

Mutual Information and Time-Interleaved Analog-to-Digital Conversion

Andrew Singer, Andrew Bean, Jun Won Choi
Department of Electrical and Computer Engineering
University of Illinois at Urbana Champaign
Urbana, IL 61822, USA
Email: {acsinger,ajbean,jwchoi}@illinois.edu

Abstract—Analog to digital conversion is often a critical component of a digital communication link. However, the figures of merit that are used in the design of the components that comprise this step are more appropriate for signal reconstruction applications than for digital communication. This paper considers the design of time-interleaved analog-to-digital converters using the mutual information between the transmitted symbols and the outputs of the paths of the converter as a design criteria. Specific attention is paid to the sensitivity of the mutual information through the converter as a function of the relative sampling phases of the time-interleaved samplers. Mutual information is evaluated for a variety of converter strategies, channel conditions, and noise sources. It is shown that for oversampling converters, the optimal sampling phases not in general equispaced, as is conventionally assumed in converter design.

I. INTRODUCTION

A great deal of care is taken in the design of the modulation and coding for modern digital communication links, where the tools used for such designs often focuses on metrics of performance that naturally align with the problem at hand, namely the bit error rate (BER) of the link. However some of the critical components in many such designs include circuit and system components that are often designed targeting other metrics of performance, such as the signal to noise plus distortion ratio, or the total harmonic distortion[1]. For example, as 10Gb/s transceivers for copper links and optical fibers move rapidly into DSP-based implementations [2], [3], a critical component in the front-end of the receive link has become the analog-to-digital converter (ADC). Examples of such links include optical transceivers that incorporate electronic dispersion compensation for single-mode[3] and multi-mode [4] fiber employing ADCs with increasing resolution followed by maximum likelihood sequence detection circuitry. To increase resolution, digital calibration has been increasingly employed in such high-performance ADCs [6], with an emphasis on minimizing nonlinearities induced by ladder offsets in flash converter designs, and those induced by gain and phase mismatches in time-interleaved ADCs [5]. In[1], a flash converter structure was considered in which the sampling and reconstruction levels for the ADC were adjusted to minimize the link BER, rather than minimizing the harmonic distortion, as is typically used in ADC designs. The result of this study was an architecture that was able to dramatically improve the link BER performance for ISI-dominated links, where

the largest gains were achieved in the low resolution regime, as is typical for links in the 10Gb/s range. An information-theoretic study of the closely-related problem of the AWGN-quantized output channel is taken in[8], [9], [7] exploring the capacity of such links, as well as strategies for reducing converter resolution with minimal impact on communication link performance.

One approach to simplifying the design of such high-speed converter structures is the use of time-interleaved architectures, where a rate $1/T$ converter is constructed by using M rate $1/MT$ converters operating in parallel, each operating on a different phase of an M -phase clock. While each of the sample-and-hold circuits within the M branches in such a design must still have the instantaneous bandwidth of the overall rate, i.e. $1/T$, the design can be relaxed by providing MT seconds for the conversion process of the samples. One of the key challenges in the design of such time-interleaved ADCs is maintaining constant gain and sampling phase across the branches, and as a result, considerable calibration and processing circuitry is employed in such architectures[2]. However, once again, the focus of such extended efforts in calibration are not on minimizing the BER of the communication link, but rather in maximizing the signal to noise plus distortion ratio (SNDR) or in minimizing the total harmonic distortion induced by such offsets, as measured by the spurious free dynamic range, or SFDR. This amounts to assuming that the signals of primary interest to the receiver are sinusoidal signals, and that it is this distortion from a pure sinusoidal tone that is to be avoided. As a result, much effort is spent in ensuring that the relative phases of a time-interleaved ADC are equally spaced and that these relative delays be precisely maintained at these levels. However, for a digital communication link, rather than spending valuable resources within the analog front-end preserving the quality of transmitted tones, these resources might be better spent in attempts to either minimize the overall link bit error rate, or to maximize the information capacity of the link. The goal of this paper, therefore is to begin to analyze the information capacity of such ADCs, with a particular focus on studying the mutual information between the transmitted bits through such a link and the sampled outputs of such a time-interleaved ADC.

