Minimal Total Harmonic Distortion Post-Correction of ADCs

Henrik Lundin, Mikael Skoglund and Peter Händel
KTH Signals, Sensors and Systems
Royal Institute of Technology, SE-100 44 Stockholm, Sweden
Fax: +46 8 790 7260, e-mail: henrik.lundin@s3.kth.se

Abstract— Dynamic digital post-correction of AD converters is considered. A generalized dynamic correction method is proposed and a framework for analyzing the related bit allocation problem is introduced. Finally, this framework is employed in an optimization problem. The solution to the problem indicates which ADC output bits to use in order to maximize the THD in a post-correction system with a constraint on memory size.

The proposed methods are accompanied by exemplary results obtained using experimental ADC data.

Keywords—dynamic correction, optimization, bit allocation.

I. Introduction

This paper studies the problem of digital post-correction of analog-to-digital converters (ADCs). Today, ADCs are widely used in many different applications. A-to-D conversion in radio receivers impose special demands on the converter, and the trend in receiver design has moved the digitization closer to the receiving antenna. Meanwhile, the carrier frequency and the bandwidth is increasing, calling for higher sample rates and increasing analog input bandwidth. Linearity of the ADC is also a key parameter, and the specifications of the system of which the ADC is a part, e.g., signal-to-noise ratio, impose requirements on linearity of the converter. Currently the performance of all-digital receivers, also known as software radio receivers, is limited by the distortion produced by the ADC, and typical dynamic range requirements cannot be met with present commercially available converters [1].

In this paper, a generalized post-correction scheme based on look-up tables is introduced. The configuration of the scheme has a significant impact on the correction results, and an analysis tool is therefore presented. The analysis tool is employed in an optimization problem where total harmonic distortion (THD) is used as cost function.

II. Post-Correction Structure

Several methods to correct ADCs using look-up tables have been proposed over the years, e.g., [2]–[4]. The method used in the present paper was proposed in [5], [6] and is outlined in Fig. 1. The b-bit output word, $x(n)$, from the ADC is concatenated together with K delayed output words ($x(n-1)$ through $x(n-K)$) to form the index (or address) word I of length $B = Kb + b$ bits.

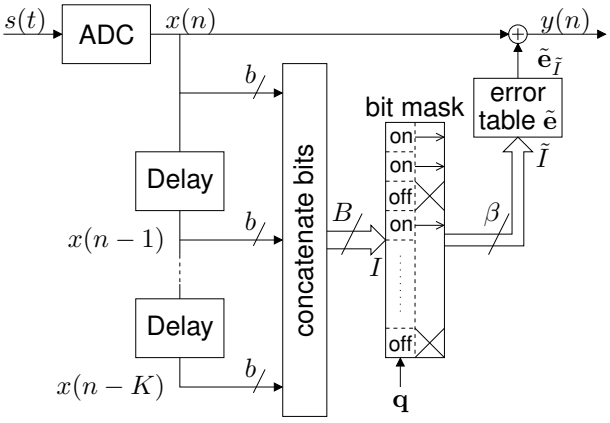

Fig. 1. Dynamic post-correction system outline.

From this index, a reduced size index $\tilde{I}$ is formed by selecting $\beta < B$ bits from I. Which β bits to select is decided by the *bit mask* (this is represented with 'on' and 'off' in Fig. 1). The bit mask vector $\mathbf{q}$ of length B is introduced here. It consists of β ones and the remaining elements zero; a '1' in the i-th position dictates that the i-th bit in I should be propagated to $\tilde{I}$. The new index $\tilde{I}$ is now used to address an error table $\tilde{\mathbf{e}}$ of size 2^β, producing the correction term $\tilde{\mathbf{e}}_{\tilde{I}}$ finally used to form the corrected output $y(n) = x(n) + \tilde{\mathbf{e}}_{\tilde{I}}$. The rationale for the proposed structure is twofold. First, since the ADC errors sought to mitigate are indeed dynamic, any correction scheme should also be dynamic. The proposed method produces *different* correction terms for a given ADC output depending on the K previous samples, and hence it is dynamic. Second, a multi-dimensional table using *all* available bits for addressing will quickly become practically infeasible due to the huge memory requirements ($2^B = 2^{Kb+b}$ entries), especially for high-resolution converters. The proposed method reduces the memory size by using only a subset of the available bits for addressing, but still takes information from K delayed samples into account.

