{j2\pi\boldsymbol{x}'^T\boldsymbol{u}f/c}\, d\boldsymbol{v}'.$$

where we use the far-field approximation $\|\boldsymbol{x}' - \boldsymbol{x}\| \approx \|\boldsymbol{x}\|$ in the denominator (where the approximation is less critical) and the more precise approximation $\|\boldsymbol{x}'-\boldsymbol{x}\| \approx \|\boldsymbol{x}\|-\boldsymbol{x}'^T\boldsymbol{u}$ in the exponent. Recognizing a spatial Fourier transform and dividing out factors dependent on the actual distance $\|\boldsymbol{x}\|$ yields the *normalized far-field* electric field

$$\frac{j2\|\boldsymbol{x}\|}{\mu} e^{j2\pi\|\boldsymbol{x}\|f/c} E(\boldsymbol{x}, f) \approx f\,\hat{Y}(-\boldsymbol{u}f/c, f). \quad (5)$$

Letting $\boldsymbol{v}$ denote spatial frequency, then, $\boldsymbol{v} = -\boldsymbol{u}f/c$ represents the frequencies corresponding to propagating radiation. Taking norms on both sides gives a version of the Helmholtz equation, $\|\boldsymbol{v}\| = |f|/c$. This defines the cone of propagating frequencies in $(\boldsymbol{v}, f)$ space. Spatial frequencies not on the cone do not propagate into the far field, and in this way the array acts as a particular ideal bandpass spatial filter. If we consider only a single spatial dimension x_z and a single corresponding spatial frequency v_z (appropriate for array elements on a line), then the projection of the positive-frequency half of the cone into the (v_z, f) plane shown in Fig. 2 suggests how to choose the noise shaping in a space-time $\Delta\Sigma$ modulator. The center trapezoid represents the intersection of the temporal frequency band of interest and the cone, and represents the spectral support of the desired signal where quantization noise should be suppressed. The surrounding region then holds the shaped quantization errors. The rectangle represents the positive-frequency half of one period of the frequency response of the loop filter when the sampling rate is $1/T$ and the element spacing is d.

Now the far-field response of the array can be written as

$$\begin{aligned} f\,\hat{Y}(\boldsymbol{v}, f) &= f\,\hat{\mathbf{P}}^T(\boldsymbol{v}, f) T\,\mathbf{R}(fT) S(fT) \\ &\quad + f\,\hat{\mathbf{P}}^T(\boldsymbol{v}, f) T\,\mathbf{H}(fT)\mathbf{E}(fT). \end{aligned} \quad (6)$$

A convenient simplification results if the antenna elements are identical, so that $p_k(\boldsymbol{x}, t) = p_0(\boldsymbol{x} - \mathbf{x}_k, t)$ where $\mathbf{x}_k$ is

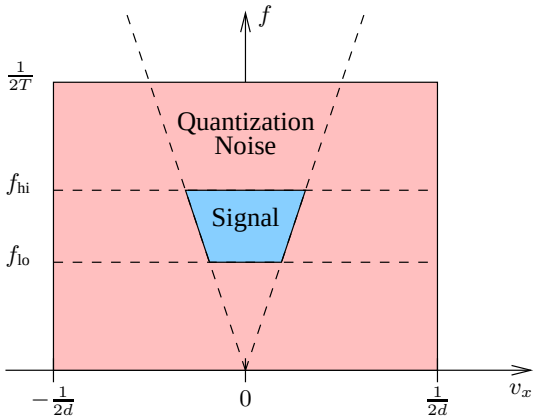


Figure 2: Spatio-temporal spectral regions.

the location of the kth element. Rewriting (6) yields

$$\begin{aligned} f\,\hat{Y}(\boldsymbol{v}, f) &= f\,\hat{P}_0(\boldsymbol{v}, f) T \sum_k R_k(fT) e^{-j2\pi\mathbf{x}_k^T\boldsymbol{v}} S(fT) \\ &\quad + f\,\hat{P}_0(\boldsymbol{v}, f) T \sum_m \sum_k H_{k,m}(fT) e^{-j2\pi\mathbf{x}_k^T\boldsymbol{v}} E_m(fT), \end{aligned} \quad (7)$$

and recognizing the spatial Fourier transforms and rewriting in terms of *array factor* $\hat{R}(\boldsymbol{v}, fT)$ and loop filter column frequency responses $\{\hat{H}_m(\boldsymbol{v}, fT)\}$ yields

$$\begin{aligned} f\,\hat{Y}(\boldsymbol{v}, f) &= f\,\hat{P}_0(\boldsymbol{v}, f) T \hat{R}(\boldsymbol{v}, fT) S(fT) \\ &\quad + f\,\hat{P}_0(\boldsymbol{v}, f) T \sum_m \hat{H}_m(\boldsymbol{v}, fT) E_m(fT). \end{aligned} \quad (8)$$

Each element of the error vector is shaped by the frequency response of the corresponding column of the loop filter.

