SIGNAL RECOVERY IN TIME-INTERLEAVED ANALOG-TO-DIGITAL CONVERTERS

Vijay Divi and Gregory Wornell

Dept. EECS, MIT
Cambridge, MA 02139
{vdivi,gww}@mit.edu

ABSTRACT

Calibration is a serious challenge in the design of high-speed time-interleaved analog-to-digital converters (ADCs). We develop an iterative blind calibration technique for such converters. In particular, an Expectation-Maximize (EM) algorithm is used to estimate the associated unknown gains and time-offsets, from which the calibrated signal is recontructed. Tradeoffs between the calibration time, reconstruction quality, and the oversampling factor are developed. The proposed algorithm can also be used in a variety of other applications, including problems of distributed sampling in sensor networks.

1. INTRODUCTION

Many modern digital systems require the sampling of very high bandwidth analog inputs. Because it has been difficult to scale individual analog-to-digital converters (ADCs) to operate at these speeds, time-interleaved ADCs are an attractive alternative.

Time-interleaved ADCs increase the sampling rate of a system by sending the analog input signal to multiple ADCs in a round-robin manner. As depicted in Fig. 1, this is accomplished by feeding the input simultaneously to the constituent ADCs and having their timings staggered relative to one another. In this way, a system with sampling period of T_s can be realized from M constituent ADCs each operating with a sampling period MT_s. More specifically, the system designer aims for the ith ADC to produce a quantization of $y_i[n] = x(nMT_s + (i-1)T_s)$, where $x(t)$ is the input signal.

In practice, however, there are both variations in gain among the constituent converters, and variations in sampling phase due to signal path length differences on the chip. Without compensation or calibration, these gains and timing offsets seriously degrade overall converter performance. These errors and their effects are increasingly well-understood; see, e.g., [1]. Hardware methods for compensation have been proposed — see, e.g., [2]. However, the analog components involved make such approaches often difficult to use in practice. Other approaches requires the use of a training signal; see, e.g., [3]. With such systems, normal operation must be periodically suspended to perform recalibration, which can be problematic in some applications.

In general, there are two steps to the calibration. The first involves determining the associated gain and timing offset parameters. The second is the compensation based on knowledge of these parameters. The second step is relatively straightforward, particularly in the high signal-to-noise ratio (SNR) regime of primary interest in this paper. Indeed, this step corresponds to a variation on the problem of signal reconstruction from non-uniform samples. As is well known, bandlimited signals can be reconstructed in such scenarios provided the average sampling rate exceeds the Nyquist rate [4]. Since in this application the non-uniform sampling is recurrent, algorithms such as those of Eldar, *et al* [5], can be used to produce the desired uniformly spaced samples. Thus, our primary focus in this paper is the parameter estimation problem of the first step.

This work has been supported in part by Texas Instruments through the Leadership Universities Program, MARCO/DARPA C2S2 under Contract No. 2001-CT-888, and a National Defense Science and Engineering Graduate (NDSEG) Fellowship.