We can immediately identify two special cases: $K = 1$ and a "transparent" bit mask (all bits selected) corresponds to a state-space correction structure, while $K = 0$ yields a "static" correction.

It should be made clear at this point that the bit

mask vector $\mathbf{q}$ is a mathematical construct included to facilitate the performance analysis; it is likely that in a typical implementation of the correction structure the address bit selection would be hardwired, after deciding upon a beneficial configuration.

In order to calibrate the error table $\tilde{\mathbf{e}}$, both the calibration signal $s(t)$ and the resulting ADC output $x(n)$ must be available on the "digital side" of the ADC. Usually this is accomplished by using a reference device, i.e., an ADC with superior characteristics. Alternatively, a digitally generated reference signal could be fed to the ADC under test through a DAC (also with superior characteristics). Both these methods require extra hardware (ADC or DAC). An alternative method is used here. The calibration signal is a sinusoid which can be reconstructed using optimal filtering in the digital domain, thus omitting the need for extra hardware [6].

III. Bit Allocation Analysis

The choice of bit mask configuration, i.e., the allocation of ones and zeros in $\mathbf{q}$, has a *significant effect* on the corrected ADC performance. A mathematical tool to analyze this problem was derived in [6]. The key results from the analysis are reviewed here.

The problem which was analyzed in [6] is the following. Assume that a correction table $\mathbf{e}$ has been calibrated using a specific set of calibration samples (all tables are henceforth considered to be represented as column vectors). The table $\mathbf{e}$ is addressed with an index I, created through concatenation of the present sample $x(n)$ with K delayed samples $x(n-1)$ through $x(n-K)$, but without bit-masking. Thus, the index I is $B = Kb + b$ bits long resulting in $\mathbf{e}$ having $M = 2^B = 2^{Kb+b}$ elements. Assume also that a second table $\tilde{\mathbf{e}}$ has been calibrated with the *same calibration samples*. The table $\tilde{\mathbf{e}}$ is, however, indexed with an index $\tilde{I}$ which is produced by deselecting a number of bits from I. The bit mask vector $\mathbf{q}$ dictates which β bits to select from I, according to the method proposed in Sect. II; $\mathbf{q}$ is said to define a mapping from the integer $I \in [0,\, 2^B - 1]$ to the integer $\tilde{I} \in [0,\, 2^\beta - 1]$. Consequently, the table $\tilde{\mathbf{e}}$ is smaller than $\mathbf{e}$, viz. 2^β elements. The question at hand is: *what is the difference between correcting an ADC using* $\mathbf{e}$ *or* $\tilde{\mathbf{e}}$?

One result from [6] was that a *bit reduction matrix* $\mathbf{R}(\mathbf{q})$ could be used to answer the question. This special matrix can be used to define a table

$$\mathbf{f} = \mathbf{R}(\mathbf{q})\mathbf{e}, \tag{1}$$

the same size as $\mathbf{e}$, but with the special property that the I-th element of $\mathbf{f}$ (i.e., $\mathbf{f}_I$) equals the $\tilde{I}$-th element of $\tilde{\mathbf{e}}$ (i.e., $\tilde{\mathbf{e}}_{\tilde{I}}$) when the index I is mapped to $\tilde{I}$. In other words, $\mathbf{f}$ and $\mathbf{e}$ share the same address space, but $\mathbf{f}$ addressed with I yields the same values as $\tilde{\mathbf{e}}$ addressed with $\tilde{I}$ if I is mapped through the bit mask $\mathbf{q}$ to $\tilde{I}$. Fig. 2 illustrates this relationship in a signal flowchart. The M-by-M matrix $\mathbf{R}(\mathbf{q})$ is [6]