The block diagram shown in figure 1 depicts a simple digital communication link with a PAM transmitter and a time-

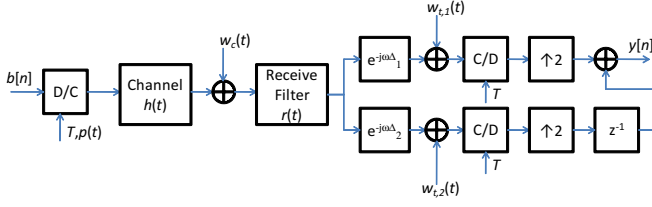


Figure 1. Block Diagram of a digital communication link with a time-interleaved analog to digital converter in the receiver. When the C/D converters each operate at one sample every T seconds (as shown), the front-end is fractionally-spaced; when the converters operate at one sample every $2T$ seconds, the converter is symbol spaced.

interleaved ADC front-end to the receiver. The sequence of transmitted symbols, $b[n]$, are modulated onto the channel as indicated through the D/C converter, such that the transmitter modulates the transmit pulse $p(t)$, as

$$x(t) = \sum_{n=-\infty}^{\infty} b[n]p(t - nT),$$

at a rate of one symbol every T seconds, or $1/T$ symbols per second. The dispersive effects of the communication channel are modeled by the equivalent baseband impulse response $h(t)$, and additive channel noise is modeled by the wide sense stationary process $w_c(t)$. At the receiver, the receive waveform is given by

$$s(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau + w_c(t),$$

which is then processed with the receive shaping filter with impulse response $r(t)$, which also serves to bandlimit (or attenuate the out of band components of) the additive channel noise $w_c(t)$. For example, in the absence of a dispersive channel response $h(t)$, the transmit and receive shaping filters might be selected as square-root raised cosine filters to provide an ISI-free link. The subsequent stages shown to the right in figure 1 depict a two-phase time-interleaved ADC. The received signal is passed through one converter along the top branch, which takes a single sample every T seconds at a relative delay of Δ_1 . The lower branch includes a different delay of Δ_2 seconds prior to sampling at a rate of one sample every T seconds. Within the branch k of the receiver, independent additive thermal noise is modeled as additive stationary white Gaussian noise, $w_{t,k}(t)$, out to the resolution bandwidth of the sample and hold circuitry (technically, thermal noise is only one component out of three dominant sources of noise that contribute to SNR degradation within the IC; these are thermal noise, aperture jitter, and comparator ambiguity, which we lump into a single noise term [10]). These samples are then interleaved to produce a net of two samples every T seconds, or a fractionally-spaced converter. If each of the converters were to operate at a single sample every $2T$ seconds, then the converter would be symbol-spaced, rather than fractionally spaced. This process could be further parallelized to include

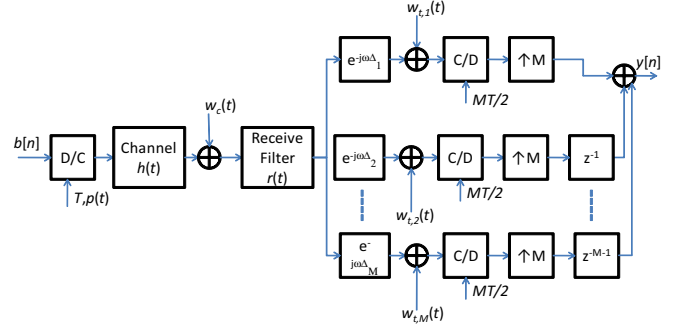


Figure 2. An M -fold time-interleaved receiver structure, shown with symbol-spaced sampling.