2.2 Optimizing the Loop Filter

The basic goal of the loop-filter design is to minimize the quantization error power in the spatio-temporal passband—that is, propagating spatial frequencies in the temporal frequency band of interest. Further, some constraints are required to maintain stability of the modulator. The design is complicated somewhat by the fact that the loop filter, while time invariant, is not space invariant. Thus, there is not a single spatio-temporal transfer function to be optimized in the usual sense. However, under the common assumption that power spectral density $\mathbf{S}_e(fT) = \sigma^2\mathbf{I}$, so that the error vector is uncorrelated in both time and space, we can find the far-field in-band noise power in terms of the component error shaping transfer functions $\{\hat{H}_k(\boldsymbol{v}, fT)\}$. For simplicity we consider identical elements as before. Now the power in the combined array output is just the sum of the powers due to each element of the error vector, found by integrating the squared far-field pattern over propagating

frequencies and the signal band:

$$P_\epsilon = \sigma^2 \int_{\mathcal{F}} \int_{\Omega} |fT\hat{P}_0(\boldsymbol{v}, f)|^2 \sum_k |\hat{H}_k(\boldsymbol{v}, fT)|^2 \, dA(\boldsymbol{u}) \, df, \tag{9}$$

where $\mathcal{F}$ denotes the set of in-band temporal frequencies, Ω is the unit sphere, and dA is the surface-area measure. This expression is quadratic in the loop filter coefficients $[\mathbf{h}(n)]_{k,m}$ and can be minimized using a number of optimization techniques. The transfer functions $\{|\hat{H}_k(\boldsymbol{v}, fT)\}$ can be optimized independently since they come from independent columns of the loop filter, a smaller problem in general than optimizing the whole loop filter at once. Stability can be constrained by limiting the maximum gain of each column transfer function.

3. COMPUTABILITY CONSTRAINTS

Having shown how to optimize the loop filter given its structure, we now turn to the topic of *computability* of the system. For our purposes computability means that for any input signal to the modulator there exists a unique, consistent output signal. As an example of a system that is not computable, consider that $[\mathbf{h}(0)]_{k,k} \neq 0$. Now the kth element of the filter output depends on the kth input at the current time, but because of the feedback we have a circular dependency. In temporally-based scalar feedback systems computation proceeds along the time axis and computability equals *strict causality* in the feedback response—that is, the feedback filter output can only depend on past inputs, not current or future inputs. In our vector $\Delta\Sigma$ system, however, computation need not proceed strictly along the time dimension, as we assume our vector quantizer is a simple scalar quantizer in disguise and operates on each vector element independently. So computability in general ensures that there exists some order in which the elements of the output vector can be computed, but that order will depend on the structure of the loop filter impulse response.

In this section we consider several structural constraints on $\mathbf{h}(n)$ that serve to ensure computability. We assume FIR responses throughout. To illustrate the middle column of an example 5×5 matrix $\mathbf{h}(n)$ is shown in Fig. 3 for the several different computability constraints to be considered. In the figure 0 represents a zero element, X represents a nonzero element, and $\bullet$ represents the main diagonal at the time origin where $[\mathbf{h}]_{k,k}(0) = 1$ always. The measure of "goodness" for these constraints comes down to performance and parallelization. In the examples we examine how different constraints affect the ability to minimize the in-band quantization noise in the modulator. In this section we consider the implications for a parallel implementation of the filter. Since the feedback loop will generally be the critical delay in a real-time implementation of a $\Delta\Sigma$ modulator, we seek a parallel implementation that can compute an entire vector of inputs at once, rather than sequentially.

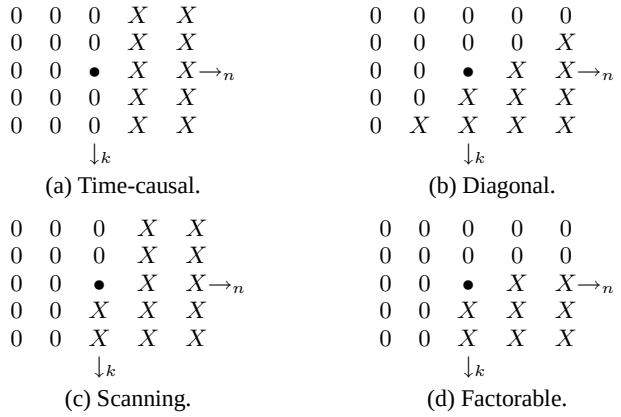


Figure 3: Possible loop filter computability restrictions.