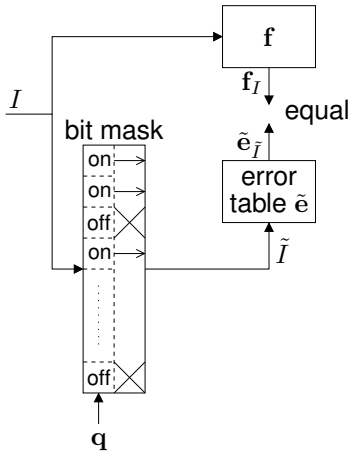

Fig. 2. An illustration of the relationship between the 2^B-size table $\mathbf{f}$ and and the 2^β-size table $\tilde{\mathbf{e}}$.

$$\mathbf{R}(\mathbf{q}) = \mathrm{Diag}\{\mathbf{P}(\mathbf{q})\mathbf{a}\}^\dagger \mathbf{P}(\mathbf{q})\,\mathrm{Diag}\{\mathbf{a}\}, \tag{2}$$

where

$$\mathbf{P}(\mathbf{q}) = \frac{1}{M}\mathbf{H}_B \left(\begin{bmatrix} 1 & 0 \\ 0 & \mathbf{q}_B \end{bmatrix} \otimes \cdots \otimes \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{q}_1 \end{bmatrix} \right) \mathbf{H}_B \tag{3}$$

and $\mathbf{q}_i$ is the i-th element of $\mathbf{q}$. The symbol $\otimes$ denotes the *Kronecker product*, $\mathbf{H}_B$ is a *Hadamard matrix* of size $M = 2^B$, and $\cdot^\dagger$ denotes the *pseudoinverse*.

The observation above can now be used to evaluate the outcome of different bit mask settings *without re-calibrating a new correction table*.

IV. THD-Optimized Bit Mask

Total harmonic distortion (THD) is defined in [7, §4.4.5.1]. The ADC under test should be exercised with a spectrally pure, near full-scale sine wave, $s(t) = A\sin(2\pi f_0 t + \Phi) + C$, with C and A chosen so that the signal is centered within and spans a major part of the ADC's analog input range. The fundamental frequency f_0 is in $[0,\, f_s/2]$ and the initial phase Φ is arbitrary. A record of N successive samples are collected in a vector $\mathbf{x} = [x(0)\; x(1)\; \ldots\; x(N-1)]^{\mathrm{T}}$. The THD is defined as

$$\mathrm{THD} = \frac{1}{N}\sqrt{\sum_{h=2}^{H} |X(f_h)|^2}, \tag{4}$$

where $X(f_h)$ is the discrete Fourier transform (DFT) of the vector $\mathbf{x}$, evaluated at the h-th harmonic of the fundamental frequency f_0. The THD can be expressed in a matrix notation. Let the row vector $\mathbf{w}_h$ be

$$\mathbf{w}_h = [e^{-j2\pi h \frac{f_0}{f_s}\,0}\; e^{-j2\pi h \frac{f_0}{f_s}\,1}\; \ldots\; e^{-j2\pi h \frac{f_0}{f_s}\,(N-1)}], \tag{5}$$

and form a matrix $\mathbf{W}$ as

$$\mathbf{W} \triangleq [\mathbf{w}_2^* \; \mathbf{w}_3^* \; \ldots \; \mathbf{w}_H^*]^*, \tag{6}$$

where $*$ denotes complex conjugate transpose. Then, we can write

$$|X(f_h)|^2 = |\mathbf{w}_h \, \mathbf{x}|^2 = \mathbf{x}^* \, \mathbf{w}_h^* \, \mathbf{w}_h \, \mathbf{x}, \tag{7}$$

and upon inserting (7) and (6) in (4) we get

$$\text{THD} = \frac{1}{N} \sqrt{\sum_{h=2}^{H} \mathbf{x}^* \, \mathbf{w}_h^* \, \mathbf{w}_h \, \mathbf{x}} = \frac{1}{N} \sqrt{\mathbf{x}^* \, \mathbf{W}^* \, \mathbf{W} \, \mathbf{x}}. \tag{8}$$

This is the THD of the *uncorrected* ADC at the frequency f_0.

