

Slicer Architectures for Analog-to-Information Conversion in Channel Equalizers

Aseem Wadhwa, Upamanyu Madhow, and Naresh R. Shanbhag

Abstract—The scaling of analog-to-digital converter (ADC) power consumption with communication bandwidth imposes severe limits on its precision, which significantly impacts receiver performance. In this paper, we consider a “space-time” generalization of the flash architecture by allowing a fixed number of slicers to be dispersed in time (i.e., sampling offset) as well as space (i.e., amplitude), with the goal of investigating its capabilities for *analog-to-information conversion* (i.e., enabling reliable recovery of digital information, rather than faithful reproduction of the input signal) in the context of channel equalization for binary signaling over a dispersive channel. We first study standard symbol-spaced ADC with severe quantization constraints, estimating the minimum number of slicers needed to avoid error floors. We observe that the performance is sensitive to channel realization and sampling phase, which motivates a more flexible space-time architecture. Using ideas similar to those underlying compressive sensing, we prove that such architectures have no fundamental limitations in theory: randomly dispersing enough one-bit slicers over space and time does provide information sufficient for reliable equalization. We then focus on practical designs for symbol-spaced and fractionally-spaced sampling subject to a constraint on the number of slicers, and propose an algorithm for optimizing slicer thresholds, which significantly improves performance over a standard design.

Index Terms—Equalization, analog to digital converter, quantization, high-speed link, fractionally spaced sampling.

I. INTRODUCTION

A CRUCIAL component of a modern communication receiver is the analog-to-digital converter (ADC), which translates the analog received waveform into a digital signal, enabling implementation of sophisticated receiver algorithms using digital signal processing. As signal bandwidths scale up to multiples of GHz, however, the cost and power consumption of high-resolution ADCs become prohibitive [1]. Once the

number of bits per sample becomes constrained by such considerations, it becomes natural to consider alternatives to the general-purpose ADC that are tailored to the communications application. Thus, we are interested in the design of *analog-to-information converters* enabling reliable recovery of the transmitted data, rather than accurate reproduction of the received signal as for a standard ADC. In this paper, we investigate this approach in the context of channel equalization, with the goal of exploring the fundamental limits imposed by ADC precision constraints.

Our starting point is the flash ADC, a popular architecture for high sampling rates (GHz) and relatively low resolutions (2-6 bits) (for instance, see [2], [3] for some recent high-speed flash ADC designs). Other architectures such as successive approximation and pipelined are also built upon low resolution flash ADCs. The standard flash architecture comprises of several parallel 1 bit slicers. An n -bit flash ADC consists of $2^n - 1$ comparators sampling *synchronously*, with slicer thresholds spread uniformly over the input signal voltage range. Fractional sampling is known to be more robust than symbol-spaced sampling for systems in which ADC resolution is not an issue. However, in the regimes of interest to us, the Nyquist sampling rate is already stressing the state of the art, hence the conventional approach is to sample at the Nyquist rate. If the objective is signal reconstruction, high precision quantization is desired (typically $n > 10$), which is prohibitive at high speeds due to the large number of slicers required. However for equalization much less precision is enough for preserving information to invert the channel and ensure error free reception at high SNR. Key questions we address in this paper are characterizing the minimum precision required for *information preservation* and whether, for a fixed number of comparators, we can do better by generalizing beyond uniform spacing and synchronous sampling of a standard flash ADC. We explore designs which are “space-time” generalization of the flash architecture, in which all slicers sample at the Nyquist rate, but can have different sampling times and non-uniform thresholds over the input dynamic range. Depending on the values of the sampling offset, we either get relatively high precision Nyquist rate samples (same offsets) or lower precision faster than Nyquist rate samples (different offsets).

As a first step in obtaining initial insights, we restrict attention to the simplest possible setting of binary antipodal signaling over a time-invariant, real baseband channel with relatively small delay spread. High-speed backplane wireline channels provide an interesting set of running examples that conform to this model: the impulse responses are fairly stationary, and are typically 5-10 symbol periods long. Backplane

Manuscript received April 27, 2016; revised August 28, 2016 and November 12, 2016; accepted December 2, 2016. Date of publication December 19, 2016; date of current version March 15, 2017. This work was supported in part by the Systems On Nanoscale Information fabriCs (SONIC), one of six centers supported by the STARnet phase of the Focus Center Research Program (FCRP), a Semiconductor Research Corporation Program sponsored by MARCO and DARPA. The associate editor coordinating the review of this paper and approving it for publication was H. Minn.

A. Wadhwa was with the Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106 USA. He is now with Apple, Cupertino, CA 95104 USA (e-mail: aseem@ece.ucsb.edu).

U. Madhow is with the Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106 USA (e-mail: madhow@ece.ucsb.edu).

N. Shanbhag is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, IL 61801 USA (e-mail: shanbhag@illinois.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCOMM.2016.2641445

communication generally employs uncoded binary antipodal modulation, in moderate to high SNR regimes which are interference limited.

The objective of this work is to explore system level design. Hence we employ a standard Gaussian noise model, ignoring issues such as clock jitter, noise in slicer thresholds and other circuit impairments which result in other sources of noise when a complete bottom-up system is considered.

Contributions and Outline: We summarize our contributions, and the organization of the paper, as follows.

We discuss related work in section II and the system model in section III.

(1) In section IV we first consider the standard Nyquist synchronously sampled, uniformly spaced design. For a given channel, we derive easily computable lower and upper bounds for the smallest number of comparators to avoid an error floor in the bit error rate (BER). The results give insight into the kind of channels that are worse in terms of requiring a larger number of comparators; for example, mixed-phase channels are worse than minimum/maximum phase channels. We also demonstrate via an example how, for the standard design, the BER can be sensitive to the sampling phase, and that more robust performance can be obtained by spreading the same number of slicers across time. This motivates a more systematic study of space-time architectures.

(2) In Section V, we establish that, in theory there are no fundamental performance limitations imposed by spreading slicers out in space and time, by proving that the ℓ_1 distance between a pair of waveforms is preserved upon quantization by n slicers spread across time and having randomly distributed thresholds, if n is larger than a lower bound. The proof of this general result employs the Chernoff bound and the union bound, analogous to the Johnson-Lindenstrauss (JL) lemma [4]. Its application to our equalization problem guarantees the absence of an error floor if sufficiently many 1-bit measurements are obtained with random thresholds. While this result provides a sound theoretical underpinning for space-time slicer architectures, in practice, good performance is obtained with fewer slicers with carefully chosen thresholds.

(3) We present (in Section VI) an approximate optimization technique for adapting, as a function of the channel, the slicer thresholds for symbol-spaced and fractionally-spaced (at $T_s/2$, where T_s denotes the symbol interval) architectures. For a fixed number of slicers, the performance gains over a standard symbol-spaced uniform ADC are significant. Depending on the choice of channel, sampling phase and number of available slicers, the procedure allocates all slicers to one sampling phase or distributes them among the two phases.

Our conclusions are presented in Section VII.

Caveats: While the results discussed in this paper are, in principle, directly applicable to backplane communication, some caveats are in order. First, since our focus is on understanding the impact of analog-to-information conversion, we do not restrict the complexity of the digital backend. Further simplifications to the digital processing might be required for practical implementation. Second, in order to derive systems-level insights, we make the simplifying assumption of equating power consumption to the number of slicers employed

in the ADC, and strive to minimize that in our design. However, the exact power consumption will depend on many specific implementation details that are beyond our current scope. We also note that, while our formulation provides insight into information conversion for dispersive channels under stringent quantization constraints, extensions to more complicated scenarios such as higher constellations, OFDM modulation and time varying channels are bound to present a richer set of challenges, which are again beyond our present scope.

II. RELATED WORK

It is known that Nyquist sampling, even for strictly band-limited inputs, is not optimal for finite precision measurements. For example, Gilbert [5], Shamai [6] have shown that the capacity of bandlimited systems with 1-bit measurements increases as we sample faster than the Nyquist rate. A related result is discussed by Kumar *et al.* [7]. The effect of heavily quantized measurements on communication systems design and performance has received significant attention recently. For non-dispersive channels, the effect of coarse quantization has been studied for the ideal AWGN channel [8] and carrier-asynchronous systems [9], [10]. Reference [11] discusses channel estimation with coarsely quantized samples.

The analog-to-information (AIC) converter discussed in [12] is “compressive”: it consists of a front-end that applies a known pseudo-random noise to the input, followed by a sub-Nyquist sampling rate ADC. Such a design preserves the information in the input signal if it is *frequency sparse*, and is therefore of interest for applications such as cognitive radio. In such settings, the sparser the signal, the more the sampling frequency can be reduced, which potentially leads to power savings. Analog to information conversion as discussed in this paper has a different context: the focus here is on equalization of non-sparse signals, since we are trying to utilize the available degrees of freedom to the signal as rapidly as possible.

A number of papers [13]–[16] consider the problem of equalization with low-precision analog front ends, and propose methods for designing ADC quantizer levels. However, the emphasis in all of these papers remains on designing multiple slicer thresholds for a given sample, rather than dispersing slicers over time as we allow. Moreover, none of these focus on designing the front end to optimize the minimum BER (based on MAP decoding) as we do.

Mutual information computation for ADCs is considered in [13], [17]. Reference [17] focuses on time-interleaved ADCs, unlike the flash ADC architecture discussed here. Reference [13] considers the problem of designing non-contiguous quantizers for maximizing the mutual information between i.i.d. inputs and quantized outputs. However, mutual information quickly saturates with SNR, and is therefore not a good measure to optimize for the uncoded or lightly coded systems typical at high speeds. Moreover, non-contiguous quantization, if implemented by parallel comparators, does not fully utilize the available number of slicers. References [14]–[16], [18] also optimize BER as we do, but they restrict attention to simpler processing (based on a linear

transmit filter and DFE rather than the optimal BCJR algorithm employed here), hence their performance degrades quickly for heavy quantization and heavy precursor ISI. Our use of optimal nonlinear decoding enables significant reduction in the number of slicers while avoiding error floors: for instance, with an FR4 channel similar to the one used in [15] and [18], the BER that we obtain using only 5 slicers (equivalent to using a $\log_2(6)$ -bit ADC) is much smaller than what is reported there using a 3-bit ADC (7 slicers). Of course, the potential power savings in the analog front end from reducing the number of slicers must be balanced against the more complex digital backend. Such detailed tradeoffs are beyond our present scope, but as noted in the conclusions, are an important topic for future work. Even though we do not discuss implementation details of our space-time ADC architecture, circuit implementations with non-uniform thresholds have been discussed in [14]–[16] and [18] which provide evidence for the feasibility of such designs.

References [12], [19], [20], which discuss “compressive” AICs for sparse signals, present models for analysis of different types of circuit noise such as clocking jitter, amplifier distortions etc. Although their analysis is specific to the AIC design and the related pseudorandom noise component, some of the circuit impairment models can serve as useful starting points for analyzing the architecture described in this paper as well, which is an important topic for future work.