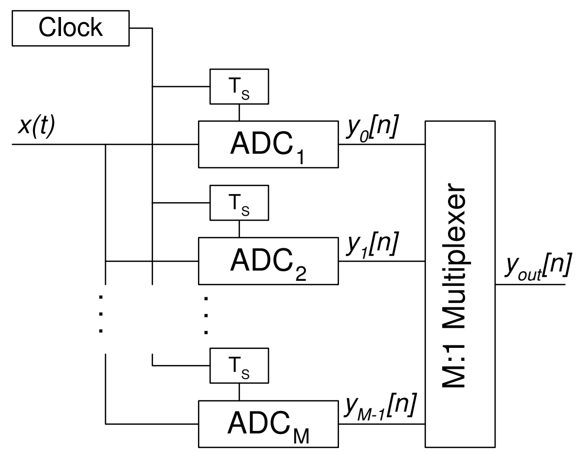

Fig. 1. Time-Interleaved ADC System

We focus on a blind parameter estimation approach that doesn't require knowledge of the input signal. To make this possible, we sample at a overall rate which is faster than the Nyquist rate and take advantage of deliberately introduced excess bandwidth in the signal. One such blind method has previously been developed to perform signal reconstruction [6]. This method uses correlation properties of a slowly varying input in order to determine the timing offsets. However, this method requires the use of a large amount of oversampling to be effective. By contrast, our approach can be effective even with arbitrarily small oversampling factors. For clarity of exposition, we restrict our attention in the development to the case of two interleaved ADCs ($M = 2$), and defer a discussion of the natural generalization to larger numbers of converters until the end of the paper.

2. PROBLEM FORMULATION AND SIGNAL MODELS

We model the converter input $x(t)$ as a stationary bandlimited zero-mean Gaussian process with power spectral density

$$S_{xx}(j\Omega) = \begin{cases} \gamma^2/\Omega_c & \text{for } \Omega \in [-\Omega_c, \Omega_c], \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

The overall sampling period T_s of the system is chosen to ensure that the sampling rate strictly exceeds the Nyquist frequency, i.e., $T_s < \pi/\Omega_c \triangleq T_{\mathrm{N}}$. The signal recovery problem is to estimate $x(nT_s)$, which is bandlimited to $\omega_c = \Omega_c T_s < \pi$, as accurately as possible from the ADC outputs.

We model the output of the ith constituent ADC as

$$y_i[n] = g_i x(2T_s n + \tau_i) + c_i + w_i[n], \qquad i = 0, 1, \quad (2)$$

where the g_i and τ_i model the unknown gains and timing offsets, and where $w_i[n]$ represents the quantization noise, which we model as zero-mean, white Gaussian noise. Its variance σ^2 depends on the number of bits to which the input is quantized. Without loss of generality, we let $g_0 = 1$ and $\tau_0 = 0$, so we may eliminate subscripts on the remaining parameters, i.e., $g_1 = g$, and $\tau_1 = \tau$. Finally, the c_i model unknown DC offsets which are also often present. To simplify our development, we let $c_0 = c_1 = 0$, though their inclusion in the algorithm is straightforward.

3. ITERATIVE PARAMETER ESTIMATION

In this section, we develop an iterative method for producing maximum likelihood estimates of the desired parameter vector $\theta = (g, \tau)$. In particular, we develop an approach based on the Expectation-Maximization (EM) algorithm [7]. This algorithm alternates between two steps:

E-Step

$$Q(\theta, \hat{\theta}^{(n)}) = E[\ln f_Z(z; \theta) \mid Y = y; \hat{\theta}^{(n)}] \quad (3)$$

M-Step

$$\hat{\theta}^{(n+1)} = \arg\max_{\theta} Q(\theta, \hat{\theta}^{(n)}) \quad (4)$$

where $\hat{\theta}^{(n)}$ is the parameter estimate from the nth iteration, and where Y and Z are the incomplete and complete data for the problem, respectively.

The incomplete data are the ADC outputs corresponding to a block of data of length $N_\circ$, i.e.,

$$Y = \{y_0[n], y_1[n], n = 0, 1, \ldots, N_\circ - 1\}.$$

Our complete data for the problem are Nyquist-rate interpolated versions of the individual ADC outputs. In particular,

$$z_0[n] = x(T_{\mathrm{N}} n) + w_0'[n] \quad (5)$$

$$z_1[n] = g\, x(T_{\mathrm{N}} n + \tau) + w_1'[n], \quad (6)$$

with T_{N} again denoting the Nyquist period, and with the $w_i'[n]$ denoting the corresponding bandlimited interpolations of the quantization noises $w_i[n]$. We make the convenient approximation that the $w_i'[n]$ are each white and of variance σ^2. Hence, $Z = \{z_0[n], z_1[n], n = 0, 1, \ldots, N-1\}$, where $N = N_\circ \cdot (2\omega_c)/\pi$.