Assume now that a correction table $\mathbf{e}$ has been calibrated as described in Sect. III, i.e., employing K delay elements but no bit mask. The correction sequence for the recorded output $\mathbf{x}$ can also be described in a matrix notation. Assign a selection matrix $\mathbf{S_x}$ of size N-by-2^B. Each row n of $\mathbf{S_x}$ corresponds to a sample time index $n-1$, and each column I corresponds to a correction table entry $\mathbf{e}_{I-1}$ (row and column indices start at 1). If the table address for the time index n is I, then the element $[\mathbf{S_x}]_{n+1,\,I+1}$ is 1 and the remaining entries in the same row are 0. That is, the matrix $\mathbf{S_x}$ selects the appropriate correction term for each time n. The correction for the sequence $\mathbf{x}$ can thus be written $\mathbf{S_x} \, \mathbf{e}$.

In order to obtain a description for the sample vector $\mathbf{x}$ conforming with the selection matrix notation above, a column vector $\mathbf{r}$ of size 2^B is introduced. Each index I is uniquely decodable to an ADC output level x_j. Thus, if the index I corresponds to the output level x_j, then let $\mathbf{r}_I = x_j$ so that the vector $\mathbf{x}$ can be written $\mathbf{x} = \mathbf{S_x} \, \mathbf{r}$.

Now, we can write the *corrected* ADC output corresponding to the record $\mathbf{x}$ in a new vector as

$$\mathbf{y} = \mathbf{x} + \mathbf{S_x} \, \mathbf{e} = \mathbf{S_x} \, \mathbf{r} + \mathbf{S_x} \, \mathbf{e} = \mathbf{S_x} \, (\mathbf{r} + \mathbf{e}), \tag{9}$$

and, inserting (9) into (8), the expression

$$\begin{aligned} \text{THD}_\mathbf{y} &= \frac{1}{N} \sqrt{\mathbf{y}^* \, \mathbf{W}^* \, \mathbf{W} \, \mathbf{y}} \\ &= \frac{1}{N} \sqrt{(\mathbf{r} + \mathbf{e})^* \, \mathbf{S_x^*} \, \mathbf{W}^* \, \mathbf{W} \, \mathbf{S_x} \, (\mathbf{r} + \mathbf{e})} \end{aligned} \tag{10}$$

is obtained for the resulting THD after correction.

Having established a matrix expression for the THD after correction with the table $\mathbf{e}$, we are now interested in how the THD is affected when an arbitrary bit mask $\mathbf{q}$ is employed. Using the results of Sect. III, the vector $\mathbf{e}$ in (10) should be replaced with the vector $\mathbf{f} = \mathbf{R}(\mathbf{q}) \, \mathbf{e}$ in order to evaluate the effect of reducing the address space. That is, the THD after correction with a table $\tilde{\mathbf{e}}$ calibrated with the same set of calibration signals as $\mathbf{e}$, but this time with a specific bit mask $\mathbf{q}$, is

$$\begin{aligned} &\text{THD}_{\tilde{\mathbf{y}}} = \\ &\frac{1}{N} \sqrt{(\mathbf{r} + \mathbf{R}(\mathbf{q}) \, \mathbf{e})^* \, \mathbf{S_x^*} \, \mathbf{W}^* \, \mathbf{W} \, \mathbf{S_x} \, (\mathbf{r} + \mathbf{R}(\mathbf{q}) \, \mathbf{e})}. \end{aligned} \tag{11}$$

For example, we can evaluate the resulting THD after correction with a state-space table $\mathbf{e}$ versus that of a static table, simply by setting the appropriate bit mask $\mathbf{q}$ in (11).