As already mentioned, the proof of our theoretical result on ℓ_1 distance preservation is analogous to that of the JL lemma [4] which provides a theoretical basis for compressed sensing. The result also appears at first glance to be similar to the *bit-conservation* principle articulated in [7], but the details and implications are completely different. The result in [7] considers signal reconstruction, and can be roughly paraphrased as saying that n 1-bit observations are equivalent to $n/2^k$ k -bit measurements. In contrast, our result says that n 1-bit measurements are equivalent to n infinite-precision measurements in terms of guaranteeing the feasibility of reliable data recovery in the low-noise regime (albeit with a smaller error exponent).

III. SYSTEM MODEL

We focus on uncoded transmission of binary symbols $\mathbf{b} = \{b_i\}$, with b_i chosen independently and equiprobably from $\{-1, +1\}$, at rate $1/T_s$ over a real baseband dispersive channel. The continuous time received signal at the input of the A/D conversion block is given by

$$x_c(t) = \sum_{i=-\infty}^{\infty} b_i h(t - iT_s) + w_c(t) \quad (1)$$

where $h(t) = (h_{TX} * h_c * h_{RX})(t)$ is the *effective* channel impulse response obtained by convolving the transmit filter $h_{TX}(t)$, the physical channel $h_c(t)$, and the receive filter $h_{RX}(t)$. Assuming white noise $n(t)$ with PSD σ^2 at the input to the receive filter, the noise $w_c(t) = (n * h_{RX})(t)$ at the input to the A/D block is zero mean Gaussian with autocorrelation function

$$R_{w_c}(\tau) = \sigma^2 \int h_{RX}(t) h_{RX}(t - \tau) dt \quad (2)$$

Input to Quantizer: Let $x(k) = x(s_k)$ denote the continuous-valued discrete time samples obtained by sampling at times $\{s_k\}$. For Nyquist sampling at rate $1/T_s$, we set $s_k = (k + \tau)T_s$, where $\tau \in [0, 1)$ is the sampling phase (suppressed in subsequent notation for simplicity of exposition). We assume that the receive filter is square root Nyquist (e.g. square root raised cosine) at rate $1/T_s$, so that the noise samples $w_c(kT_s)$ are uncorrelated. However, sampling irregularly, or faster than $1/T_s$, both of which we allow, yields correlated noise samples.

Quantizer: We denote by $q(x; \mathbf{T})$ the output of a quantizer mapping a real-valued sample x to $N + 1$ values using thresholds $\mathbf{T} = \{t_1, \dots, t_N\}$. For a classical n -bit quantizer, we have $N = 2^n - 1$. For a uniform quantizer over the range $[-R, R]$, we have

$$t_i = R \left(-1 + i \frac{2}{N+1} \right), \quad i = 1, \dots, N \quad (3)$$

Our goal here is to explore more flexible designs, in terms of choice of both N and \mathbf{T} .

In this paper, we consider three different scenarios:

1) *T-Spaced Equalization (TSE):* We consider regularly spaced samples at rate $1/T_s$, and we use a fixed quantizer for all samples. The effective discrete time channel is denoted by $\mathbf{h} = [h(0), h(T_s), \dots, h((L-1)T_s)]^T = [h_1, h_2, \dots, h_L]^T$, where $L-1$ is the channel memory. We note that

$$x(k) = \langle \mathbf{h}, \mathbf{b}_k^{k-L+1} \rangle + w(k) \quad (4)$$

where $\mathbf{b}_k^{k-L+1} = (b_k, b_{k-1}, \dots, b_{k-L+1})^T$ denotes the set of bits affecting the k th sample, $\langle \cdot \rangle$ the dot product, and $w(k)$ are i.i.d. $N(0, \sigma^2 \|h_{RX}\|^2)$. We assume that the same quantizer \mathbf{T} is used for all samples, so that the quantized samples are given by

$$x_q(k) = q(x(k); \mathbf{T}) \quad (5)$$

The key question in this setting is how the performance depends on \mathbf{T} , where we allow channel-dependent choices of \mathbf{T} .

2) *Fractionally Spaced Equalization (FSE):* We consider samples spaced by $T_s/2$ (the typical choice for FSE), which yields two parallel symbol rate observations, which can be modeled as two parallel discrete time channels \mathbf{h}_1 and \mathbf{h}_2 operating on the same symbol stream:

$$x_i(k) = \langle \mathbf{h}_i, \mathbf{b}_k^{k-L+1} \rangle + w_i(k), \quad i = 1, 2 \quad (6)$$

where L is the larger of the memory of the two parallel channels. The noise streams $w_i(k)$ are each white, but are correlated with each other. The correlations can be computed based on the autocorrelation function (2) of the continuous-time noise w_c . We also allow the quantizers for the two streams to differ, with thresholds \mathbf{T}_1 and \mathbf{T}_2 , so that the two-dimensional quantized observation at time k is given by $x_q(k) = [q(x_1(k); \mathbf{T}_1), q(x_2(k); \mathbf{T}_2)]^T$.

3) *General Space-Time Equalization:* Here we allow the sampling times $\{s_k\}$ to be arbitrary, and also allow the quantizer \mathbf{T}_k for each sample to vary.

Thus, our goal is to understand how to rethink equalizer design in the classical settings of scenarios 1 and 2 when we have severe quantization constraints. In considering scenario 3,

we try to provide a theoretical perspective on how flexible quantizer design can be, in terms of choice of sampling times and quantizers. In particular, we focus on high rate fractionally spaced sampling with randomly chosen and scalar \mathbf{T}_k , corresponding to one-bit quantization with time-varying thresholds.

We assume that the discrete time channels corresponding to the sampling points $\{s_k\}$ are known (e.g., see [11] and [21, Ch. 6] for approaches for channel estimation with low-precision quantization). We employ the BCJR [22] or the Viterbi MLSE algorithm [23] to evaluate various quantizer designs (for completeness, a quick review of how these apply to our setting is provided in the appendix). For irregular or faster than Nyquist sampling, the noise samples at the quantizer input are correlated, but we ignore these in running the BCJR or MLSE algorithm, which means that the performance in these settings could potentially be improved further by accounting for these correlations. However, accounting for such correlations in severely quantized observations is difficult, and we do not expect the gains to be significant at the high SNRs (typical for high-speed wireline links) considered here.

Example Channels: We use three channels as running examples (see Figures 1(a), 1(b), 1(c)) throughout the paper. Channel A models a 20 inch FR4 backplane channel at 10GHz [15], and has discrete time channel impulse response (CIR) $\mathbf{h}_{A,0} = [.1, .25, .16, .08, .04]$ (maximum phase, as is typical for backplane channels). Channel B, taken from [24], is mixed phase with CIR $\mathbf{h}_{B,0} = [.23, .46, .69, .46, .23]$. For simulations with irregular or faster than Nyquist sampling, the continuous channel impulse waveform is required. We generate it using interpolation with a raised cosine waveform with roll-off factor 0.5. This may be interpreted as using matched square root raised cosine (SRRC) pulses for the transmit and receive filters with physical channel impulse response $h_c(t) = \sum_{i=1}^L h_i \delta(t - i)$ (setting $T_s = 1$ without loss of generality). Channel C is generated by SRRC transmit and receive pulses as above, with physical channel $h_c(t) = .2\delta(t - 1) + .3\delta(t - 1.85) + .15\delta(t - 2.55) + .25\delta(t - 3.35) + .05\delta(t - 4.6)$. This gives a channel with a broader peak (formed from the merging of two peaks) than the other two. The impulse responses ($h(t)$) of the 3 channels are shown in the subfigures 1(a), 1(b), 1(c). The notation $\mathbf{h}_{A,\tau}$, $0 \leq \tau < 1$ is used to denote the CIR obtained by sampling at the sampling phase τ (i.e., the sampling times are at $(k + \tau)T_s$). For instance $\mathbf{h}_{C,1/2} = [-.03, .24, .3, .22, .03, .01]$.

IV. NYQUIST SAMPLED UNIFORM ADC

We first consider the standard setting of Nyquist sampling with uniform ADC with N thresholds as in Eq. (3), and ask how small N can be for a given channel while avoiding an error floor (i.e., error-free reception at infinite SNR)? An analytical characterization is intractable, but it is possible to evaluate N_{\min} numerically by fixing $\sigma^2 = 0$, and increasing N until the information rate reaches its maximum value (for binary signaling) of one. The information rate can be evaluated via Monte Carlo methods using BCJR as described in [25]. However, it is interesting to explore whether there are analytical insights to be obtained by examining the

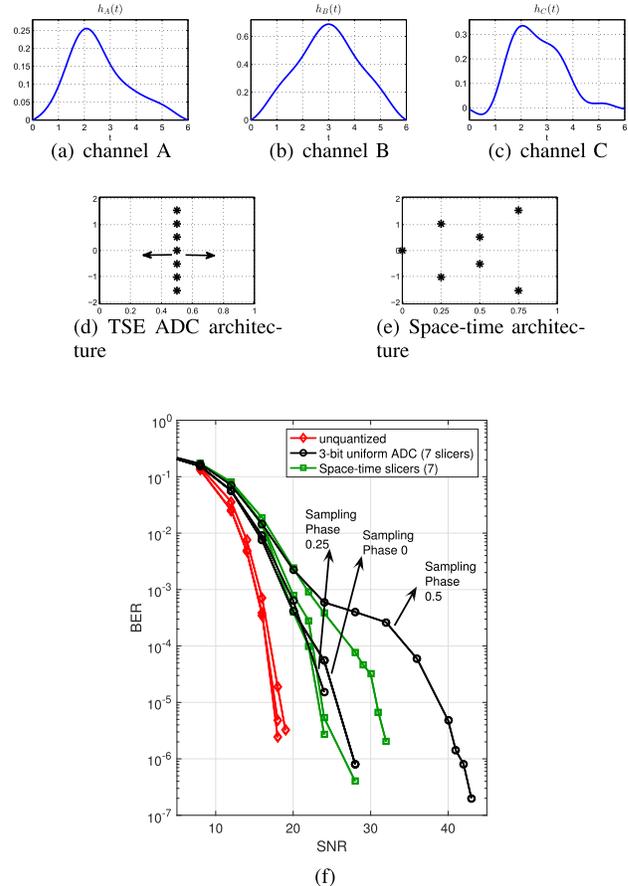


Fig. 1. (d),(e) Each star symbol denotes a slicer, x and y axis correspond to the sampling phase and threshold of the slicer, respectively. All slicers still operate at the Nyquist symbol rate (normalized to 1 in the figures). (f) Bit error rate curves for channel B corresponding to different sampling phases 0, 0.25, 0.5.

channel coefficients. Intuitively, we expect that a channel with a strong dominant tap should have a lower value of N_{\min} , compared to a channel where the taps are comparable. The placement of the dominant tap should also have a significant effect. We make these intuitions concrete via the lemma stated next, which provides easily computable bounds for N_{\min} when all the channel taps have the same sign (which is often a good approximation for backplane channels, for example). The proof of the lemma, given in the appendix, is based on bounds on information rate derived by Zeitler *et al.* [13].

