<table>
<thead>
<tr>
<th>Title</th>
<th>Linear model-based testing of ADC nonlinearities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Wegener, Carsten; Kennedy, Michael Peter</td>
</tr>
<tr>
<td>Publication date</td>
<td>2004-01</td>
</tr>
<tr>
<td>Type of publication</td>
<td>Article (peer-reviewed)</td>
</tr>
<tr>
<td>Link to publisher's version</td>
<td><a href="http://dx.doi.org/10.1109/TCSI.2003.821281">http://dx.doi.org/10.1109/TCSI.2003.821281</a></td>
</tr>
<tr>
<td>Rights</td>
<td>©2004 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.</td>
</tr>
<tr>
<td>Item downloaded from</td>
<td><a href="http://hdl.handle.net/10468/161">http://hdl.handle.net/10468/161</a></td>
</tr>
</tbody>
</table>

Downloaded on 2019-03-17T17:32:00Z
Linear Model-Based Testing of ADC Nonlinearities

Carsten Wegener and Michael Peter Kennedy

Abstract—In this brief, we demonstrate the procedures of linear model-based testing for the example of a 12-b Nyquist-rate analog-to-digital converter (ADC). In a production test environment, we apply this technique to two wafer lots of devices, and we establish that the model is robust with respect to its ability to reduce the uncertainty of the test outcome. Reducing this uncertainty is particularly beneficial for higher resolution devices, for which measurement noise increasingly corrupts the measured “signal” that is the nonlinearity of the device under test.

Index Terms—Linear modeling, noise-induced test uncertainty, specification test, cost reduction.

I. INTRODUCTION

Testing the static linearity specifications of an N-bit ADC requires one to estimate the integral nonlinearity (INL) at each of the $2^N$ converter codes. Traditionally, this estimate is obtained from an all-codes measurement of the INL characteristic.

When developing a production test for an ADC, a trade-off between the measurement uncertainty and the data acquisition time has to be

Manuscript received January 15, 2003; revised October 14, 2003. This work was supported by Enterprise Ireland and Analog Devices B.V. This paper was recommended by Guest Editors A. Rodríguez-Vázquez, F. Mediero, and O. Feely.

The authors are with the Department of Microelectronic Engineering, University College Cork, Cork, Ireland (e-mail: carsten@ee.ucd.ie). Digital Object Identifier 10.1109/TCSI.2003.821281

One significant bit (LSB), the IEEE Standard 1241 [1] provides the following relationship:

$$N_r \sim \frac{1}{\Delta^2}$$

between the desired value of $\Delta$ and the number $N_r$ of converter samples to be acquired. Thus, in order to halve the uncertainty $\Delta$, the number of acquired samples needs to be quadrupled, which, for high-resolution parts, soon exhausts the typical test time budget [3].

Linear model-based testing has been proposed as an alternative approach to lower the test uncertainty without the need to increase the number of acquired samples [4]. An overview of this approach is depicted in Fig. 1. The all-codes INL measurement result, denoted by a vector $\hat{b} \in \mathbb{R}^N$, is used to estimate the model parameter vector $\hat{x}$ of the device. Based on the estimated model parameter vector $\hat{x}$, the INL characteristic of the device is predicted as $\hat{b} \in \mathbb{R}^N$.

Note that the estimate $\hat{x}$ is obtained by solving a linear system of $2^N$ equations in $n$ unknowns and, in the (typical) case of $n < 2^N$, the solution is obtained in the least-squares sense which provides the noise-reduction capability of model-based testing. Thus, the uncertainty of the predicted all-codes INL characteristic $\hat{b}$ can be lower than the uncertainty of the measured characteristic $b$.

Reducing the uncertainty of the test outcome by model-based testing comes at a computational premium for estimating the model parameters and predicting the all-codes INL. In the $N = 12$-b case study presented in the remainder of this brief, the uncertainty is, by a factor of four, reduced for an additional computation time equivalent to half the data acquisition time. Achieving this reduction in uncertainty of the test outcome without model-based testing, one would need to increase the number of acquired samples by a factor of 16 based on (1).

The contributions of this brief are as follows:

- we quantify the influence of measurement noise on the test outcome and demonstrate the capability of the linear model to reduce this influence;
- we distinguish between contributions to the uncertainty of the predicted INL characteristic by the model’s lack-of-fit and by the measurement noise;
- we develop an approach to reduce the dominant noise-induced contribution; and
- we verify the robustness of the model-based approach in a production test environment.