Our algorithm is most naturally developed in the frequency domain. Accordingly, we work with the discrete Fourier transform (DFT) coefficients of our incomplete and complete data, which we denote using $Y_i[k]$ and $Z_i[k]$, respectively.

Since our complete data vector

$$Z = \begin{bmatrix} Z_0[0] & \cdots & Z_0[N-1] & Z_1[0] & \cdots & Z_1[N-1] \end{bmatrix}^{\mathrm{T}} \quad (7)$$

with $Z_i[k] = (1/\sqrt{N}) \sum_{n=0}^{N-1} z_i[n]\, e^{-j2\pi kn/N}$ is a circularly-symmetric complex Gaussian vector, its log-likelihood function takes the form

$$\ln f_Z(z) = -N\ln(2\pi) - \ln(|\Lambda|) - Z^{\mathrm{H}}\Lambda^{-1}Z \quad (8)$$

$$= -N\ln(2\pi) - N\ln\alpha - \frac{1}{\alpha}\sum_k \Big[(g^2\gamma^2 + \sigma^2)|Z_0[k]|^2 + (\gamma^2+\sigma^2)|Z_1[k]|^2 - 2\,\mathrm{Re}\{g\gamma^2\beta_k Z_0^*[k] Z_1[k]\}\Big] \quad (9)$$

where

$$\alpha = (\gamma^2 + \sigma^2)(g^2\gamma^2 + \sigma^2) - g^2\gamma^4 \quad (10)$$

and

$$\beta_k = \begin{cases} e^{jk\zeta 2\pi/N} & 0 \le k < N/2 \\ e^{j(N-k)\zeta 2\pi/N} & N/2 \le k < N \end{cases} \quad (11)$$

with $\zeta = \tau/T_{\mathrm{N}}$, and where the $2N \times 2N$ covariance matrix $\Lambda = E[ZZ^{\mathrm{H}}]$ of the complete data takes the simple tridiagonal form

$$\Lambda = \begin{bmatrix} * & 0 & 0 & * & 0 & 0 \\ 0 & \ddots & 0 & 0 & \ddots & 0 \\ 0 & 0 & * & 0 & 0 & * \\ * & 0 & 0 & * & 0 & 0 \\ 0 & \ddots & 0 & 0 & \ddots & 0 \\ 0 & 0 & * & 0 & 0 & * \end{bmatrix},$$

with the nonzero entries are indicated by $*$ symbols. More specifically,

$$\Lambda_{i,j} = \begin{cases} \gamma^2 + \sigma^2 & i = j, i < N \\ g^2\gamma^2 + \sigma^2 & i = j, i \ge N \\ g\gamma^2 e^{j\zeta 2\pi j/N} & i = N + j, j < N/2 \\ g\gamma^2 e^{j\zeta 2\pi (j-N)/N} & i = N + j, j \ge N/2 \\ g\gamma^2 e^{-j\zeta 2\pi i/N} & j = N + i, i < N/2 \\ g\gamma^2 e^{-j\zeta 2\pi (i-N)/N} & j = N + i, i \ge N/2 \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

where $0 \le i, j \le 2N - 1$. The inverse of this matrix is also straightforward to obtain via N smaller (2×2) matrix inversions, and the result is also tridiagonal with the same diagonal locations.

The linear mapping from complete to incomplete data is straightforward to develop. In particular, the incomplete data are obtained by downsampling the complete data by a factor of $2\omega_c/\pi$. Since this factor is not generally an integer, this sampling rate conversion is implemented in practice by upsampling, lowpass filtering, then downsampling. In the DFT domain, these steps in the transformation can be expressed as a sequence of matrix mutiplications. Specifically,

$$Y_i = H_{\mathrm{D}} H_{\mathrm{LPF}} H_{\mathrm{U}}\, Zi \quad (13)$$

where

$$Y_i = \begin{bmatrix} Y_i[0] & Y_i[1] & \cdots & Y_i[N_\circ - 1] \end{bmatrix}^{\mathrm{T}} \tag{14}$$

$$Z_i = \begin{bmatrix} Z_i[0] & Z_i[1] & \cdots & Z_i[N - 1] \end{bmatrix}^{\mathrm{T}} \tag{15}$$