Before stating the lemma, we note that the symmetric information rate is invariant under time reversal and scaling (under fixed SNR) of the channel. The scaling result is standard, and the time reversal result follows because the same output is generated by feeding a time reversed bit stream (which is another valid i.i.d. input) to the time reversed channel. Naturally, the bounds in the lemma also exhibit these invariances. Define $\mathbf{g} = \frac{\mathbf{h}}{\|\mathbf{h}\|_1}$ as a normalized version of \mathbf{h} with unit ℓ_1 norm, and set $\tilde{\mathbf{g}}$ as the time-reversed version of \mathbf{g} , so that $\tilde{g}_i = g_{L-i+1}$, $i = 1, \dots, L$. Define

$$N_l = \left\lceil \frac{1}{\max_i (g_i)} - 1 \right\rceil \quad (7)$$

TABLE I
MINIMUM NUMBER OF THRESHOLDS REQUIRED TO DECODE WITH NO ERROR AT HIGH SNR. ALSO LISTED ARE THE LOWER AND UPPER BOUNDS COMPUTED USING LEMMA 1

\mathbf{h}	N_l	N_u	N_{\min}
$\mathbf{h}_{B,0} = [.23 .46 .69 .46 .23]$	2	8	5
$[\.46 .69 .46 .23 .23]$	2	4	2
$[\.69 .46 .46 .23 .23]$	2	2	2
$\mathbf{h}_{B,1/4} = [.04 .29 .54 .67 .39 .16]$	3	8	5
$\mathbf{h}_{B,1/2} = [.09 .34 .61 .61 .34 .09]$	3	6	6
$\mathbf{h}_{A,0} = [.1 .25 .16 .08 .04]$	2	4	3
$\mathbf{h}_{C,0} = [.05 .33 .26 .11 .02]$	2	2	2

$$N_u = \min \left(\{[u_i], 2 \leq i \leq L-1\}, \{[v_i], 2 \leq i \leq L-1\}, \left[\frac{1}{g_1} - 1 \right], \left[\frac{1}{\tilde{g}_1} - 1 \right] \right) \quad (8)$$

where

$$u_i = \frac{1}{(g_i - \sum_{j=1}^{i-1} g_j)_+} - 1; \quad v_i = \frac{1}{(\tilde{g}_i - \sum_{j=1}^{i-1} \tilde{g}_j)_+} - 1$$

where $(x)_+ = x$ if $x > 0$ and $(x)_+ = 0$ if $x \leq 0$ and $\lceil \cdot \rceil$ denotes the *ceil* function i.e. rounding to the next integer. Thus, we allow u_i, v_i to take the value $+\infty$, but the value of N_u is guaranteed to be finite because of the last two terms in the minimum. It is also easy to see that $v_{L-i+1} = \frac{1}{(g_i - \sum_{j=i+1}^L g_j)_+} - 1$.

Lemma 1: The minimum number of levels for avoiding an error floor is bounded as $N_l \leq N_{\min} \leq N_u$.

The lower and upper bounds capture the effect of the strength and the location of the dominant tap, respectively. An examination of the expression (8) for N_u shows that, if we can permute a given set of channel coefficients, maximum or minimum phase channels (most of the energy in ending or beginning taps) will generally have smaller N_{\min} compared to mixed phase channels (most of the energy in the taps in the middle). Table I lists the values of N_{\min} (computed numerically) for a few different channels along with the lower and upper bounds. The channel in the first row is mixed phase and has the highest N_{\min} compared to rows 2 and 3 channels which have the same tap values but are instead maximum phase. We find that for a fixed channel, varying the sampling phase may slightly change N_{\min} . However, as we show next, the shape of the BER curve and the performance at moderate SNRs may be far more sensitive to the sampling phase.

While the bounds of the number of ADC levels are not always tight (see Table I, channels $\mathbf{h}_{B,0}$ and $\mathbf{h}_{B,1/4}$), they provide valuable insights on what is needed to avoid error floors. The lower bound is the inverse of the dominant tap strength minus 1 (assuming that the channel coefficient magnitudes are normalized to sum to one), and the upper bound depends on the difference of the strength of the dominant tap(s) and the taps before and after. Thus, if the dominant tap is much larger than the other taps, then we need fewer quantization levels. This makes intuitive sense, since the effect of ISI is less severe, and we need to devote fewer resources to modeling it. On the other hand, the number of quantization levels is at least $L-1$ for a uniform channel of length L .

For suboptimal linear equalization with unquantized samples, it is well known [24], [26] that fractionally spaced equalizers (FSE) are superior to symbol-spaced equalizers, providing robustness to sampling phase and avoiding error floors due to residual interference. However, when optimal BCJR or MLSE equalization is employed, the difference is not as drastic, but FSE is still more insensitive to sampling phase, which is attractive because hardware-based control of sampling phase is not always feasible. We would like to investigate if similar trends hold with severe quantization, with a quick exploration in this section followed by more detailed theory and algorithms in later sections. In order to have a fair comparison between TSE and FSE, we take the number of slicers used in a TSE and disperse them across different sampling phases to obtain a *space-time* architecture.

As an example, we plot in Fig. 1(f) the BER over channel B with TSE. We consider unquantized samples, as well as samples quantized using a 3-bit ADC (i.e., with 7 slicers), for sampling phases $\tau = 0, 0.25, 0.5$. In the unquantized setting, there is a small degradation in performance (~ 1 dB at 10^{-5}) at sampling phase 0.5. However, the degradation with quantization is much larger, even though there is no error floor (see the $\mathbf{h}_{B,1/2}$ entry in Table I). Even for channels with similar dynamic ranges, the performance of TSE/uniform-ADC with a fixed set of thresholds can show significant sensitivity to sampling phase. As a quick remedy, we try spreading the *same* set of slicers across time, as shown in Fig. 1(e). Each slicer still samples at the symbol rate but with a different sampling phase. Changing the sampling phase now corresponds to shifting the whole space-time slicer structure. We see that now the performance (the BER curves in green) is much less sensitive to the phase, although there is still some degradation for one of the sampling phases. This was a specific configuration, obtained without any design, which demonstrated the potential of space-time slicers. However, there are numerous ways in which the slicers can be spread across time, hence it is of interest to develop automated procedures for arriving at good designs. It is also natural to ask the question as to whether there is any fundamental disadvantage to spreading slicers across time.

In the next section, we show that even randomly distributed slicers spread across time suffice to avoid error floors as long as the number of slicers is large enough. While such a design would be impractical, it shows that, in principle, there is no fundamental performance limitation imposed by quantization using one-bit comparators spread out in space and time. Of course, the number of slicers predicted by this theoretical result is much larger than what is required when the space-time architecture is optimized for a particular channel, and we consider this problem in Section VI.

V. ONE-BIT MEASUREMENTS WITH RANDOM THRESHOLDS

In this section, we consider the special case of 1-bit measurements spread over time. Without loss of generality, consider reliable demodulation of bit b_0 . We restrict attention to measurements in the interval $[0, LT_s]$ affected by this bit. This choice of observation interval is sensible but arbitrary,

and our approach applies to other choices as well. The measurements in this interval are also affected by $L - 1$ “past” ISI bits (b_{-L+1}, \dots, b_{-1}) and $L - 1$ “future” ISI bits (b_1, \dots, b_{L-1}). Denote the noiseless received waveform in this interval by $s(t)$, suppressing the dependence on the desired bit b_i and the ISI bits from the notation. Without loss of generality, we normalize $h(t)$ so that $s(t)$ lies in $[-1, 1]$. The main result in this section can be paraphrased as follows: for sufficiently many 1-bit measurements uniformly spaced in time but with thresholds chosen randomly over $[-1, 1]$, it is possible (at high SNR) to reliably distinguish between $b_0 = +1$ and $b_0 = -1$, as long as it is possible to do so with unquantized measurements.

Information Rate: Let \mathbf{x}_i^j denote the vector of samples (these may or may not be quantized) obtained during the interval $[iT_s, jT_s]$. For symbol spaced sampling, the length of \mathbf{x}_i^j is $j - i + 1$ (the length for general space-time slicers depends on the specific pattern of sampling times used). The information rate between the transmitted bits and the received samples is given by

$$\begin{aligned} I(\mathbf{b}; \mathbf{x}) &= \lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{b}_1^N; \mathbf{x}_1^N) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N I(b_i; \mathbf{x}_i^N | b_{i-L+1}^{i-1}) \\ &\geq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N I(b_i; \mathbf{x}_i^{i+f} | b_{i-L+1}^{i-1}) \end{aligned} \quad (9)$$

Inequality (9), derived in [13], states that the information rate is lower bounded by the average (over the past bits) mutual information between the current bit and the measurements over the next few symbols (f), conditioned on the past bits. Numerical results in [13] show that this lower bound becomes a fairly tight approximation for $f = L$ future symbols.

Let \mathbf{x}_0^L denote the vector of continuous-valued samples obtained by sampling $s(t)$ uniformly, n times, over the observation interval. Length of \mathbf{x}_0^L is n . If we denote the sampling interval by Δ , we have $n\Delta = LT_s$. Fixing the past ISI bits, we partition the noiseless waveforms corresponding to all possible realizations of the future bits into two sets, each of cardinality 2^{L-1} , corresponding to the two possible values of the “tagged bit” b_0 : $\mathcal{S}_{-1} = \{s(t) \text{ s.t. } b_0 = -1\}$ and $\mathcal{S}_{+1} = \{s(t) \text{ s.t. } b_0 = +1\}$. Denote by \mathcal{X}_{-1} and \mathcal{X}_{+1} the corresponding sampled vectors \mathbf{x}_0^L . The absence of error floors can be proved by setting the noise level to zero and checking whether the lower bound (9) on the information rate equals one. This happens as long as the set of observations generated by the two different values of the desired bit are mutually exclusive: $\mathcal{X}_{-1} \cap \mathcal{X}_{+1} = \emptyset$. Note that this property always holds for unquantized measurements, as long as at least one sample is obtained in the first symbol period ($[0, T_s]$) and the corresponding CIR value $h(0) \neq 0$. This follows from the fact that, since the past bits are fixed, and future ISI bits do not affect the waveform in the interval $[0, T_s]$, $b_0 = -1$ and $b_0 = +1$ result in different samples in the first entry of \mathbf{x}_0^L . This result is also discussed in [27], where the author considers symbol spaced samples and shows that the lower bound

(and hence the information rate) goes to one as SNR increases as long as the first element of the discrete time CIR is nonzero. In general, such guarantees cannot be provided for quantized measurements. However, we show that as long as n is large, using randomized thresholds for one-bit quantization results in similar behavior.