M converters, each operating at a rate of one sample every MT seconds, as shown in figure 2.

II. INPUT/OUTPUT MUTUAL INFORMATION

We note that the system shown in figure 2 has a linear input-output relationship. More specifically, consider a finite vector of symbols $\mathbf{b} = (b[0], \dots, b[N-1])^T$ and an observation vector $\mathbf{y} = (y[0], \dots, y[L-1])$. It is possible to relate these input and output vectors according to the relation

$$\mathbf{y} = \mathbf{A}(\mathbf{D})\mathbf{b} + \mathbf{w}_{th} + \mathbf{w}_{ch}. \quad (1)$$

In (1), $\mathbf{A}(\mathbf{D})$ is a channel matrix that depends on the relative sampling time delays of the ADC branches, $\mathbf{D} = (\Delta_1, \dots, \Delta_M)^T$, \mathbf{w}_{th} is a vector of stationary independent Gaussian noise samples that arise from the bandlimited sampling (due to the integration time of the sample and hold circuitry) of $w_{t,j}(t)$, and \mathbf{w}_{ch} is a vector of noise samples that arise from sampling the output of the receive shaping filter due to the channel noise, $w_c(t)$. The matrix $\mathbf{A}(\mathbf{D})$ is derived from the channel model as follows. First, we define the aggregate pulse shape as the three-fold convolution $q(t) = (p * h * r)(t)$. Defining the vector $\mathbf{q}(\mathbf{D}) = [q(\Delta_1), q(\Delta_2), \dots, q(\Delta_M)]^T$, we have that

$$\mathbf{A}(\mathbf{D}) =$$

$$\begin{bmatrix} \mathbf{q}(\mathbf{D}) & \mathbf{q}(\mathbf{D}-T) & \dots & \mathbf{q}(\mathbf{D}-(N-1)T) \\ \mathbf{q}(S+\mathbf{D}) & \mathbf{q}(S+\mathbf{D}-T) & \dots & \mathbf{q}(S+\mathbf{D}-(N-1)T) \\ \mathbf{q}(2S+\mathbf{D}) & \mathbf{q}(2S+\mathbf{D}-T) & \dots & \mathbf{q}(2S+\mathbf{D}-(N-1)T) \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}. \quad (2)$$

The quantity $\frac{1}{S}$ in (2) corresponds with the sampling rate of a single branch of the time-interleaved ADC. To characterize the noise, we assume that each of the noise sources are independent Gaussian random processes. For the thermal noise component, we have that the noise covariance matrix is $\sigma_{th}^2 \mathbf{I}_{L \times L}$. In order to derive the covariance matrix of the channel noise, using the vector of delays $\mathbf{D} = (\Delta_1, \dots, \Delta_M)^T$ and the vector $\mathbf{1}_{n \times m}$, the $n \times m$ matrix of ones, we define the matrix $\mathcal{T}_{\mathbf{D}}$ of pairwise time differences as

$$\mathcal{T}_{\mathbf{D}} = \mathbf{D}\mathbf{1}_{1 \times M} - \mathbf{1}_{M \times 1}\mathbf{D}^T. \quad (3)$$

Let $\sigma_{ch}^2 R(\tau)$ be the autocorrelation function of the filtered channel noise, normalized such that $R(0) = 1$. We now construct a block Toeplitz matrix $\mathcal{T} \in \mathbb{R}^{L \times L}$ of time differences as

$$\mathcal{T} = \begin{bmatrix} \mathcal{T}_{\mathcal{D}} & \mathcal{T}_{\mathcal{D}} - S\mathbf{1}_{M \times M} & \mathcal{T}_{\mathcal{D}} - 2S\mathbf{1}_{M \times M} & \cdots \\ \mathcal{T}_{\mathcal{D}+S\mathbf{1}_{M \times M}} & \mathcal{T}_{\mathcal{D}} & \mathcal{T}_{\mathcal{D}} - S\mathbf{1}_{M \times M} & \cdots \\ \mathcal{T}_{\mathcal{D}+2S\mathbf{1}_{M \times M}} & \mathcal{T}_{\mathcal{D}} + S\mathbf{1}_{M \times M} & \mathcal{T}_{\mathcal{D}} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (4)$$