For convenience, we let

$$H = \begin{bmatrix} H_{\mathrm{D}} H_{\mathrm{LPF}} H_{\mathrm{U}} & 0 \\ 0 & H_{\mathrm{D}} H_{\mathrm{LPF}} H_{\mathrm{U}} \end{bmatrix} \tag{16}$$

so we can write

$$Y = HZ. \tag{17}$$

It remains only to determine the computations for two steps of the algorithm, which we develop next.

3.1. E-Step

Using (8), the expectation (3) can be readily evaluated and is an affine function of the data. In particular, since the only part of the log-likelihood function that depends on Z is the quadratic form $Z^{\mathrm{H}}\Lambda^{-1}Z = \mathrm{tr}\{ZZ^{\mathrm{H}}\Lambda^{-1}\}$, we obtain

$$\begin{aligned} Q(\theta, \hat{\theta}^{(n)}) = &- N \ln(2\pi) - \ln(|\Lambda(\theta)|) \\ &- \mathrm{tr}\{E[ZZ^{\mathrm{H}} \mid Y; \hat{\theta}^{(n)}]\Lambda(\theta)^{-1}\}, \end{aligned} \tag{18}$$

where the notation $\Lambda(\cdot)$ denotes the covariance function (12) with the parameters values set according to the argument.

Thus, it remains only to evaluate $E[ZZ^{\mathrm{H}} \mid Y; \hat{\theta}^{(n)}]$, which is straightforward. Indeed, since Z and Y are jointly Gaussian, we obtain from routine linear estimation theory [8]

$$E[ZZ^{\mathrm{H}} \mid Y; \hat{\theta}^{(n)}] = \left[I - K(\hat{\theta}^{(n)})H\right] \Lambda(\hat{\theta}^{(n)}) + (\mu^{(n)})(\mu^{(n)})^{\mathrm{H}} \tag{19}$$

where

$$K(\theta) = \Lambda(\theta)H^{\mathrm{H}} \left[H\Lambda(\theta)H^{\mathrm{H}}\right]^{-1} \tag{20}$$

and

$$\mu^{(n)} = E[Z \mid Y; \hat{\theta}^{(n)}] = K(\hat{\theta}^{(n)})Y, \tag{21}$$

with

$$Y = \begin{bmatrix} Y_0[0] & \cdots & Y_0[N_\circ - 1] & Y_1[0] & \cdots & Y_1[N_\circ - 1] \end{bmatrix}^{\mathrm{T}}. \tag{22}$$

The block (and diagonal) structure of the matrices can be exploited to further simplify the computations in the E-Step.

3.2. M-Step

With our setup, the maximization (4) of (18) conveniently decouples. In particular, since the optimizing τ does not depend on g, one can solve for these parameters in order. In particular, the optimization for τ simplifies to

$$\begin{aligned} \hat{\tau}^{(n+1)} &= \underset{\tau \in (0, 2T_s)}{\arg\max}\, Q\left((g, \tau); (\hat{g}^{(n)}, \hat{\tau}^{(n)})\right) & (23) \\ &= \underset{\tau \in (0, 2T_s)}{\arg\max} \sum_k \mathrm{Re}\left\{\beta_k(\tau) E[Z_0^*[k]\, Z_1[k] \mid Y; \hat{\theta}^{(n)}]\right\} & (24) \end{aligned}$$

The notation $\beta_k(\cdot)$ denotes the variable (11) with the parameter value set according to the argument.

The expectation terms in this equation are obtained from the previous E-Step. Although this optimization does not have a simple closed form solution, it is possible to numerically test quantized values in the interval for one that yields the maximum argument.