In general (at any SNR), the performance depends on the amount of overlap/separability between the sets \mathcal{X}_{-1} and \mathcal{X}_{+1} . For the purpose of our proof, we employ the *normalized* ℓ_1 distance between each pair of elements $\mathbf{x}_{-1} \in \mathcal{X}_{-1}$, $\mathbf{x}_{+1} \in \mathcal{X}_{+1}$, defined as follows:

$$\|\mathbf{x}_{-1} - \mathbf{x}_{+1}\|_1 = \sum_{i=1}^n \Delta |s_{-1}(i\Delta) - s_{+1}(i\Delta)| \quad (10)$$

where $s_{-1}(t)$ and $s_{+1}(t)$ are the corresponding continuous time waveforms from sets \mathcal{S}_{-1} and \mathcal{S}_{+1} respectively and Δ is the sampling interval (for uniform sampling as assumed in this section, $n\Delta = LT_s$). The scale factor Δ is included for the normalized ℓ_1 norm $\|\mathbf{x}_{-1} - \mathbf{x}_{+1}\|_1$ to approximate the continuous time ℓ_1 norm $\|s_{-1} - s_{+1}\|_1$ as n gets large. We define the minimum normalized ℓ_1 distance between the two sets as follows:

$$d = \min_{\mathbf{x}_{-1} \in \mathcal{X}_{-1}; \mathbf{x}_{+1} \in \mathcal{X}_{+1}} \|\mathbf{x}_{-1} - \mathbf{x}_{+1}\|_1 \quad (11)$$

For unquantized observations, as noted earlier, $\mathcal{X}_{-1} \cap \mathcal{X}_{+1} = \emptyset$, and hence $d > 0$.

Let us now consider what happens when we pass the unquantized sampled vector \mathbf{x} through a series of one-bit quantizers, with the i th sample compared to threshold t_i . The vector of thresholds is denoted as $\mathbf{T} = [t_1, t_2, \dots, t_n]^T$, and defines a quantization function q as follows:

$$q(\mathbf{x}) = (2\Delta)\mathbf{y}; \quad y(i) = \begin{cases} 1 & \text{if } x(i) \geq t_i \\ 0 & \text{if } x(i) < t_i \end{cases} \quad i = 1, \dots, n \quad (12)$$

The following theorem states that, with a sufficient number of samples n , quantized with random thresholds, the quantization function $q(\cdot)$ approximately preserves the ℓ_1 norm of the unquantized differences $\|\mathbf{x}_{-1} - \mathbf{x}_{+1}\|_1$. This result bears some similarity to the JL lemma in which random projections preserve the norm for embeddings to lower dimension subspaces [28].

Theorem: If each entry of the threshold array \mathbf{T} is picked uniformly and independently from $[-1, 1]$, then for any constants $\epsilon, \beta, \delta \geq 0$, with probability at least $1 - \delta$, for all $\mathbf{x}_{-1} \in \mathcal{X}_{-1}$; $\mathbf{x}_{+1} \in \mathcal{X}_{+1}$ we have

$$(1 - \epsilon) \|\mathbf{x}_{-1} - \mathbf{x}_{+1}\|_1 \leq \|q(\mathbf{x}_{-1}) - q(\mathbf{x}_{+1})\|_1 \leq (1 + \epsilon) \|\mathbf{x}_{-1} - \mathbf{x}_{+1}\|_1 \quad (13)$$

for

$$n \geq \frac{4T_s}{d\epsilon^2} \left(\log 2 \cdot (2L^2 + L) + L \log \delta^{-1} \right) \quad (14)$$

where d is the minimum ℓ_1 distance defined in (11).

Proof: Consider a particular pair of sampled measurements $\mathbf{x}_{-1} \in \mathcal{X}_{-1}$; $\mathbf{x}_{+1} \in \mathcal{X}_{+1}$ (corresponding to $s_{-1}(t) \in \mathcal{S}_{-1}$; $s_{+1}(t) \in \mathcal{S}_{+1}$). Define $\mathbf{z} = |q(\mathbf{x}_{-1}) - q(\mathbf{x}_{+1})|$, so

that $z(i) = 2\Delta$ if t_i lies between (and hence can distinguish between) $s_{+1}(i\Delta)$ and $s_{-1}(i\Delta)$, and $z(i) = 0$ otherwise. Since t_i is uniformly picked from $[-1, 1]$, $z(i)$ is a (scaled version of a) Bernoulli random variable with parameter $p_i = \frac{1}{2} |s_{-1}(i\Delta) - s_{+1}(i\Delta)|$ and mean $2\Delta p_i$. Thus, from (10)

$$\begin{aligned} E(\|\mathbf{z}\|_1) &= E\left(\sum_{i=1}^n z(i)\right) \\ &= 2\Delta \sum_i \frac{|s_{-1}(i\Delta) - s_{+1}(i\Delta)|}{2} = \|\mathbf{x}_{-1} - \mathbf{x}_{+1}\|_1 \end{aligned} \quad (15)$$

so that the quantization function $q(\cdot)$ preserves the norms of the differences in expectation. It remains to prove a concentration result using a Chernoff bound to show that the probability of deviation from the expectation goes to zero for large enough n . Given that the $z(i)$ are independent scaled Bernoulli random variables, derivation of the Chernoff bound is a straightforward exercise and we state the final result, omitting the details (see [29]). To simplify notation, we use the shorthand $\mu = \|\mathbf{x}_{-1} - \mathbf{x}_{+1}\|_1$ in the following.

$$\Pr(\|\mathbf{z}\|_1 > (1 + \epsilon)\mu) \leq e^{-\frac{\mu}{2\Delta}((1+\epsilon)\log(1+\epsilon) - \epsilon)} \leq e^{-\frac{\mu n \epsilon^2}{4LT_s}} \quad (16)$$

where we have substituted $\Delta = \frac{LT_s}{n}$ and used $\log(1 + \epsilon) \geq \epsilon$ (for $\epsilon \geq 0$) to obtain the last inequality. Proceeding along similar lines, we obtain an analogous bound for the probability of deviation below the expectation: $\Pr(\|\mathbf{z}\|_1 < (1 - \epsilon)\mu) \leq e^{-\frac{\mu n \epsilon^2}{4LT_s}}$. Combining with (16) yields

$$\begin{aligned} \Pr(\|\mathbf{z}\|_1 < (1 - \epsilon)\mu \text{ or } \|\mathbf{z}\|_1 > (1 + \epsilon)\mu) &\leq 2e^{-\frac{\mu n \epsilon^2}{4LT_s}} \\ &\leq 2e^{-\frac{dn\epsilon^2}{4LT_s}} \end{aligned} \quad (17)$$

where the last inequality follows from the definition of d in (11). There are 2^{L+1} pairs of distances given the past bits (i.e. $|X_{-1}| = |X_{+1}| = 2^L$), and varying the L past bits, $|X_{-1}| = |X_{+1}| = 2^L$, and taking the union bound over all possible pairs $\mathbf{x}_{-1} \in X_{-1}$; $\mathbf{x}_{+1} \in X_{+1}$, we obtain

$$\begin{aligned} \Pr(\|\mathbf{z}\|_1 \leq (1 - \epsilon)\mu \text{ or } \|\mathbf{z}\|_1 \geq (1 + \epsilon)\mu) &\leq 2^{2L} \cdot 2e^{-\frac{dn\epsilon^2}{4LT_s}} \\ &\leq \delta \end{aligned} \quad (18)$$

which can be bounded as tightly as desired (18) by decreasing δ and ensuring that n meets the condition (14). ■

Remarks: While we have considered uniform sampling for simplicity, this is not required for the theorem to hold. Using the continuity of the CIR, any non-uniform sampling strategy that provides sufficient density of samples to capture the separation of $s_{-1}(t)$ and $s_{+1}(t)$ in the regions where the waveforms are apart suffices. The independence of the choice of thresholds is crucial for the concentration result.

Simulations: Due to the looseness of the union bound used to prove the theorem, picking n based on the theorem is excessively conservative. We now show via simulations that moderate values of n suffice to provide good equalization performance. Figure 2(b)) shows the BER curve obtained

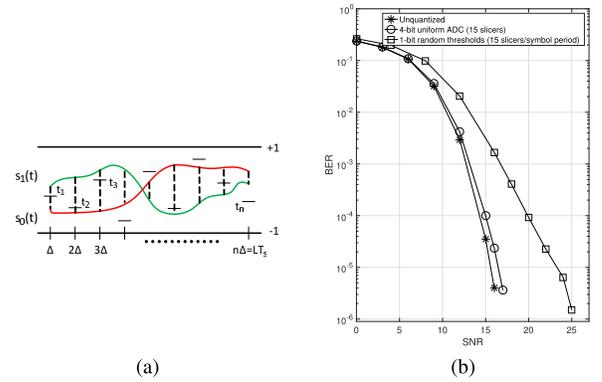


Fig. 2. (a) One-bit measurements with randomly varying thresholds (b) Bit error rates for the channel $\mathbf{h}_{A,0} = [1, .25, .16, .08, .04]$.

by employing 90 randomly selected 1 bit slicers for the FR4 channel. This translates to 15 slicers per symbol period as FR4 is 6 symbols long. The SNR is defined as $\frac{\|\mathbf{h}\|^2}{\sigma^2}$. The BER curves vary slightly for different instances of slicer thresholds, the general behavior remains the same for a fixed number of slicers and we find that ~ 15 slicers suffice to avoid the error floor. BER for random 1-bit slicers is worse than for uniform ADC thresholds since no optimization has been performed. Carefully picking slicer thresholds for minimizing BER is discussed in the next section. The bit error rates are computed empirically using BCJR. Note that the BER obtained for the random slicers case is actually an upper bound of the minimum BER as the BCJR algorithm used ignores the noise correlations and hence is not optimal. As also mentioned in the appendix, it is non-trivial to extend BCJR for the case with quantization and colored noise (even though each these 2 scenarios alone can be handled).

The theorem states that the L1 norm of difference between two waveforms is preserved after quantization through a random set of thresholds, in a manner analogous to the preservation of norms of randomly projected vectors via the JL Lemma, which is used in compressed sensing. As in the latter application, we use union and Chernoff bounds, which are quite loose. However, the result provides the key insight that information can be preserved using one-bit comparators with random quantization thresholds, and motivates the search for constructive strategies. While the theorem is reassuring testimony to the flexibility of space-time architectures, in practice, it is often simpler to place slicers at fewer locations. In the next section, we consider optimization of slicer locations for TSE and FSE.