3.1 Strict Causality

The most straightforward extension of the causality requirement in scalar $\Delta\Sigma$ modulators to our case is the *strict causality* constraint

$$\mathbf{h}(n) - \delta(n)\mathbf{I} = \mathbf{0}, \qquad n \leq 0, \tag{10}$$

so that at each time instant the filter output depends only on previous values of the filter input for all vector indices. A representative column of $\mathbf{h}(n)$ is shown in Fig. 3(a). This represents the best structural case for implementation, as parallelization of the filter computation is straightforward, and the quantization of all vector elements can be done simultaneously. As seen in the examples, however, it results in poor performance, and performs scarcely better than individual $\Delta\Sigma$ modulators operating on each element of the input with no spatial noise shaping at all.

3.2 Diagonal Computability

We define a *diagonally computable* loop filter impulse response by

$$\big[\mathbf{h}(n) - \delta(n)\mathbf{I}\big]_{k,m} = 0, \qquad n + k - m \leq 0, \tag{11}$$

a representative column of which is shown in Fig. 3(b). This structure forces computation to proceed along a diagonal path through space and time; to compute all the outputs in space at a given time requires that some outputs in the future first be calculated. Although this would seem to pose a problem in implementation, in fact diagonal computability and strict causality are related through a unitary transformation. Defining diagonal delay matrix

$$\mathbf{d}(n) \triangleq \mathrm{diag}\{\delta(n), \delta(n-1), \ldots, \delta(n-K+1)\}$$

with inverse

$$\mathbf{d}^{-1}(n) = \mathrm{diag}\{\delta(n), \delta(n+1), \ldots, \delta(n+K-1)\}$$

we can see that if $\mathbf{h}(n)$ is diagonally computable, then $(\mathbf{d} * \mathbf{h} * \mathbf{d}^{-1})(n)$ is strictly causal. Likewise, if $\mathbf{h}(n)$ is

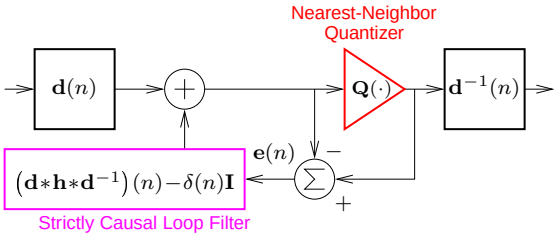


Figure 4: Transformation of diagonal loop filter to strictly causal.

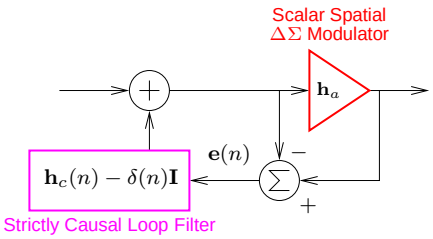


Figure 5: Decomposition of scanning loop filter

strictly causal, then $(\mathbf{d}^{-1} * \mathbf{h} * \mathbf{d})(n)$ is diagonally computable, an equivalent definition to (11). This relation can be used to transform the system of Fig. 1 with a diagonally computable loop filter to the equivalent system of Fig. 4 which has a strictly causal loop filter. The delay matrix at the input skews the elements of the input signal through progressive delays, so that the vector traveling around the loop at any instant n is a diagonal cut through the input signal. At the output of the quantizer the signal is realigned by the inverse delay matrix. To make this last step causal in an implementation an additional delay of $K-1$ would be required, but this modest additional system delay is the only penalty incurred relative to a strictly causal system.