In turn, the optimization for g takes the form

$$\hat{g}^{(n+1)} = \arg\max_g Q\left((g, \hat{\tau}^{(n+1)}); (\hat{g}^{(n)}, \hat{\tau}^{(n)})\right). \tag{25}$$

It follows that the maximizing value of g is the non-negative solution to the following cubic equation:

$$\begin{aligned} -N(\gamma^2\sigma^2 + \gamma^4)g^3 - S_{01}\gamma^2 g^2 - N(\gamma^2\sigma^2 + \sigma^4)g& \\ + S_1(\gamma^2 + \sigma^2)g + S_{01}\sigma^2 = 0& \end{aligned} \tag{26}$$

where

$$S_1 = \sum_k E[Z_1[k]|^2 \mid Y; \hat{\theta}^{(n)}] \tag{27}$$

and

$$S_{01} = \sum_k \mathrm{Re}\{\beta_k(\hat{\tau}^{(n+1)}) E[Z_0^*[k] Z_1[k] \mid Y; \hat{\theta}^{(n)}]\}. \tag{28}$$

4. ANALYSIS AND SIMULATIONS

Simulations were performed to evaluate the performance of the signal recovery algorithm. The inputs used for testing were realizations of stationary, bandlimited, Gaussian signals. For our tests, we simulated a 12-bit ADC, corresponding to a signal-to-quantization-noise power ratio of 70 dB. We set $\tau = 3T_s/4$ and $g = 1.17$; similar results were produced with other values. The system was initialized with normalized estimates for the unknown parameters: $\hat{\tau}^0 = T_S$ and $\hat{g}^0 = 1$, and reconstruction was performed after 15 iterations of the EM algorithm.

To measure performance, we calculated the effective number of bits the output quality represents, via $(\mathrm{SNR}_{\mathrm{OUT}} - 1.76)/6.02$. Fig. 2 shows the performance of the algorithm as a function of the calibration time, i.e., block size $N_\circ$. Each curve in the graph pertains to a different degree of oversampling, expressed in terms of the "excess number of converters." For example, for an output cutoff frequency of $\omega_c = 0.6\pi$, the signal can be accurately reconstructed with only 1.2 converters, so that with a two-converter system there are 0.8 excess converters.

The upper dashed curve in Fig. 2 depicts the number of effective bits when correct values of gain and time offset parameters are used to reconstruct the signal. The lower dashed curve is equal to the number of effective bits when no calibration is used. Different levels of oversampling are represented through the three solid curves. The top curve represents 0.8 excess converters, middle curve represents 0.6 excess converters, and the bottom curve represents 0.4 excess converters.

As expected, tests show increasing reconstruction quality with calibration time. After $N_\circ = 800$ samples, the system shows a marginal amount of improvement for increased calibration time. Moreover, larger amounts of oversampling reduces the amount of calibration time to reach a target output fidelity. In practice, it is reasonable to use 0.8 excess converters in the 2 ADC system. For long calibration times, this amount of oversampling is ample to estimate gain and offset parameters which produce a reconstructed signal within one bit from the ideal. Reducing the amount of excess converters from this point decreases overall performance, as seen in the 0.4 and 0.6 excess converter cases.

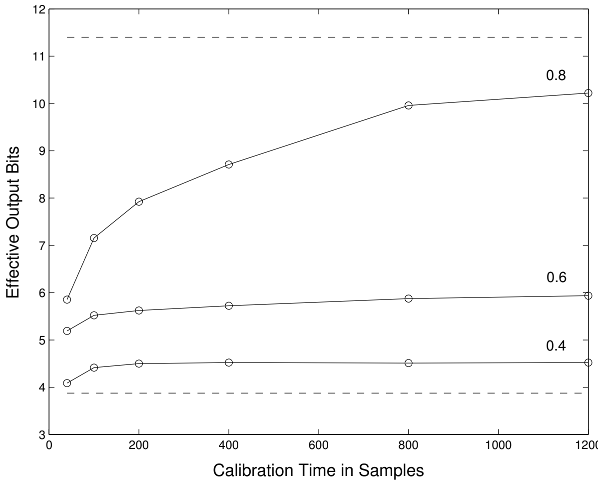

Fig. 2. Reconstructed signal quality, in terms of effective number of bits, as a function of calibration time for 2 time-interleaved 12-bit ADCs. Upper and lower dashed curves correspond to perfectly calibrated and uncalibrated converters, respectively. The solid curves (from top to bottom) correspond to decreasing number of excess converters, i.e., smaller oversampling factors.