1For the 12-b test vehicle used here, we have verified this relationship experimentally as reported in [2].
II. TEST UNCERTAINTY AND LINEAR MODELING

As the aim of model-based testing is to reduce the uncertainty of the test outcome, we first quantify the noise-induced uncertainty associated with the test measurements. Then we build a linear model of the manufacturing process-induced INL-causing mechanisms. This model allows us to predict a device’s all-codes INL characteristic from a noisy test measurement. Finally, we quantify the uncertainty associated with this prediction and compare this uncertainty with the previously determined measurement uncertainty.

A. Quantifying Test Uncertainties

Let $b'$ denote a set of test characteristics, where $b' \in B'$ is a column vector whose elements represent the INL characteristic determined for a converter under test; thus, for an $N = 12$-b analog-to-digital converter (ADC), $b' \in \mathbb{R}^{12\times1}$. Furthermore, let $B$ denote a set of reference characteristics with $b \in B$ being the noise-free characteristic corresponding to $b' \in B'$.

A performance test for a min/max-type data sheet specification limit $b_{\pm}$ rejects all devices for which

$$||b'||_{\infty} + \Delta > b_{\pm}$$

is determined. With a guard-band [5] of $\Delta \geq ||b' - b||_{\infty}$, this test implies that the device reference characteristic $b$ meets the data sheet performance specification $||b||_{\infty} < b_{\pm}$. In the following, we quantify the uncertainty $\Delta$ based on

$$e_{\max}(B', B) = \text{mean}_{(b', b) \in (B', B)} (||b' - b||_{\infty}) + 3 \cdot \text{std}_{(b', b) \in (B', B)} (||b' - b||_{\infty}).$$

The initial test implementation produces a histogram of $2^{18}$ converted samples applied to the 12-b ADC under test. From this histogram, the INL characteristic $b' \in \mathbb{R}^{12\times1}$ is determined as described in [6]. In order to obtain a reference characteristic $b$ for the tested device, we perform a second measurement with the number of histogrammed samples increased by a factor of two hundred. For a sample set of $w = 191$ devices, we measure the test and reference characteristics in order to form the columns of two matrices $B$ and $B'$, respectively. This notation allows us to determine $e_{\max}(B', B)$ which characterizes the measurement uncertainty, denoted as $e_{\max} = 260$ mLSB, for the given measurement setup.

Note that the larger the test uncertainty $\Delta$, the more devices which meet the data sheet performance limits $||b||_{\infty} < b_{\pm}$ are rejected based on (2). Thus, it is desirable to reduce the test uncertainty $\Delta$ which, in the case of a test being based on the measured characteristic $b'$, evaluates to $e_{\max} = 260$ mLSB. In order to halve the value of $e_{\max}$, the number of samples in the histogram needs to be quadrupled, which is costly in terms of test time. An alternative approach to reducing the test uncertainty is linear model-based testing described in the following subsection.

B. Linear Modeling

The nonlinearity characteristic of a device is approximated as follows:

$$b \approx A x + b^0$$

where $b^0$ denotes the “nominal” device characteristic, the deviation from which is modeled as a weighted sum of the columns of the model matrix $A$. The weights are the elements of the model parameter vector $x$.

The nominal characteristic $b^0$ is determined as the average of the reference characteristics contained in the sample set of devices, i.e., $b^0$ denotes the row-mean of the matrix $B$. For convenience, we form a matrix $B^0$ of the same shape as $B$, but whose columns represent the nominal device characteristic $b^0$.

With this notation, the model matrix $A$ is derived from the first $n$ columns of $U$ which is a result of singular value decomposition (SVD) [7].

$$B - B^0 \approx U \Sigma \Sigma^T.$$ (5)

This choice of the model matrix minimizes, for a particular value of $n$, the lack-of-fit in the rms norm [7]. However, we are testing for min/max-type specifications and thus the required model order $n$ should be determined based on the lack-of-fit in the maximum-norm $||b'||_{\infty}$. To determine the model’s fitted characteristic $\hat{b}$ for a device characteristic $b \in B$, we derive the model matrix $A$ from $B \setminus \{b\}$, solve (4) in the least-squares sense for $x \in \mathbb{R}^p$, and compute

$$\hat{b} = A \hat{x} + b^0.$$ (6)

Iterating this delete-one-cross-correlation procedure [8] for all the devices $b \in B$, we obtain a matrix $\hat{B}$ of fitted characteristics which allows us to determine the lack-of-fit $\hat{e}_{\max} = e_{\max}(\hat{B}, B)$.