VI. OPTIMIZING SLICER THRESHOLDS

In the example discussed in Section IV, we observed that the uniform ADC performed very poorly at the sampling phase 0.5 with channel B ($\mathbf{h}_{B,1/2}$). A closer look at the error events (at 25 dB) reveals that most of the errors are caused due to poor threshold locations rather than large noise samples. Fig. 3(a) plots the continuous-valued signals corresponding to the correct and incorrect bit sequences from a simulation run in which bits 1 and 2 have been incorrectly decoded. Both noiseless and noisy signals are plotted, but they are

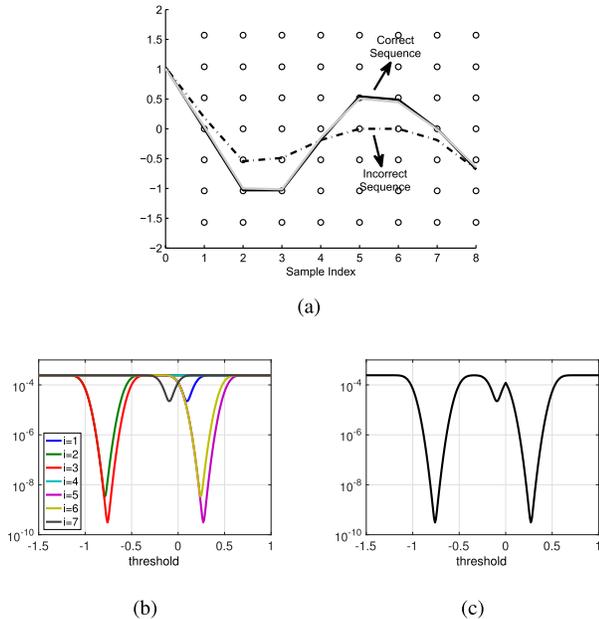


Fig. 3. (a) Example of an error event with channel $\mathbf{h}_{B,1/2}$ at 25 dB. Plot in gray is after noise addition. The small circles denote slicers. (b) Probability of error for different indices Eq. (22) (c) $g(\Omega, t)$ for the sequence shown in (a) at 25 dB.

barely discernible from each other (i.e., the noise samples are small). The noiseless sequences differ significantly at 4 sample locations (locations 2, 3, 5, 6) affected by bits 1 and 2, but at all of these, the thresholds separating the two waveforms are very close to at least one of them, hence even a small deviation due to noise greatly increases the possibility of an incorrect detection. This shows that, for low-precision quantization, it is critical to choose thresholds that are compatible with the channel at hand, since “off-the-shelf” uniform ADCs may not effectively separate out the waveforms corresponding to different bit sequences. Uniform thresholds are more compatible with Channel B with a different sampling phase, $\mathbf{h}_{B,0}$, but here too, the performance can be improved by choosing channel-specific thresholds. In this section, we present a procedure for designing a non-uniform ADC with thresholds chosen based on the channel, given a constraint on the number of slicers. We first consider a TSE, and then extend the algorithm to an FSE sampled at twice the Nyquist rate. We assume that the sampling phase is fixed for the receiver, which determines the channel tap values. Slicer thresholds are designed for a given channel response.

A. Threshold Design for TSE

Ideally, we would like to choose the thresholds, $\mathbf{T} = [t_1, \dots, t_M]$, to minimize the *minimum* BER attained by MAP/BCJR decoding. However, this cost function is analytically intractable, hence we consider the union bound for MLSE performance and truncate it to a few dominant terms, targeting a high SNR regime. We use as our cost function an upper bound of this truncated sum, which can be computed easily for quantized observations.

The MLSE bit error probability, P_e , can be upper bounded using the union bound, which in its general form can be stated

as follows [30, Sec. 5.8.1]

$$P_e \leq P_u = \sum_{\mathbf{e} \in \mathcal{E}} \sum_{\mathbf{b}, \mathbf{b}' } P_B(\mathbf{b}, \mathbf{b}') w(\mathbf{e}) 2^{-w(\mathbf{e})} \quad (19)$$

where $\mathbf{b}' = \mathbf{b} + 2\mathbf{e}$

where \mathcal{E} denotes the set of error events. As defined in [30] an error event is a simple error sequence whose first nonzero entry is at a fixed time, say at index 0. The elements of \mathbf{e} take values in $\{0, \pm 1\}$, and are nonzero at indices where the bit sequences \mathbf{b} and \mathbf{b}' differ. The number of nonzero elements in \mathbf{e} , or its weight, is denoted by $w(\mathbf{e})$. We denote by $P_B(\mathbf{b}, \mathbf{b}')$ the pairwise error probability for binary hypothesis testing between \mathbf{b} and \mathbf{b}' , which are separated by the error event expressed by \mathbf{e} . For continuous-valued measurements, $P_B(\cdot)$ depends only on \mathbf{e} , which reduces the summation $\sum_{\mathbf{b}, \mathbf{b}'}$ to a single term that can be expressed as a function of the standard normal complementary CDF (or Q function; see [30, eq. (5.76)]). Exact evaluation of $P_B(\cdot)$ is difficult for quantized observations, hence we bound it from above. This, together with a restriction on the set of error events, yields an approximate upper bound that serves as our cost function for threshold design using K-means.

1) *Truncated Union Bound*: While there are infinitely many error events in \mathcal{E} , at high SNR, it suffices to consider a small set of most likely events which dominate the summation (19). For continuous-valued measurements, these correspond to the most slowly decaying Q function terms, which correspond to low weight error sequences [30]. For quantized observations, it is more difficult to identify the dominant error events, but for the channels considered here, and using the uniform quantizer starting point, simulations yield the expected result: weight one and two error patterns, $\mathbf{e}_1 = \{\pm 1, 0, 0, 0, \dots\}$ and $\mathbf{e}_2 = \{\pm 1, \pm 1, 0, 0, 0, \dots\}$, are by far the most dominant. We therefore restrict attention to these in truncating the union bound (19), as follows:

$$P_u \approx P_{ut} = \sum_{\Omega \in E_1} P_B(\Omega) w(\mathbf{e}_1) 2^{-w(\mathbf{e}_1)} + \sum_{\Omega \in E_2} P_B(\Omega) w(\mathbf{e}_2) 2^{-w(\mathbf{e}_2)} \quad (20)$$

where $E_i = \{\mathbf{b}, \mathbf{b}' \text{ s.t. } \mathbf{b}' = \mathbf{b} + 2\mathbf{e}_i\}$, $i = 1, 2$ and we have denoted pairs of bit sequences $(\mathbf{b}, \mathbf{b}')$ by Ω for brevity and $w(\mathbf{e}_1) = 1$, $w(\mathbf{e}_2) = 2$. Note that $|E_1| = 2^{(L-1)} \cdot 2^{(L-1)}$. This is because the observations that depend on the bit in error, b_0 , are only affected by the truncated bit sequence $\mathbf{b}_{-(L-1)}^{L-1}$. Similarly we get $|E_2| = 2^{(L-1)} \cdot 2^{(L-1)} \cdot 2$. For a channel with $L = 6$, $|E_1| = 1024$, $|E_2| = 2048$ which gives the total terms to be summed over to be $N = |E_1| + |E_2| = 3072$.

2) *Bounding the Pairwise Error Probability*: Note that in the notation so far we have suppressed the dependence on thresholds \mathbf{T} . We make that explicit now. Hence in Eq. 20 above, we really have pair wise error probabilities that depend on the bit sequences and the thresholds i.e. $P_B(\Omega, \mathbf{T})$. We now wish to bound the pairwise error probabilities $P_B(\Omega, \mathbf{T})$ for a particular set of thresholds \mathbf{T} . Consider the corresponding noiseless unquantized signals $\mathbf{x} = \langle \mathbf{h}, \mathbf{b} \rangle$ and $\mathbf{x}' = \langle \mathbf{h}, \mathbf{b}' \rangle$. Since we are only interested in simple error sequences,

\mathbf{x} and \mathbf{x}' differ at most in, say K , consecutive locations. That is, $x(i) = x'(i) \quad \forall i \leq 0, i \geq K + 1$. Note that $K = L$ for $\mathbf{b}, \mathbf{b}' \in E_1$ and $K = L + 1$ for E_2 (changing a given bit can have an effect over at most L output samples when convolved with a channel of length L). The binary hypothesis problem of choosing one of \mathbf{b} and \mathbf{b}' then reduces to selecting one of the two vectors, \mathbf{X}_0 or \mathbf{X}_1 given by

$$\begin{aligned} H_0: \quad \mathbf{X}_0 &= \mathbf{x}(1:K), \quad H_1: \quad \mathbf{X}_1 = \mathbf{x}'(1:K); \\ P_B(\Omega, \mathbf{T}) &= P_B(\mathbf{b}, \mathbf{b}', \mathbf{T}) = P_B(\mathbf{X}_0, \mathbf{X}_1, \mathbf{T}) \end{aligned}$$

Fig. 3(a) shows an example of \mathbf{X}_0 and \mathbf{X}_1 corresponding to a particular bit sequence pair in E_2 . The vectors \mathbf{X}_0 and \mathbf{X}_1 are of length K , after quantization each element takes one of $M + 1$ values, as there are M thresholds. We can now obtain a simple upper bound on the pairwise error probability by considering the probability of error in separating the scalars $X_0(i)$ and $X_1(i)$. The pairwise error probability if we only use the i th component depends only on a single threshold in the array \mathbf{T} . That is,

$$P_B(X_0(i), X_1(i), \mathbf{T}) = \min_{t \in \mathbf{T}} P_B(X_0(i), X_1(i), t) \quad (21)$$

Note that the value of t that minimizes the above expression is the one that is closest to $\frac{X_0(i)+X_1(i)}{2}$. As a function of this scalar threshold t , we obtain that

$$\begin{aligned} P_B(X_0(i), X_1(i), t) \\ = 2^{-(2L-2)} \left(Q\left(\frac{t - X_{\min}}{\sigma}\right) + Q\left(\frac{t - X_{\max}}{\sigma}\right) \right) \end{aligned} \quad (22)$$

where

$$\begin{aligned} X_{\min} &= \min(X_0(i), X_1(i)) \\ X_{\max} &= \max(X_0(i), X_1(i)) \end{aligned}$$

The factor of $2^{-(2L-2)}$ is included due to the prior on the truncated bit sequences. Fig. 3(b) plots this function for different indices $i = 1, \dots, 7$ for the pair of sequences X_0 and X_1 shown in Fig. 3(a). The probability of error for deciding between the hypothesis H_0 and H_1 can be upper bounded by each of the probabilities of error based on the scalar components as we vary i , hence minimizing over i provides an upper bound. Applying this and Eq. 21 we get:

$$\begin{aligned} P_B(\Omega, \mathbf{T}) &= P_B(\mathbf{X}_0, \mathbf{X}_1, \mathbf{T}) \leq \min_{i=1, \dots, K} P_B(X_0(i), X_1(i), \mathbf{T}) \\ &= \min_i \min_{t \in \mathbf{T}} P_B(X_0(i), X_1(i), t) \\ &= \min_{t \in \mathbf{T}} \min_i P_B(X_0(i), X_1(i), t) \end{aligned} \quad (23)$$