Then the channel noise covariance matrix is given by $E\{\mathbf{w}_{ch}\mathbf{w}_{ch}^T\} = \sigma_{ch}^2 R(\mathcal{T})$, where the autocorrelation is evaluated element-wise on the matrix \mathcal{T} . Finally, since the channel noise and thermal noise processes are assumed independent, we have that the composite noise covariance is given by $\Sigma = \sigma_{th}^2 \mathbf{I}_{L \times L} + \sigma_{ch}^2 R(\mathcal{T})$, where $\Sigma = E\{(\mathbf{w}_{ch} + \mathbf{w}_{th})(\mathbf{w}_{ch} + \mathbf{w}_{th})^T\}$.

In order to gain some insight into the potential effects of different values of \mathbf{D} on the effectiveness of the ADC samples for communication, we examine the input to output mutual information, per channel use, as a function of the delays \mathbf{D} . As such, we investigate the following:

$$C(\mathbf{D}) \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{b}^N; \mathbf{y}^L) \quad (5)$$

In equation (5), it is implicit that the observation vector \mathbf{y}^L depends on the particular choice of \mathbf{D} , and the number of observations depends on the number of symbols. For simplicity of analysis and computation, we assume that the distribution of input symbols is fixed such that each symbol is an independent Gaussian random value with mean zero and unit variance. We then have that

$$\begin{aligned} C(\mathbf{D}) &= \lim_{N \rightarrow \infty} \frac{1}{N} \left\{ \frac{1}{2} \log_2 \left((2\pi e)^N |\mathbf{A}(\mathbf{D})\mathbf{A}(\mathbf{D})^H + \Sigma| \right) \right. \\ &\quad \left. - \frac{1}{2} \log_2 \left((2\pi e)^N |\Sigma| \right) \right\} \\ &= \lim_{N \rightarrow \infty} \frac{1}{2N} \left(\log_2 |\mathbf{A}(\mathbf{D})\mathbf{A}(\mathbf{D})^H + \Sigma| - \log_2 |\Sigma| \right) \end{aligned} \quad (6)$$

We note that methods for limiting forms of Toeplitz matrices may be used to solve for this limit analytically. However, for our purposes, it is sufficient to closely approximate $C(\mathbf{D})$ as follows:

$$C(\mathbf{D}) \approx \frac{1}{N_2 - N_1} [I(\mathbf{b}^{N_2}; \mathbf{y}^{L_2}) - I(\mathbf{b}^{N_1}; \mathbf{y}^{L_1})]. \quad (8)$$

In order to produce our figures, we chose $N_2 = 120$ and $N_1 = 30$.

Based on our model of the time-interleaved analog-to-digital converter, there are three natural symmetry properties that we expect to observe for $C(\mathbf{D})$. We will refer to these as permutation symmetry, symbol phase symmetry, and branch phase symmetry. Permutation symmetry means that $C(\mathbf{D}) = C(\mathbf{P}_\pi \mathbf{D})$, where π is any permutation, and \mathbf{P}_π is the corresponding permutation matrix. Symbol phase symmetry means that $C(\mathbf{D}) = C(\mathbf{D} + T)$. Finally, branch phase symmetry means that $C(\mathbf{D}) = C(\mathbf{D} + S\mathbf{e}_i)$ for any i ,