As the lack-of-fit depends on the model order $n$, we graph in Fig. 2 the value of $e_{\max}$ versus $n$. The plot shows that, with increasing model order $n$, the lack-of-fit decreases as the model is increasingly able to represent the INL-causing error mechanisms. However, for $n > 46$, the lack-of-fit levels off; this leads us to choose $n = 46$, for which $e_{\max}$ evaluates to 49.3 mLSB.

In production testing, the model parameter vector can be estimated by solving

$$\hat{b} \approx A \hat{x} + b^0.$$ (7)

in the least-squares sense for $\hat{x} \in \mathbb{R}^p$. The predicted device characteristic $\hat{b} = A \hat{x} + b^0$ is an approximation of $b$, the uncertainty of which we refer to as “prediction error” $\hat{e}_{\max} = e_{\max}(\hat{B}, B)$. For $n = 46$, this error evaluates to 81 mLSB, which establishes the potential of model-based testing to reduce the measurement uncertainty.
\[ \hat{e}_{\text{max}} = 260 \text{ mLSB} \] to the smaller uncertainty \( \hat{e}_{\text{max}} \) associated with the model’s prediction.

### III. Reducing the Prediction Error

There are two primary contributions to the model’s prediction error \( \hat{e}_{\text{max}} \): one is the model’s lack-of-fit \( \hat{e}_{\text{max}} \) and the other is the noise-induced uncertainty of the estimated model parameter estimate which gives rise to the following:

From this notation, it becomes obvious that the prediction error is due to two contributions: one from the lack-of-fit and another from the uncertainty of the model parameter estimate \( \hat{x} \). In this section, we develop an approach to reducing the dominant noise-induced contribution.

#### A. Prediction Error Contributions

For a device, the prediction error is characterized by \((\hat{b} - b)\), which we can also write as

\[ (\hat{b} - b) = A(\hat{x} - x) + \epsilon \]

From this notation, it becomes obvious that the prediction error is due to two contributions: one from the lack-of-fit and another from the uncertainty of the model parameter estimate \( \hat{x} \). Using (4) and (7), we determine that this uncertainty

\[ (\hat{b} - b) \approx A(\hat{x} - x) \]

is merely due to the measurement noise contribution \((\hat{b} - b)\).

In the following, we refer to \( \hat{x} = (\hat{x} - x) \) as the noise-induced uncertainty of the model parameter estimate which gives rise to the noise-induced prediction error \( \hat{e}_{\text{max}} = e_{\text{max}}(A\hat{x}, 0) \), with \( \hat{x} \) being the solution of \((\hat{B} - B) \approx A \hat{X} \). For the model of order \( n = 46 \) derived previously, the value of \( e_{\text{max}} \) is 67 mLSB, which is larger than \( \hat{e}_{\text{max}} \) = 67 mLSB and, therefore, dominates the model’s prediction error \( \hat{e}_{\text{max}} \).

An obvious way of reducing the noise-induced prediction error is to reduce the measurement uncertainty \( \hat{e}_{\text{max}} \). This approach is costly in terms of test time, however, due to the required increase in the number of acquired samples. An alternative is to adapt the model in order to reduce the measurement noise-induced uncertainty of the model parameter estimate.

#### B. Measurement Noise Influence

When estimating the model parameter vector by solving (7) in the least-squares sense, the assumption is that the measurement noise-induced INL values are independent and identically distributed. In this case, the least-squares solver performs optimally in reducing the measurement uncertainty \((\hat{b} - b)\) to the uncertainty \((\hat{x} - x)\) of the model parameter estimate.

This reduction becomes explicit in (9), which we can write for all devices \( b \in B \) as follows: 

\[ (\hat{B} - B) \approx A \hat{X} \]

In order to test the validity of the Gaussian noise assumption on \((\hat{B} - B)\), we perform SVD as follows:

\[ (\hat{B} - B) = U \Sigma V^T \]

and we denote the singular values, i.e., the diagonal elements of \( \Sigma \), as \( \tilde{\sigma}_1, \tilde{\sigma}_2, \ldots, \tilde{\sigma}_w \). The Gaussian-noise assumption leads to an estimate \( \tilde{\sigma}_1 \) for the first (and largest) singular value. With \( m = 4096 \) and \( w = 191 \) denoting the number of rows and columns of \((\hat{B} - B)\), respectively, the estimate \( \tilde{\sigma}_1 \) is given in [9] as follows:

\[ \tilde{\sigma}_1 = \tilde{\sigma} \left( 1 + \frac{w}{m} \right) \]

with \( \tilde{\sigma} = \sqrt{\frac{1}{w} \sum_{i=1}^{w} \tilde{\sigma}_i} \).