3) *Truncated Upper Bound:* Define

$$g(\Omega, t) = \min_i P_B(X_0(i), X_1(i), t) \quad (24)$$

Fig. 3(c) shows an example plot of the function $g(\Omega, t)$ for the pair of sequences shown in Fig. 3(a). We can rewrite the upper bound as

$$P_B(\mathbf{b}, \mathbf{b}', \mathbf{T}) = P_B(\Omega, \mathbf{T}) \leq \min_{t \in \{t_1, \dots, t_M\}} g(\Omega, t) \quad (25)$$

Applying Eq. (25) to Eq. (20), we get an upper bound on the truncated union bound, which is our cost function

$$\begin{aligned} P_e &\leq P_u \approx P_{ut} = \sum_{\Omega \in E_1 \cup E_2} P_B(\Omega, \mathbf{T}) w(\Omega) 2^{-w(\Omega)} \\ &\leq \sum_{n=1}^N \min_{t \in \mathbf{T}} g(\Omega_n, t) w(\Omega_n) 2^{-w(\Omega_n)} \\ &= \sum_n \min_{t \in \mathbf{T}} f(\Omega_n, t); \end{aligned} \quad (26)$$

where

$$f(\Omega, t) = g(\Omega, t) w(\Omega) 2^{-w(\Omega)} \quad (27)$$

$w(\Omega)$ denotes the weight of the error event $\mathbf{e} = \frac{\mathbf{b}' - \mathbf{b}}{2}$ corresponding to $\Omega = (\mathbf{b}, \mathbf{b}')$. Our objective now is to find thresholds that minimize this cost function (Eq. 26).

4) *Optimization Using K-Means:* The problem of finding the thresholds now reduces to the following minimization problem

$$\mathbf{T}^* = \underset{\mathbf{T}}{\operatorname{argmin}} \sum_{n=1}^N \min_{t \in \{t_1, \dots, t_M\}} f(\Omega_n, t) = \underset{\mathbf{T}}{\operatorname{argmin}} \sum_{n=1}^N f(\Omega_n, t_n^*) \quad (28)$$

We note that the above formulation is identical to the *clustering* problem where we are given N data points Ω_n , which are required to be grouped into M clusters to minimize the total distortion. The distortion function is specified by $f(\Omega, t)$ and the M cluster centers represent the thresholds. We can therefore apply the standard K-means [31] algorithm to obtain candidate solutions. This involves two alternating steps, as follows:

$$\begin{aligned} j^* &= \underset{j=1, \dots, M}{\operatorname{argmin}} f(\Omega_n, t_j^i) \text{ (assignment)} \\ t_j^{i+1} &= \min_t \sum_{\Omega \in t_j^i} f(\Omega, t) \text{ (update)} \end{aligned}$$

At the i^{th} iteration there are M cluster centers/thresholds $\{t_1^i, \dots, t_M^i\}$. Each 'data point' Ω_n gets assigned to its closest threshold. Then each threshold value is updated by summing over all the data points assigned to it. The functions $f(\Omega, t)$ can be easily computed for each pair of weight 1 and 2 error sequences, and we compute and store them for each Ω over a grid for the parameter t . This makes the minimization in the update step straightforward. We use a grid of size 200, after first normalizing the channel to limit the range of the unquantized channel output to $[-1, 1]$, and then using a grid of size .01 for t . The K-means algorithm typically converges in a small number of iterations (< 10). Since the K-means algorithm tends to get stuck in local minima, we run several iterations of it using different starting points such as uniform thresholds, Lloyd-Max thresholds and random perturbations of these. Finally we pick the thresholds that give the lowest approximate upper bound P_{ut} . Precomputing functions $f(\Omega, t)$ for a grid of threshold values and Ω makes it possible to run several iterations of K-means from different starting points very quickly. It also enables rapid computation of the truncated upper bound P_{ut} . Using P_{ut} as a proxy for the

real BER we can quickly compare a given set of thresholds. This is leveraged for the FSE threshold design.

B. Threshold Design for FSE $T_s/2$

Now, consider the problem of designing thresholds for slicers spread across two sampling phases separated by half a symbol period i.e. an FSE $T_s/2$ architecture. We now have two parallel discrete channels, \mathbf{h}_1 and \mathbf{h}_2 . Fixing the total budget of slicers to M , suppose that we fix M_1 , the number of slicers placed at the first phase (so that $M - M_1$ are placed at the second phase), then the threshold values can be computed using exactly the same machinery as earlier. We then optimize by searching over the values of M_1 . The results for TSE are then a special case corresponding to $M_1 = 0$ or $M_1 = M$, and indeed, in several examples, it turns out that allocating all available slicers to one of the two sampling phases results in the lowest cost. For instance, for channel B, it is best to put all the 7 slicers at sampling phase 0 ($\mathbf{h}_{B,0}$). When we increase the number of slicers to $M = 9$ a 7-2 split configuration turns out to be the best, but it is only marginally better than having all 9 at $\mathbf{h}_{B,0}$. This makes sense, since in this case the sampling phase 0 is a good choice. For channel C we find that 2-1 is the optimal split.

C. Simulations

For the three example channels, we present the BER attained using the non-uniform threshold ADCs designed by our algorithm in Figures 4 and 5(a). We simulated 1000 runs, each with 10000 bits to get BER in the range of 10^{-6} to 1. We compare performance with ‘off-the shelf’ uniform threshold ADCs and the Lloyd-Max quantizer [32], [33] which picks thresholds to minimize the quantization error. Both the Lloyd-Max quantizer and our algorithm generate thresholds that depend on the noise level. We design for a nominal SNR of 20dB: we find that the threshold values are relatively insensitive to SNR once the latter is high enough. In all the cases we observe that our non-uniform ADC design performs the best. It is considerably better than uniform ADC and consistently better than Lloyd-Max. The latter optimizes quantization error rather than BER, and hence yields inconsistent performance. In certain cases it is almost as good as our design (channel B), while sometimes it performs almost as poorly as the uniform design (channel A, channel C phase 0). In Fig. 4 we also plot the approximate truncated upper bound P_{ut} (Eq. 26) which our algorithm is trying to minimize. Even though it is an approximate (and rather loose) upper bound, it seems to follow a shape similar to the BER curves, and the benefit of minimizing it gets translated to the actual BER. FSE design leads to the largest gains for channel C. We consider sampling phases 0 and 0.5 and a budget of $M = 3$ slicers (2 slicers are enough for this channel to ensure no error floor, see Table I). We find that the optimal configuration is a 2-1 split (Fig. 5(d)). We notice a 2dB (1dB) gain compared to using a TSE non-uniform architecture at the sampling phase 0.5 (0) (Fig. 5(a)).

Error in Channel Estimation: We have assumed perfect knowledge of the channel impulse response in our designs and performance evaluations. This is a reasonable assumption

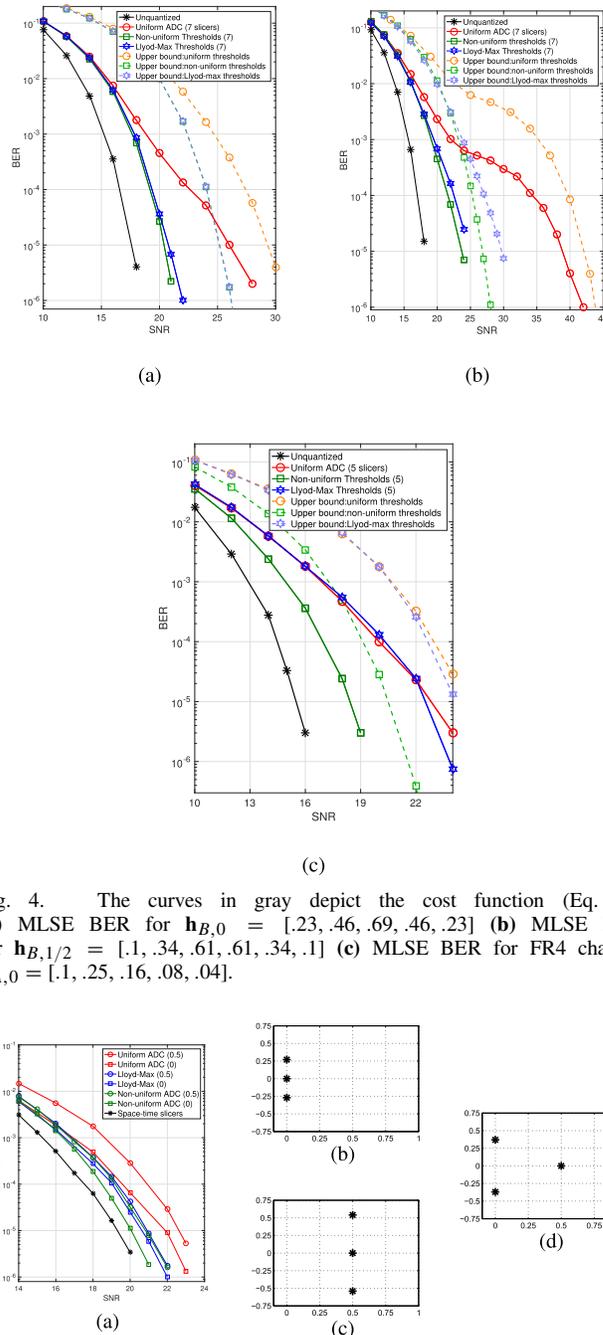


Fig. 4. The curves in gray depict the cost function (Eq. 26) (a) MLSE BER for $\mathbf{h}_{B,0} = [.23, .46, .69, .46, .23]$ (b) MLSE BER for $\mathbf{h}_{B,1/2} = [.1, .34, .61, .61, .34, .1]$ (c) MLSE BER for FR4 channel $\mathbf{h}_{A,0} = [.1, .25, .16, .08, .04]$.

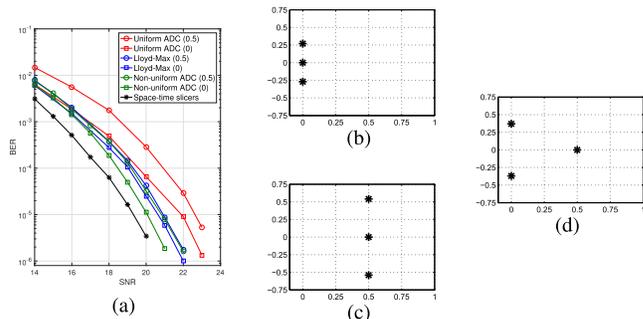


Fig. 5. (a) Bit error rate curves for channel C with sampling phases 0 and 0.5 and a budget of 3 thresholds (b) Non-uniform ADC thresholds at $t = 0$ (c) Non-uniform ADC thresholds at $t = 0.5$ (d) Optimal space-time slicers configuration.

to make, especially for static backplane channels. As shown in [11], channel taps can be estimated very accurately even with low precision ADCs. In particular, it was shown that even with 1-bit ADCs a channel MSE (mean square error) in the range of -10 to -20 dB can be obtained using a pilot signal of duration $N = 10L$ bits. Furthermore, the MSE is reduced by 10 dB for every 10 times increase in N . We performed simulations by taking channel MSE into account and observed that the BER performance remains essentially the same if MSE is below -50 dB. A pilot sequence which is 10^6 bits long (i.e. 0.1ms long for a 10 Gbps link) would easily achieve

channel MSEs lower than -50 dB. Hence it is safe to assume, as we have done in this paper, that the channel is known at the receiver with high fidelity.