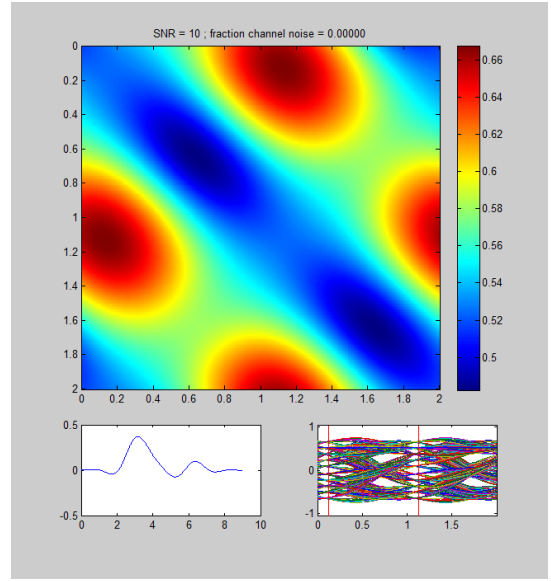


Figure 3. $C(\mathbf{D})$ for symbol-rate sampling at $SNR = 10$ and 0% channel noise.

where \mathbf{e}_i is an $M \times 1$ vector with 1 in the i^{th} position and 0 in every other position.

Figure 3 shows a basic situation for a two-channel time-interleaved ADC in which there is only circuit noise and the signal to noise ratio is 10dB. As of 1999, the overall circuit noise-induced SNR due to input referred-thermal noise, aperture jitter due to uncertainty in the sampling time, and comparator ambiguity due to regeneration time constants from the integrated circuit fabric, was around 20dB for converters in the 10Gs/s regime and would be around 10dB for converters near 40Gs/s [10]. The horizontal axis of the upper subplot is Δ_1 and the vertical axis of the upper subplot is Δ_2 . The setup uses values of $T = 1$ and $S = 2$. Hence, this is a symbol-rate sampling converter. We note that the points of maximum mutual information (per input symbol) lie on the diagonals where $|\Delta_1 - \Delta_2| = 1$, i.e., the optimal sampling scheme is equispaced sampling, i.e. one sample every $T = 1$ second. We can also observe the symmetries mentioned. Permutation symmetry is evident in the reflection symmetry across the $\Delta_1 = \Delta_2$ diagonal. Symbol phase symmetry is also evident, i.e., $C(\mathbf{D}) = C(\mathbf{D} + \mathbf{1})$. Branch phase symmetry is evident in the appearance of periodic boundary conditions in the figures. Furthermore, these symmetries, in combination, result in reflection symmetry across the $\Delta_1 = \Delta_2 \pm 1$ diagonals. The lower left subplot shows the function $q(t)$ for this channel, as defined above. The lower right subplot shows the eye diagram for the channel, as derived from $q(t)$, as well as the optimal sampling points.

Figure 4 shows a similar situation to figure 3, where the difference is that we now have 99% of the noise power coming from noise sources in the channel. Again, the optimal sampling scheme, with respect to the choice of ADC branch delays, is equispaced sampling. The same symmetry properties observed in figure 3 are also visible. What we note here is that while

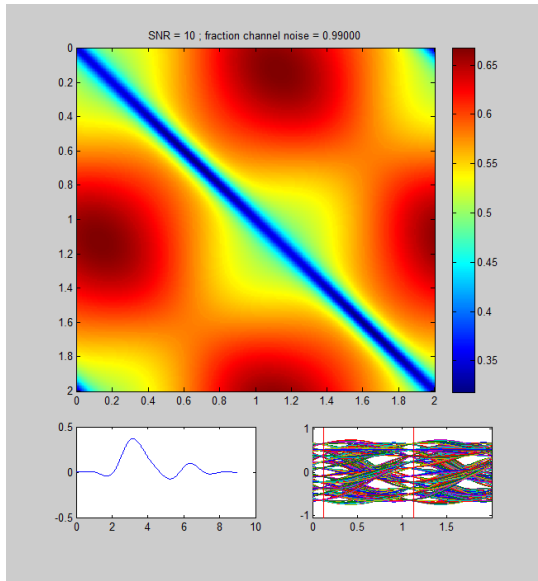


Figure 4. $C(\mathbf{D})$ for symbol-rate sampling at $SNR = 10$ and 99% channel noise.