The function in (11) is the cost function to minimize. The constraint is that a bit mask of β ones and $B - \beta$ zeros is the only allowed solution. The optimization problem for minimizing the THD becomes

$$\begin{cases} \min\limits_{\mathbf{q}} & \text{THD}_{\tilde{\mathbf{y}}} \\ \text{s.t.} & \sum\limits_i \mathbf{q}_i = \beta \quad \text{and} \quad \mathbf{q}_i \in \{0,\, 1\} \quad \forall i, \end{cases} \tag{12}$$

where $\text{THD}_{\tilde{\mathbf{y}}}$ is given by (11).

A. Exemplary Optimization Results

The optimization problem (12) has been solved and evaluated for an exemplary scenario. Experimental ADC data from an Analog Devices AD876 10-bit, pipelined flash converter, designed to operate at a sample rate of 20 MSPS were used. (In the test setup used to acquire the data, the sample frequency f_s was set to 19 972 069 Hz in order to agree with the requirements for coherent sampling [7].) A state-space table is considered in this example. The table, denoted $\mathbf{e}$, is indexed using an index building structure with one delay element (i.e., $K = 1$) and a transparent bit mask (cf. Fig. 1), so that the index I is $B = 20$ bits long and, hence, $\mathbf{e}$ consists of $M = 2^{20} = 1\,048\,576$ entries. The table is calibrated with a large number of different signals, all near full-scale sine-waves but each with a unique frequency. The vector $\mathbf{a}$ represents the number of times each entry in $\mathbf{e}$ was updated during the calibration.

An optimization frequency f_0 is selected and a near full-scale sine-wave record $\mathbf{x}$ of $N = 16\,384$ samples is taken; in the results below the frequency $f_0 = 3\,007\,273$ Hz. The matrices $\mathbf{S_x}$ and $\mathbf{W}$ are formed and the optimization problem (12) is solved for all integers $\beta \in [1,\, 19]$ (the solutions for $\beta = 0$ and $\beta = 20$ are trivial, viz. $\mathbf{q}$ set to all zeros and all ones, respectively). Fig. 3 illustrates the optimal bit masks for different choices of β: each row corresponds to a specific β and the dots indicates which of the positions in the bit mask $\mathbf{q}$ should be set to '1', or, which of the original 20 bits to use in a β-bit index $\tilde{I}$. For example, if a 10-bit index is desired, Fig. 3 suggests that the 6 most significant bits from the present sample $x(n)$, i.e., bits 10 through 5, with 10 being the msb, should be selected, together with bits 8, 7, 5 and 4 from the previous sample $x(n-1)$.

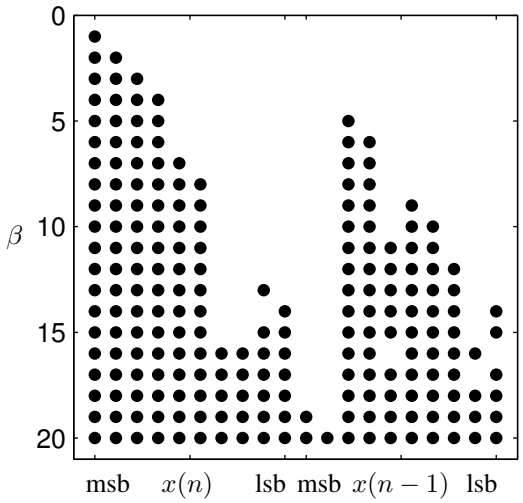

Fig. 3. Exemplary optimization results for THD. Each row corresponds to a specific choice of β, and the dots indicates which positions in the bit mask $\mathbf{q}$ should be set to '1'.