3.3 Scanning Computability

We define *scanning computability* as

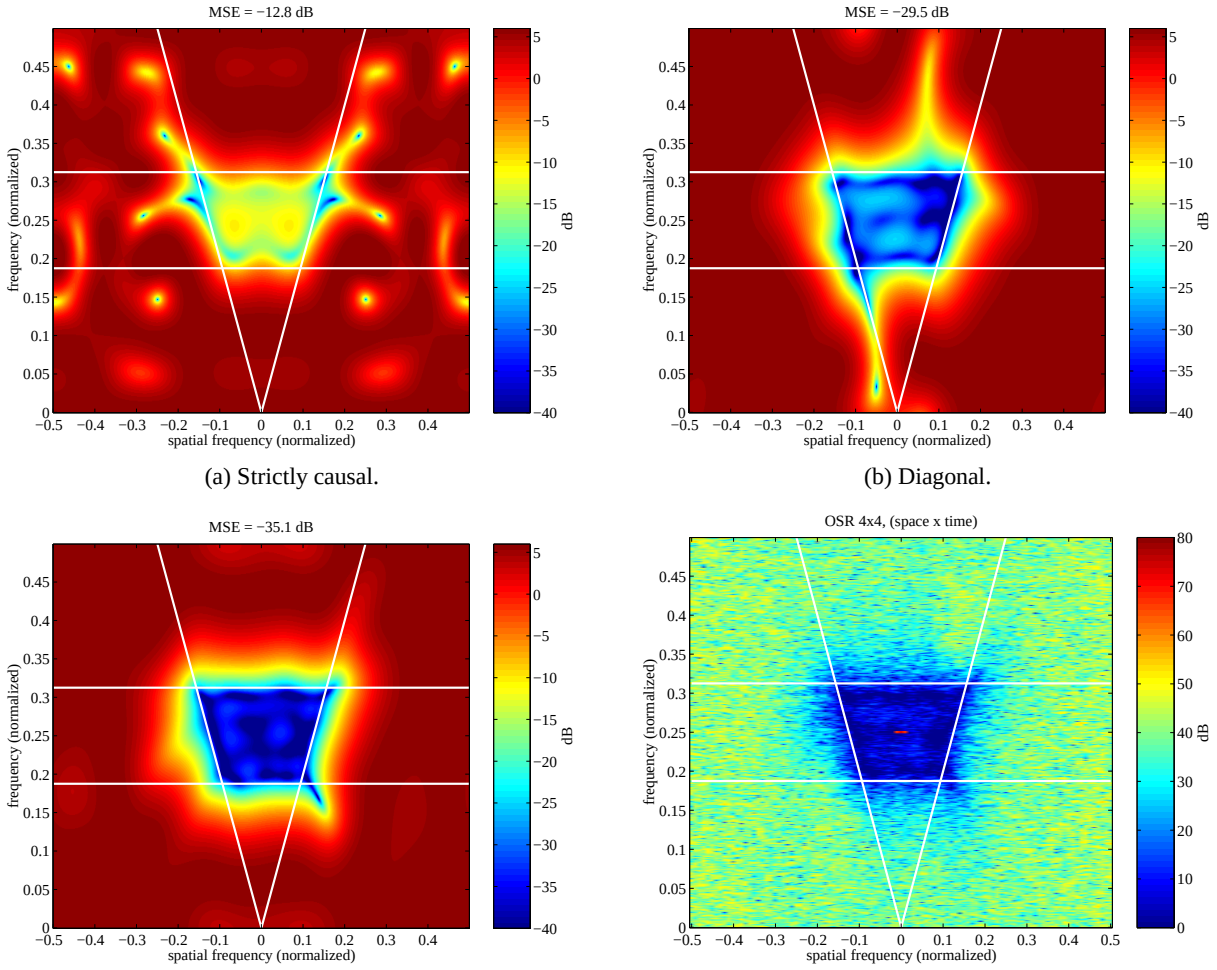
$$\begin{aligned} \mathbf{h}(n) &= \mathbf{0}, \qquad n < 0 \\ \big[\mathbf{h}(0) - \mathbf{I}\big]_{k,m} &= 0, \qquad k \le m, \end{aligned} \tag{12}$$

or in words the impulse response is causal (but not strictly so) and at time $n = 0$ the impulse response matrix is lower triangular. A representative column of such an impulse response is shown in Fig. 3(c). This is in some sense the most natural extension of causality to multiple dimensions and is commonly used for image halftoning [5]. The lower triangular structure at $n = 0$ makes the output at a given spatial index and time dependent on the output at smaller indices at the same time. This spatial dependence dictates sequential computation across vector elements at each time step, a real barrier to parallelization. In general there exists no invertible linear transformation to a strictly causal impulse response (although for some special cases some parallelism may be achieved this way). However, we can decompose a scanning-computable impulse response into its strictly causal and non-strictly causal parts, $\mathbf{h}(n) = \mathbf{h}_c(n) + \mathbf{h}_a\delta(n)$, where $\mathbf{h}_a \triangleq \mathbf{h}(0) - \mathbf{I}$. Now the non-strictly causal response can be moved into the definition of the vector quantizer. No longer a simple nearest neighbor quantizer, it becomes a spatial $\Delta\Sigma$ modulator (as for image halftoning) with loop filter matrix $\mathbf{h}_a$ as shown in Fig. 5. While now the quantizer is sequential and potentially hard to parallelize, the remaining loop filter (which in most cases contains the bulk of the computations) is strictly causal.