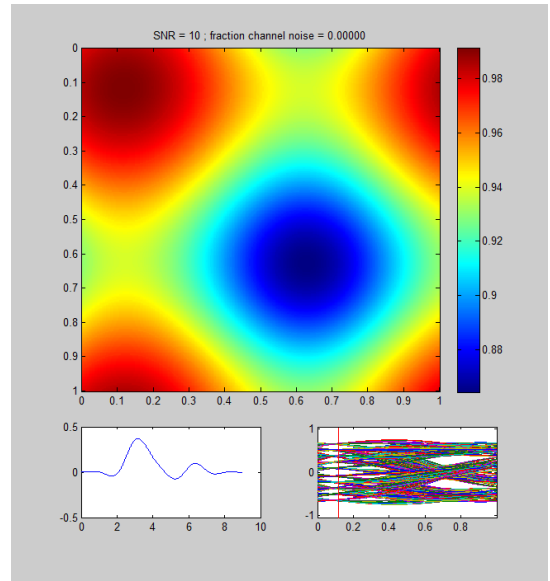


Figure 5. $C(\mathbf{D})$ for oversampling by a factor of 2 at $SNR = 10$ and 0% channel noise.

equispaced samples seem to be the optimal choice within our setup for symbol-rate sampling; irrespective of the choice of channel, SNR, or how much of the noise power is due to channel sources versus circuit noise; we lose very little by imprecisely choosing the ADC branch delays. In fact, by allowing the branch delays to be off by up to plus or minus 10% of a symbol period, we only decrease our achievable data rate by approximately 2% or 3% in the worst case.

Figures 5 and 6 show a setup with oversampling by a factor of 2. The channel is the same as that used to produce figures 3 and 4, we set $SNR = 10$ dB, and $T = 1$. However, corresponding with the oversampling setup, we have that $S = 1$. Figure 5 visualizes $C(\mathbf{D})$ where all of the noise power comes from circuit sources. We observe the symmetries in these figures as well. However, in this case the sample phase symmetry is redundant, since the branch phase symmetry accounts for it when $S = T$. Of particular interest is that the value of \mathbf{D} that maximizes $C(\mathbf{D})$ in this case is on the diagonal where $\Delta_1 = \Delta_2$, which is contrary to the conventional wisdom that the samples should be equispaced.

Now, figure 6 illustrates a channel noise dominated case, where 90% of the noise power comes from the channel. Here, we see that the optimal \mathbf{D} is neither on the equispaced sampling diagonal nor is it on the simultaneous samples diagonal, as in figure 5, where $\Delta_1 = \Delta_2$. Note that when there is no circuit noise present, due to the root-raised cosine filter, an oversampling converter that takes two samples per symbol period contains sufficient information to completely reconstruct the analog output of the channel (including the channel noise). In this limit, the mutual information becomes uniform as a function of the delays, for all pairs (Δ_1, Δ_2) such that $\Delta_1 \neq \Delta_2$.

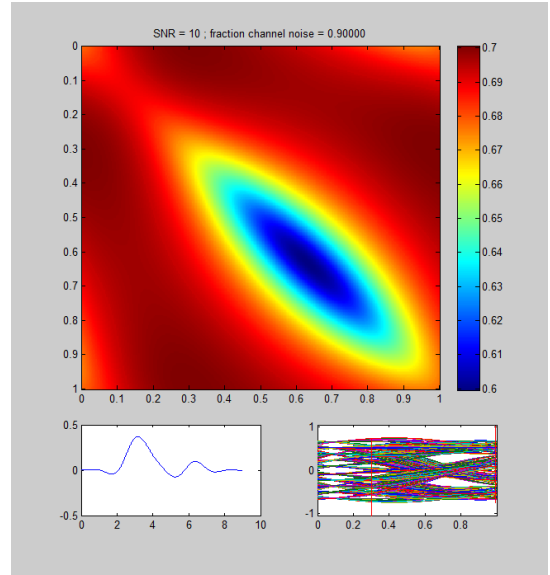


Figure 6. $C(\mathbf{D})$ for oversampling by a factor of 2 at $SNR = 10$ and 90% channel noise.