Our overall observation is that TSE with channel-optimized thresholds significantly outperforms the standard uniform ADC. The additional gain obtained by generalizing to FSE depends on the channel and the sampling phase. Of course, the trends might be quite different if BCJR decoding is replaced with lower-complexity algorithms. For example, for continuous-valued observations, FSE is much better than TSE for linear equalizers, but is typically only marginally better with BCJR decoding.

VII. CONCLUSIONS

We have shown that there is significant scope for improving on generic ADC designs for communication over dispersive channels when there are severe quantization constraints. Specifically, we propose choosing slicer thresholds for *analog-to-information conversion*, by effectively discriminating between waveforms corresponding to different bit sequences. In addition to choosing slicer thresholds as a function of the channel, spreading slicers over time can improve upon Nyquist rate sampling. We have shown that there are no error floors when we take this concept to an extreme, with one-bit comparators dispersed uniformly over time. We have also provided an algorithm for choosing slicer thresholds for TSE and FSE (sampled at twice the symbol rate), which yields designs that significantly outperform the standard Nyquist-sampled uniform ADC. In summary, our results show that, despite the increased dynamic range due to channel dispersion, it is possible to significantly reduce the number of slicers (which is taken as a proxy for the power consumption of the data conversion front end), while recovering the information encoded in the received signal.

While our focus here has been on communication-theoretic considerations, investigation of detailed cost/power tradeoffs in the context of specific circuit designs is an important topic for future work. From a circuit designer's point of view, it is essential to trade off complexity and power consumption of the analog front end and the associated digital backend (e.g., using fewer slicers in the analog front end may require complex digital processing). For example, while the BCJR algorithm provides a useful benchmark (and may be viable for short backplane channels with antipodal signaling), it is of interest to reduce the complexity of the digital equalizer, and to design the analog-to-information converter accordingly. In particular, it is of interest to explore if we can improve performance relative to prior attempts along these lines based on linear transmit filters and DFE [14]–[16], possibly using a judicious combination of the simplicity of the DFE with the more comprehensive exploration of sequence space obtained using more complex MLSE/BCJR algorithms. This becomes more important when we extend our framework to larger constellations.

We assume the availability of channel estimates in our design. As shown in [11], such estimates can be obtained even with coarsely quantized observations for time invariant channels. It is of interest to explore how best to extend

such approaches to time-varying channels. Finally, while our starting point here is the flash ADC architecture, it is of interest to explore whether the concept of analog-to-information conversion can be effectively applied to obtain more power-efficient designs starting from the pipelined or successive approximation register architectures.

APPENDIX A BCJR ALGORITHM

The BCJR algorithm relies on a Markov structure [22], and applies directly to quantized observations with Nyquist sampling. For faster sampling, the noise correlation can still be handled by state extension if the observations are unquantized [34], but the Markov structure is destroyed by quantization. Thus, for FSE/space-time architectures, we simply ignore the noise correlations, so that the BER attained is an upper bound on the minimum possible BER.

For TSE, the state at time k is $S_k = \{b_k, b_{k-1}, \dots, b_{k-L+2}\}$. From (4), the observation $x(k)$ is a function of S_{k-1} , S_k and the noise $w(k)$. The standard BCJR equations for the posterior probability of the state are given by

$$\begin{aligned} p(S_k | \mathbf{x}_0^N) &\propto p(S_k | \mathbf{x}_0^k) p(\mathbf{x}_{k+1}^N | S_k, \mathbf{x}_0^k) \\ &= p(S_k | \mathbf{x}_0^k) p(\mathbf{x}_{k+1}^N | S_k) \\ &= \alpha_k \beta_k \end{aligned} \quad (29)$$

Forward Recursion:

$$\alpha_k = p(S_k | \mathbf{x}_0^k) = \sum_{S_{k-1}} p(x_k | S_k, S_{k-1}) p(S_k | S_{k-1}) \alpha_{k-1} \quad (30)$$

Backward Recursion:

$$\beta_k = p(\mathbf{x}_{k+1}^N | S_k) = \sum_{S_{k+1}} \beta_{k+1} p(x_{k+1} | S_k, S_{k+1}) p(S_{k+1} | S_k) \quad (31)$$

Note that, for i.i.d. binary signaling, the only computation required is of $p(x_k | S_k, S_{k-1})$, since $p(S_k | S_{k-1}) = 0.5$. From (4), (5), the likelihood of the observation given the states is given by *Continuous Observations*:

$$p(x(k) | S_k, S_{k-1}) \propto \exp\left(\frac{-1}{2\sigma^2} \|x(k) - \mu\|^2\right) \quad (32)$$

Quantized Observations:

$$p(x_q(k) | S_k, S_{k-1}) = \mathcal{Q}\left(\frac{l - \mu}{\sigma}\right) - \mathcal{Q}\left(\frac{u - \mu}{\sigma}\right); \quad l \leq x(k) \leq u \quad (33)$$

where $\mu = \langle \mathbf{h}, \mathbf{b}_k^{k-L+1} \rangle$. The quantized observation $x_q(k)$ is specified via the interval $[l, u]$. $\mathcal{Q}(\cdot)$ denotes the standard normal Q -function. Note that \mathbf{b}_k^{k-L+1} is specified completely via S_k and S_{k-1} . Note that MLSE using the Viterbi algorithm [23] can be run in similar fashion, since it also involves the same core computation of the observation likelihoods (33). Since we are ignoring noise correlations, the preceding approach extends directly to FSE with quantization.

APPENDIX B
PROOF OF LEMMA 1

To prove the lemma, we utilize bounds on information rate derived by Zeitler *et al.* [13], which are valid for both unquantized and quantized measurements, assuming i.i.d. bits and symbol spaced sampling (independent noise samples).

Lower Bound

$$I(\mathbf{b}, \mathbf{z}) \geq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_i I(b_i, \mathbf{z}_i^{i+L-1} | \mathbf{b}_{i-L+1}^{i-1})$$

$$\stackrel{\text{stationarity}}{=} I(b_i, \mathbf{z}_i^{i+L-1} | \mathbf{b}_{\text{past}}) \quad (34)$$

$$= H(b_i) - H(b_i | \mathbf{z}_i^{i+L-1}, \mathbf{b}_{\text{past}})$$

$$= 1 - H(b_i | \mathbf{z}_i^{i+L-1}, \mathbf{b}_{\text{past}}) \quad (35)$$

Upper Bound

$$I(\mathbf{b}, \mathbf{z}) \leq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_i I(b_i, \mathbf{z}_i^{i+L-1} | \mathbf{b}_{i-L+1}^{i-1}, \mathbf{b}_{i+1}^{i+L-1})$$

$$\stackrel{\text{stationarity}}{=} I(b_i, \mathbf{z}_i^{i+L-1} | \mathbf{b}_{\text{past}}, \mathbf{b}_{\text{future}}) \quad (36)$$

$$= H(b_i) - H(b_i | \mathbf{z}_i^{i+L-1}, \mathbf{b}_{\text{past}}, \mathbf{b}_{\text{future}})$$

$$= 1 - H(b_i | \mathbf{z}_i^{i+L-1}, \mathbf{b}_{\text{past}}, \mathbf{b}_{\text{future}}) \quad (37)$$

Here \mathbf{z} denotes measurements at the symbol rate: $\mathbf{z} = \mathbf{x}$ (unquantized), $\mathbf{z} = \mathbf{x}_q$ (quantized). The lower bound is the average mutual information between a bit (b_i) and the set of observations it affects (which are \mathbf{z}_i^{i+L-1}), conditioned on the *past* bits ($\mathbf{b}_{\text{past}} = \mathbf{b}_{i-L+1}^{i-1}$). If we further condition on the *future* bits ($\mathbf{b}_{\text{future}} = \mathbf{b}_{i+1}^{i+L-1}$) we get the upper bound.

We set the noise variance to zero, and consider the normalized channel $\mathbf{g} = (g_1, \dots, g_L)^T$ with $g_j \geq 0$ for all j . Setting $i = 0$ without loss of generality, let $\mathbf{y} = \mathbf{z}_0^{L-1} = \mathbf{x}_0^{L-1}$ denote the portion of the continuous-valued output containing contributions from b_0 :

$$y(j) = \dots + g_{j-1}b_1 + g_j b_0 + g_{j+1}b_{-1} + \dots$$

or

$$\mathbf{y} = G_p \mathbf{b}_{\text{past}} + G_f \mathbf{b}_{\text{future}} + b_0 \mathbf{g} \quad (38)$$

where G_p and G_f are appropriately defined matrices of size $L \times (L-1)$.

In order to derive the lower bound N_l , consider the upper bound (37) on information rate. Let \mathbf{y}_{+1} denote the value of \mathbf{y} conditioned on $b_0 = +1$ and \mathbf{y}_{-1} denote the corresponding value for $b_0 = -1$. Conditioned on the past and future bits, $\Delta \mathbf{y} = \mathbf{y}_{+1} - \mathbf{y}_{-1} = 2\mathbf{g}$. Since $\|\mathbf{g}\|_1 = 1$, each output sample $y(j)$ is confined to $[-1, 1]$ (since the input bits are from ± 1). For a uniform ADC with N thresholds covering this range, the size of each quantization bin is $\frac{2}{N+1}$. If the thresholds separate even one component of $\Delta \mathbf{y}$, we can distinguish between $b_0 = +1$ and $b_0 = -1$, and the conditional entropy term in (37) is zero. This happens if N is large enough that the bin size is smaller than the biggest separation, given by $\max_k 2g_k$:

$$\frac{2}{N+1} \leq \max_k (2g_k) \Rightarrow N \geq \frac{1}{\max(\mathbf{g})} - 1 \quad (39)$$

If N is smaller than the preceding value, it is easy to see that there is at least one set of values for the past and future bits (e.g., set them all to one) for which $b_0 = +1$ and $b_0 = -1$ cannot be distinguished.