A special case of scanning computability arises when the matrix impulse response can be factored into a causal scalar function of n and a lower-triangular matrix: $\mathbf{h}(n) = h_c(n)\mathbf{h}_a$, where $h_c(0) = 1$ and $\mathbf{h}_a$ has ones down the main diagonal. A representative column of such an impulse response is shown in Fig. 3(d). In this case the loop filter impulse response of Fig. 5 can be written with $\mathbf{h}_c(n) = h_c(n)\mathbf{I}$ and instead of a general matrix-vector convolution each element on $\mathbf{e}(n)$ is separately filtered by $h_c(n)$. The result is effectively a scalar spatial $\Delta\Sigma$ modulator operating inside a scalar temporal $\Delta\Sigma$ modulator, which can result in substantial computational savings, although performance will necessarily suffer as each modulator also suppresses noise outside the signal band. Such an approach was used in [6].

4. EXAMPLES

To compare the effects of the various computability restrictions, strict causality, diagonal computability, and scanning computability were imposed on the design of the loop filter for an example modulator. The example array has 101 identical elements uniformly spaced in a line. For a sampling rate of $f_s = 1/T_s$ we wish to transmit over a bandwidth of $f_s/8$ centered at $f_s/4$. The element spacing is $d = cT_s/2$, or half the wavelength at the sampling frequency. These parameters provide four-times oversampling in both space and time. The nonzero elements of both the scanning and causal impulse response matrices occupy a diagonal band of width 11 and are 11 taps long in time. The diagonal computability impulse response has a somewhat more irregular geometry, but approximately the same number of nonzero coefficients. In all cases the geometry of most of the columns are identical except for a spatial shift, which has no effect on the optimization. Thus the center portion of the optimized response matrices is Toeplitz, and for simplicity only a center column of each was optimized and replicated to form a (slightly suboptimal) Toeplitz matrix. Second-order cone programming was used [7] to minimize the filter's mean-square frequency response in the signal band, while limiting the maximum filter gain outside the band to 6 dB for modulator stability. The resulting optimized frequency responses are shown along with their mean-square signal-band values in Fig. 6. As expected, the best response results from the scanning case. The diagonal impulse response is 6 dB worse but offers a parallel implementation. Restricting the impulse response to strict temporal causality results in rather poor performance. In fact, experimentation shows that the spatial extent of a strictly causal filter has little effect on the performance, suggesting



(a) Strictly causal.

(b) Diagonal.

(c) Scanning.

(d) Example scanning output spectrum.

Figure 6: Optimized loop filter magnitude responses with various computability constraints.

that a one-dimensional time-only filter would do as well.

To show a typical modulator output spectrum, the scanning impulse response of Fig. 6(c) was used to modulate a sinusoidal, Hanning-weighted beam directed normal to the array. The resulting spectrum out of the modulator output prior to the array is shown in Fig. 6(d). The red blip is the desired signal, surrounded by shaped quantization noise. The similarity between this plot and Fig. 6(c) serves to validate our assumption of white, uncorrelated errors.

REFERENCES

[1] D. P. Scholnik and J. O. Coleman, "Joint spatial and temporal delta-sigma modulation for wideband antenna arrays and video halftoning," in *Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Processing*, Salt Lake City, UT, May 2001.

[2] Arun Jayaraman, P. F. Chen, G. Hanington, L. Larson, and P. Asbeck, "Linear high-efficiency microwave power amplifiers using bandpass delta-sigma modulators," *IEEE Microwave and Guided Wave Letters*, vol. 8, no. 3, Mar. 1998.

[3] D. P. Scholnik and J. O. Coleman, "Superdirectivity and SNR constraints in wideband array-pattern design," in *IEEE Int'l Radar Conference*, Atlanta, GA, May 2001.

[4] C. A. Balanis, *Antenna Theory*, John Wiley & Sons, Inc., 1982.

[5] T. D. Kite, B. L. Evans, and A. C. Bovik, "Modeling and quality assessment of halftoning by error diffusion," *IEEE Trans. Image Processing*, vol. 9, no. 5, pp. 909–922, May 2000.

[6] Y. Tamura, N. Kawakami, O. Akasaka, M. Okada, and K. Koyama, "Beam-forming using multidimensional sigma-delta modulation," in *Proc. IEEE Ultrasonics Symp.*, 1998, pp. 1077–1080.

[7] D. P. Scholnik and J. O. Coleman, "Formulating wideband array-pattern optimizations," in *Proc. IEEE Int'l Symp. Phased Array Systems and Technology (ISPAST 2000)*, Dana Point, CA, May 2000.