The tests presented focused on a finite number of EM iterations. As the amount of excess bandwidth decreased, more iterations than we actually used are necessary for the algorithm to converge. Some simple tests show that increasing the number of iterations yields an increase in the recovery performance in the small excess bandwidth regime. Although we present no analysis at this point, we conjecture that the parameter estimates produced are consistent — that for even arbitrarily small amounts of oversampling, perfect calibration is possible in the long calibration time limit; however, further analysis is necessary to confirm this. Also, the question to whether the likelihood function has local minima instead of a global minimum is unresolved. Our conjecture is that at least in the high SNR regime no local minima exist.

5. CONCLUDING REMARKS

We have developed an attractive blind algorithm for signal recovery in time-interleaved analog-to-digital converter systems. The algorithm is iterative and involves alternating between estimating the input signal covariance and estimating the unknown gain and timing offset parameters. Via this process, we are able to obtain signal reconstructions of the quality one would obtain from precalibrated systems. The tradeoffs achieved between the oversampling factor, calibration time, and reconstruction quality appear promising for practical applications.

While the paper restricted attention to the two-converter case, the iterative algorithm we develop can easily be generalized to work in systems with $M > 2$ ADCs. In this case, the complete and incomplete data sets contain the DFT coefficients of M signals. The E-Step remains the same. The M-Step now computes estimates for each of the τ_i's and g_i's parameters in a sequential manner. Qualitatively similar behaviors are observed in simulations with more ADCs.

It is also straightforward to generalize the algorithm to operate with input signals that are bandlimited but not necessarily spectrally flat. This can be is accomplished by incorporating additional spectral parameters into the problem.

From a broader perspective, this general family of signal reconstruction algorithms has a variety of applications beyond the context of blind time-interleaved ADC calibration. In problems of distributed sampling in sensor networks, for example, the need for such algorithms arises. Indeed, our methods can be applied in any number of data acquisition systems where multiple sampling components are operating asynchronously.

6. REFERENCES

[1] C. Vogel, "Comprehensive, error analysis of combined channel mismatch effects in time-interleaved ADCs," in *Proc. IEEE IMTF*, 2003, pp. 733–738.

[2] K. Dyer, D. Fu, P. Hurst, and S. Lewis, "A comparison of monolithic background calibration in two time-interleaved analog-to-digital converters," in *Proc. IEEE ICASSP*, 1998, pp. 13–16.

[3] H. Jin and E. K. F. Lee, "A digital-background calibration technique for minimizing timing-error effects in time-interleaved ADCs," *IEEE Trans. Circuits, Systems II*, vol. 47, pp. 603–613, July 2000.

[4] F.J. Beutler, "Error-free recovery of signals from irregularly spaced samples," *SIAM Rev.*, vol. 8, pp. 328–335, July 1966.

[5] Y. C. Eldar and A. V. Oppenheim, "Filterbank reconstruction of bandlimited signals from nonuniform and generalized samples," *IEEE Trans. Signal Processing*, vol. 48, no. 10, pp. 2864–2875, Oct. 2000.

[6] J. Elbornsson, F. Gustafsson, and J.-E. Eklund, "Amplitude and gain error influence on time error estimation algorithm for time interleaved A/D converter system," in *Proc. IEEE ICASSP*, 2002, pp. 1281–1284.

[7] N. M. Laird, A. P. Dempster, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Ann. Roy. Stat. Soc.*, pp. 1–38, Dec. 1977.

[8] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Pt. I*, John Wiley and Sons, 1968.