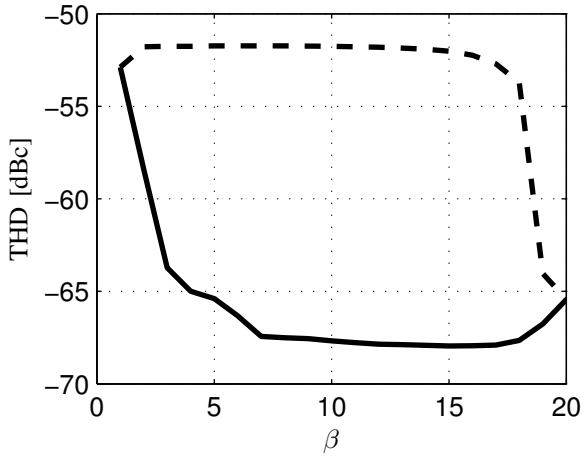

Fig. 4. Resulting THD after correction, optimized for and evaluated at $f_0 = 3.01$ MHz, as a function of table index size β. The solid line is when THD is minimized and the dashed is when THD is maximized.

These 10 bits form the index $\tilde{I}$ used to address the table $\tilde{\mathrm{e}}$ of size $2^{10} = 1024$ entries. Note that Fig. 3 illustrates results for a specific ADC at a specific frequency, and should not be taken to be optimal in general.

In Fig. 4 the resulting THD after correction with a β-bit table is plotted, indexed with the optimal choice of index bits as suggested in Fig. 3. The THD is evaluated at the frequency f_0, i.e., the same frequency as the one for which the index was optimized. At this frequency the *uncompensated* ADC has a THD of –52.9 dB. Somewhat surprisingly, the THD is not minimal at $\beta = 20$ bits, but rather at 15 bits. This phenomenon is due to the fact that in our experiment the amount of calibration data is constant, so that a smaller table will have more calibration data *per table entry*.

Included for reference is the resulting THD when the optimization problem (12) is changed to *maximize* instead of minimize. The plot shows that up to 15 dB of THD can be gained simply by selecting the appropriate indexing bits.

V. Conclusions

A generalized external dynamic correction method for AD converters has been presented. The method introduced the concept of bit-masking, and the problem of selecting a beneficial bit mask was identified. An analysis tool for the bit allocation problem was introduced. With this tool, the outcome of different bit mask configurations can be analyzed. Finally, an optimization problem was posed and solved, where the bit mask was optimized in order to minimize the THD of the corrected ADC. The exemplary results show that selecting an appropriate bit mask significantly affects the performance of the correction, and that the size of the error table can be substantially reduced without causing severe performance degradations.

References

[1] D. Hummels, "Performance improvement of all-digital wide-bandwidth receivers by linearization of ADCs and DACs," *Measurement*, vol. 31, no. 1, pp. 35–45, Jan. 2002.

[2] F. H. Irons, D. M. Hummels, and S. P. Kennedy, "Improved compensation for analog-to-digital converters," *IEEE Trans. Circuits Syst.*, vol. 38, no. 8, pp. 958–961, Aug. 1991.

[3] J. Tsimbinos and K. V. Lever, "Improved error-table compensation of A/D converters," *IEE Proceedings - Circuits, Devices and Systems*, vol. 144, no. 6, pp. 343–349, Dec. 1997.

[4] D. Moulin, "Real-time equalization of A/D converter nonlinearities," in *Proc. IEEE Int. Symp. on Circuits and Syst.*, vol. 1. Portland, OR, USA: IEEE, 1989, pp. 262–267.

[5] H. Lundin, M. Skoglund, and P. Händel, "A framework for external dynamic compensation of AD converters," in *Proceedings 7th European Workshop on ADC Modelling and Testing*, Prague, Czech Republic, June 2002, pp. 135–138.

[6] H. Lundin, "Post-correction of analog-to-digital converters," Royal Institute of Technology (KTH), Stockholm, Sweden, Licentiate thesis TRITA–S3–SB–0324, May 2003, www.s3.kth.se.

[7] *IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters*, Std. 1241, 2000.