III. CONVERTER PERFORMANCE

To support the results obtained from the mutual information analysis, we test the bit error rate performance of a receiver wherein turbo equalization is applied to the data sampled by a two-way time-interleaved ADC using the noise model under study in this paper. For data generation, source bits are encoded by a 1/2 rate recursive systematic convolutional (RSC) code with a generator polynomial (3, 5). The encoded bits are permuted by a random interleaver and modulated as BPSK symbols. A root-raised cosine filter with roll-off factor of 0.5 is used for pulse-shaping and the signal is transmitted through the ISI channel whose impulse response is shown

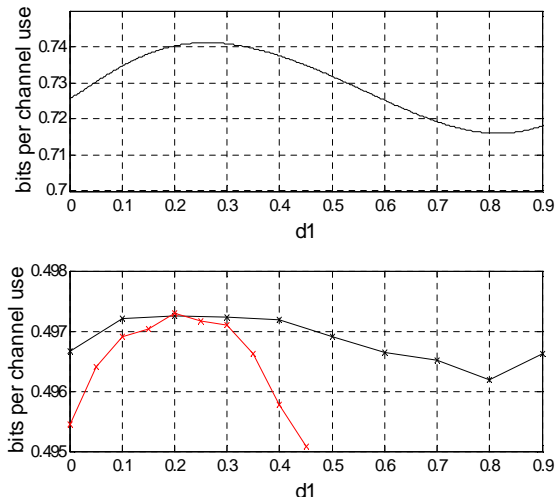


Figure 7. Performance of a turbo-equalization based receiver processing the outputs of a fractionally-spaced, time-interleaved ADC. The average fraction of successfully transmitted bits is shown as a function of the relative delay between the two arms in the converter. Both the known delay case, and the mismatched case are shown.

in Fig. 6. Stationary white Gaussian channel noise is added to the output of the channel, which is then filtered with a root-raised cosine receive shaping filter and then sampled by the time-interleaved ADC. A circuit noise component, modeled by white Gaussian noise is then added to the sampled data. The sequence of the sampled data is processed by a linear minimum mean square error (MMSE) soft-input/soft-output equalizer (SISO) for use in turbo equalization[11]. A SISO max-log-MAP channel decoder is used for turbo equalization iterations. It is assumed that the receiver has a perfect knowledge of the system response including the ISI channel response and the transmit and receive filters as well as exact sampling phase timing (Δ_1, Δ_2) of both branches of the ADC. In the simulations, a total of 50 frames are generated, where each processing frame contains 10,000 coded bits. The BER is measured after the fifth iteration and the ratio of channel noise power to thermal noise power is set to 70%. Rather than BER, figure 7 depicts the average fraction of successfully transmitted bits as a function of delay Δ_2 , for two cases: known Δ_2 and mismatched Δ_2 – *i.e.*, the receiver is designed assuming $\Delta_2 = 0.2$, regardless of its actual value. The value of Δ_1 is fixed at 0 and Δ_2 is varied. As seen in the figure, the location of the minimum BER occurs at the same location at which the mutual information takes on its maximum from figure 6, a slice of which is indicated in the top subfigure, for reference. We note that performance degrades gracefully away from the optimal sampling phases, even in the case of mismatched design, *i.e.* when the receiver has wide uncertainty in the actual value of the sampling phases.

IV. CONCLUSION

This paper considered the design of time-interleaved analog to digital converters for use in digital communication links.