For deriving N_u , we consider the lower bound (35) on the information rate. Conditioned on the past bits, the possible values of the components of \mathbf{y}_{+1} and \mathbf{y}_{-1} are given by

$$y_{+1}(j) = \dots + g_{j-2}b_2^l + g_{j-1}b_1^l + g_j + g_{j+1}b_{-1} + \dots$$

$$y_{-1}(j) = \dots + g_{j-2}b_2^k + g_{j-1}b_1^k - g_j + g_{j+1}b_{-1} + \dots$$

where the superscripts l and k are used to denote that the future bits b_1, b_2, \dots need not be the same. The *minimum* value of $y_{+1}(j)$ and the *maximum* value of $y_{-1}(j)$ are given by

$$y_1^*(i) := \min_{\mathbf{b}_{\text{future}}^i} y_1(i) = - \sum_{t=1}^{i-1} g_t + g_i + g_{i+1}b_{-1} + \dots$$

$$y_{-1}^*(i) := \max_{\mathbf{b}_{\text{future}}^i} y_{-1}(i) = \sum_{t=1}^{i-1} g_t - g_i + g_{i+1}b_{-1} + \dots$$

We have an open eye at sample j if $y_{+1}^*(j) - y_{-1}^*(j) > 0$, which happens if $2(g_j - \sum_{t=1}^{j-1} g_t) \geq 0$. If there is a threshold between $y_{+1}^*(j)$ and $y_{-1}^*(j)$, then we can separate $b_0 = +1$ and $b_0 = -1$ irrespective of the value of the future bits. This corresponds to the following condition on N :

$$\frac{2}{N+1} \leq 2 \left(g_j - \sum_{t=1}^{j-1} g_t \right) \Rightarrow N \geq \frac{1}{g_j - \sum_{t=1}^{j-1} g_t} - 1 \quad (40)$$

We get a set of upper bounds on N for each $j = 1, \dots, L$, along with a corresponding set of bounds for the time-reversed channel. Minimizing across these gives the bound N_u stated in the lemma.

REFERENCES

- [1] B. Murmann. "ADC performance survey 1997-2015," accessed on Apr, 2016. [Online]. Available: <http://web.stanford.edu/~murmann/adcsurvey.html>
- [2] A. Varzaghani *et al.*, "A 10.3-gs/s, 6-bit flash ADC for 10G Ethernet applications," *IEEE J. Solid-State Circuits*, vol. 48, no. 12, pp. 3038–3048, Dec. 2013.
- [3] V. H.-C. Chen and L. Pileggi, "An 8.5 mW 5GS/s 6B flash ADC with dynamic offset calibration in 32nm CMOS SOI," in *Proc. Symp. VLSI Circuits (VLSIC)*, Oct. 2013, pp. C264–C265.
- [4] S. Dasgupta and A. Gupta, "An elementary proof of the Johnson–Lindenstrauss Lemma," Int. Comput. Sci. Inst., Berkeley, CA, USA, Tech. Rep. 006–99, 1999.
- [5] E. N. Gilbert, "Increased information rate by oversampling," *IEEE Trans. Inf. Theory*, vol. 39, no. 6, pp. 1973–1976, Nov. 1993.
- [6] S. Shamai (Shitz), "Information rates by oversampling the sign of a bandlimited process," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1230–1236, Jul. 1994.
- [7] A. Kumar, P. Ishwar, and K. Ramchandran, "High-resolution distributed sampling of bandlimited fields with low-precision sensors," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 476–492, Jan. 2011.
- [8] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3629–3639, Dec. 2009.
- [9] J. Singh and U. Madhow, "On block noncoherent communication with low-precision phase quantization at the receiver," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2009, pp. 2199–2203.
- [10] A. Wadhwa and U. Madhow, "Blind phase/frequency synchronization with low-precision ADC: A Bayesian approach," in *Proc. 51st Annu. Allerton Conf. Commun. Control Comput. (Allerton)*, Oct. 2013, pp. 181–188.

- [11] O. Dabeer and U. Madhow, "Channel estimation with low-precision analog-to-digital conversion," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Apr. 2010, pp. 1–6.
- [12] X. Chen, Z. Yu, S. Hoyos, B. M. Sadler, and J. Silva-Martinez, "A sub-Nyquist rate sampling receiver exploiting compressive sensing," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, no. 3, pp. 507–520, Mar. 2011.
- [13] G. Zeitler, A. C. Singer, and G. Kramer, "Low-precision A/D conversion for maximum information rate in channels with memory," *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2511–2521, Sep. 2012.
- [14] E.-H. Chen, R. Yousry, and C.-K. Yang, "Power optimized ADC-based serial link receiver," *IEEE J. Solid-State Circuits*, vol. 47, no. 4, pp. 938–951, Apr. 2012.
- [15] R. Narasimha, M. Lu, N. R. Shanbhag, and A. C. Singer, "BER-optimal analog-to-digital converters for communication links," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3683–3691, Jul. 2012.
- [16] R. Narasimha, G. Zeitler, N. Shanbhag, A. C. Singer, and G. Kramer, "System-driven metrics for the design and adaptation of analog to digital converters," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Oct. 2012, pp. 5281–5284.
- [17] A. Singer, A. Bean, and J. W. Choi, "Mutual information and time-interleaved analog-to-digital conversion," in *Proc. Inf. Theory Appl. Workshop (ITA)*, 2010, pp. 1–5.
- [18] Y. Lin *et al.*, "A study of BER-optimal ADC-based receiver for serial links," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 63, no. 5, pp. 693–704, May 2016.
- [19] O. Abari, F. Chen, F. Lim, and V. Stojanovic, "Performance trade-offs and design limitations of analog-to-information converter front-ends," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Aug. 2012, pp. 5309–5312.
- [20] P. Daponte, S. Rapuano, L. De Vito, and I. Tudosa, "Challenges for aerospace measurement systems: Acquisition of wideband radio frequency using analog-to-information converters," in *Proc. IEEE Metrol. Aerosp. (MetroAeroSpace)*, Oct. 2014, pp. 377–382.
- [21] G. Zeitler, "Low-precision quantizer design for communication problems," Ph.D. dissertation, Inst. Commun. Eng. (LNT), Tech. Univ. Munich, München, Germany, 2012.
- [22] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.
- [23] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.
- [24] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw-Hill, 2000.
- [25] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3498–3508, Aug. 2006.
- [26] R. D. Gitlin and S. B. Weinstein, "Fractionally-spaced equalization: An improved digital transversal equalizer," *Bell Syst. Tech. J.*, vol. 60, no. 2, pp. 275–296, Feb. 1981.
- [27] G. Zeitler, "Low-precision analog-to-digital conversion and mutual information in channels with memory," in *Proc. 48th Annu. Allerton Conf. Commun. Control Comput. (Allerton)*, Oct. 2010, pp. 745–752.
- [28] D. Achlioptas, "Database-friendly random projections," in *Proc. 29th ACM SIGMOD-SIGACT-SIGART Symp. Principles Database Syst.*, 2001, pp. 274–281.
- [29] A. Wadhwa, U. Madhow, and N. Shanbhag, "Space-time slicer architectures for analog-to-information conversion in channel equalizers," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 2124–2129.
- [30] U. Madhow, *Fundamentals of Digital Communication*, vol. 518. New York, NY, USA: Cambridge Univ. Press, 2008.
- [31] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, vol. 1. Berkeley, CA, USA, 1967, p. 14.
- [32] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.
- [33] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.
- [34] A. Kavčić and J. M. F. Moura, "The Viterbi algorithm and Markov noise memory," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 291–301, Jan. 2000.



communication.



communications, signal processing and networking, with current emphasis on millimeter wave communication, and on distributed and bio-inspired approaches to networking and inference. He is a recipient of the 1996 NSF CAREER award, and a co-recipient of the 2012 IEEE Marconi prize paper award in wireless communications. He has served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON INFORMATION THEORY, and the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY. He has authored two textbooks *Fundamentals of Digital Communication* (Cambridge University Press, 2008) and *Introduction to Communication Systems* (Cambridge University Press, 2014).



Naresh R. Shanbhag (F'06) was with AT&T Bell Laboratories, Murray Hill, on VDSL chip-sets, from 1993 to 1995. In 2000, he Co-founded and served as the Chief Technology Officer of Intersymbol Communications, Inc. (acquired by Finisar Corporation, Inc., in 2007), a semiconductor start-up providing DSP-enhanced mixed-signal ICs for electronic dispersion compensation of OC-192 optical links. He has held visiting faculty positions with the National Taiwan University in 2007, and with Stanford University in 2014. He is the founding Director of the Systems On Nanoscale Information fabriCs Center, a five-year multi-university center Funded by DARPA and SRC under the STARnet Program from 2013 to 2017. He is currently the Jack S. Kilby Professor of Electrical and Computer Engineering with the University of Illinois at Urbana-Champaign. He holds 12 U.S. patents. His research interests are in the design of energy-efficient systems, architectures, and integrated circuits for machine learning, signal processing, and communications. He has over 200 publications in this area. He received the 2010 Richard Newton GSRC Industrial Impact Award, the 2006 IEEE Journal of Solid-State Circuits Best Paper Award, the 2001 IEEE TRANSACTIONS ON VLSI Best Paper Award, the 1999 IEEE Leon K. Kirchmayer Best Paper Award, the 1999 Xerox Faculty Award, the Distinguished Lectureship from the IEEE Circuits and Systems Society in 1997, the National Science Foundation CAREER Award in 1996, and the 1994 Darlington Best Paper Award from the IEEE Circuits and Systems Society. He has served as the General Chair of the IEEE Workshop on Signal Processing Systems in 2013, and the IEEE International Symposium on Low-Power Design in 2012, the Technical Program Co-Chair of the 2010 ISLPED, and served on the Technical Program Committee of a number of conferences, including the International Solid-State Circuits Conference from 2007 to 2011. He has served as an Associate Editor for the IEEE JOURNAL ON EXPLORATORY SOLID-STATE COMPUTATION DEVICES AND CIRCUITS from 2014 to 2016, the IEEE TRANSACTION ON CIRCUITS AND SYSTEMS: PART II from 1997 to 1999, and the IEEE TRANSACTIONS ON VLSI from 1999 to 2009.

Aseem Wadhwa received the bachelor's degree in electrical engineering from IIT Delhi in 2009, and the master's and Ph.D. degrees in electrical and computer engineering from the University of California (UC) at Santa Barbara, Santa Barbara, CA, USA, in 2011 and 2014, respectively. He was a Post-Doctoral Researcher with Prof. Upamanyu Madhow with UC at Santa Barbara from 2015 to 2016. He is currently with Apple Inc., Cupertino. He is interested in machine learning, neuroscience inspired learning, and ADC constrained high speed

Upamanyu Madhow received the bachelor's degree in electrical engineering from the IIT Kanpur, Kanpur, in 1985, and the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign, Champaign, in 1990. He was a Research Scientist with Bell Communications Research, Morristown, NJ, USA, and as a Faculty with the University of Illinois at Urbana-Champaign. He is currently a Professor of Electrical and Computer Engineering with the University of California at Santa Barbara, Santa Barbara. His research interests broadly span