Specifically, the mutual information between the transmitted symbols and the output of the converter at the receiver was used as a guide to measure the optimal relative sampling phases for symbol-spaced and fractionally-spaced receiver structures. It was observed that for symbol-spaced receivers, the conventional approach of equispaced samples coincides with the locations of maximum mutual information. However for oversampling receivers, the relative delays of the arms of the converter were not equispaced, but rather were a function of the relative amounts of channel noise and circuit noise present in the converter samples. Simulations of a link using a turbo-equalization based receiver confirmed that the location of the optimal sampling points observed via mutual information coincided with the sampling phases that achieved the minimum bit error rate.

ACKNOWLEDGMENT

This work was supported in part by the department of the Navy, Office of Naval Research, under grants ONR MURI N00014-07-1-0738 and ONR N00014-07-1-0311 and by the National Science Foundation under grant NSF CCF 07-29092 and by the Gigascale System Research Center (GSRC), one of five research centers funded under the Focus Center Research Program (FCRP), a Semiconductor Research Corporation program.

REFERENCES

- [1] M. Lu, N.R. Shanbhag, and A.C. Singer, "BER-Optimal Analog-to-Digital Converters for Communication Links," *Proc. IEEE International Symposium on Circuits and Systems*, Paris, France, 30 May-2 June, 2010, to appear.
- [2] A. Nazemi, et al., "A 10.3GS/s 6bit (5.1 ENOB at Nyquist) Time-Interleaved/Pipelined ADC Using Open-Loop Amplifiers and Digital Calibration in 90nm CMOS" 2008 Symposium on VLSI Circuits, Technical Digest of Papers, pp. 18-19.
- [3] H. M. Bae, J. Ashbrook, J. Park, N. Shanbhag, A.C. Singer, and S. Chopra, "An MLSE receiver for electronic-dispersion compensation of OC-192 links," *Journal of Solid-State Circuits Conference, Journal of Solid State Circuits*, Volume 41, Issue 11, Nov. 2006, pp. 2541 - 2554.
- [4] O. Agazzi, et al., "A 90nm CMOS DSP MLSD Transceiver with Integrated AFE for Electronic Dispersion Compensation of Multi-mode Optical Fibers at 10Gb/s," *International Solid-State Circuits Conference*, 2008, pp. 232-233.
- [5] C. Grace, P. J. Hurst, S. H. Lewis, "A 12b 80MS/s Pipelined ADC with Bootstrapped Digital Calibration," *ISSCC Dig. Tech. Papers*, pp. 460-461, Feb., 2004.
- [6] P. Nikaeen and B. Murmann, "Digital Compensation of Dynamic Acquisition Errors at the Front-End of High-Performance A/D Converters," *IEEE Journal of Selected Topics in Signal Processing*, 3(3):499-508, 2009.
- [7] O. Dabeer, J. Singh and U. Madhow, "On the Limits of Communication Performance with One-Bit Analog-To-Digital Conversion," In *Proc. IEEE Workshop on Signal Proc. Advances in Wireless Comm. (SPAWC)*, 2006, Cannes, France.
- [8] J. Singh, O. Dabeer and U. Madhow, "Communication Limits with Low Precision Analog-To-Digital Conversion at the Receiver," In *Proc. IEEE Intl. Conf. on Communications (ICC)*, 2007, Glasgow, Scotland.
- [9] J. Singh, O. Dabeer and U. Madhow, "Capacity of the Discrete-Time AWGN Channel Under Output Quantization," In *Proc. IEEE Intl. Symp. on Info. Theory (ISIT)*, 2008, Toronto, Canada.
- [10] R. Walden, "Analog-to-Digital Converter Survey and Analysis," *IEEE J. Select. Areas Comm.*, 17(4):539-550, Apr. 1999.
- [11] M. Tuechler, R. Koetter, and A. C. Singer, "Turbo Equalization: Principles and New Results," *IEEE Transactions on Communications*, vol. 50, no. 5, pp.754-767, May 2002.