In Fig. 3, we show the first ten singular values of the measurement noise matrix \((\hat{B} - B)\) and indicate by a horizontal dashed line the estimate \( \tilde{\sigma}_1 \) of the first singular value.

That the first two singular values \( \tilde{\sigma}_1 \) and \( \tilde{\sigma}_2 \) are significantly above the maximum singular value \( \tilde{\sigma}_1 \) expected for the Gaussian measurement noise assumption indicates that the elements in \((\hat{b} - b)\) are not independent; this degrades the noise reduction achievable when solving (9). In the following, we present a method to whiten the noise by first estimating and subsequently removing these dependencies from \((\hat{b} - b)\) before estimating the model parameters using the weighted least-squares method [10].

#### C. Whitening Measurement Noise

Using SVD in (10), we can extract the two components in the measurement noise matrix \((\hat{B} - B)\) for which the singular values are above \( \tilde{\sigma}_1 \). The associated singular vectors, i.e., the first two columns of \( \hat{U}_r \), are shown in the top and middle plot of Fig. 4.

The columns of \( \hat{U}_r \) that are associated with singular values which are significantly above the noise floor are abbreviated by \( D \). With \( D \) being orthogonal by construction, \( D^T \) is the pseudo-inverse of \( D \). Thus, we can extract from a given measurement noise vector \((\hat{b} - b)\) the non-Gaussian noise components from \((\hat{b} - b)\) results in

\[ b^* = (\hat{b} - b) - D(D^T(\hat{b} - b)) \]

whose elements are independent.

The standard deviation of the elements in \( b^* \) versus code is shown in the bottom part of Fig. 4. The bow-shape can be explained by the pdf of the sinewave histogrammed for deriving the INL characteristics of the devices. At mid-scale of the sinewave, the pdf assumes its minimum, causing a maximum of the measurement uncertainty, i.e., the standard deviation of the noise. At either end of the code range, the pdf of the sinewave approaches its maximum and, thus, minimizes the measurement noise-induced uncertainties.

The method of weighted least-squares [7] equalizes the standard deviation at each code before estimating the least-squares solution for the model parameters. For the \( i \)th element in \( b^* \), we assign

\[ s_i = \begin{cases} 
1, & \text{if } \text{std}(b^*_i) = 0 \\
\frac{1}{\text{std}(b^*_i)}, & \text{otherwise}
\end{cases} \]
and denote $S = \text{diag}(s_1, s_2, \ldots, s_{1024})$. Each element in $b^*$ is scaled by $s_i$, i.e., we multiply $S \cdot b^*$ with the result that the standard deviations of the elements in the column vector $(S \cdot b^*)$ are equalized. Thus, the vector

$$ S \cdot ((\hat{b} - b) - D(D^T(\hat{b} - b))) $$

approximates the white-noise assumption of the least-squares method more closely than the measurement noise vector $(\hat{b} - b)$.

The model extraction method (5) needs to be extended to account for the impact on the model parameter estimation of our method of whitening the measurement noise. Instead of (5), we perform SVD as follows:

$$ \begin{pmatrix} B - B_0 \\ S \cdot [(B - B_0) - D(D^T(B - B_0))] \end{pmatrix} = U \Sigma V^T $$

and extract the model that is an $(8192 \times n)$-matrix from the first $n$ columns of $U$. From this matrix, we form two matrices $\overline{A}$ and $\overline{A}^T$ by taking the top and bottom 4096 rows, respectively.

The noise-induced prediction error is computed as $\tilde{b} = \overline{A} \overline{x}$ based on $\overline{x}$, the least-squares solution of

$$ S \cdot [(\tilde{b} - b) - D(D^T(\tilde{b} - b))] \approx \overline{A} \tilde{x}. $$

For $n = 46$, the noise-induced prediction error evaluates to $\tilde{\epsilon}_{\text{max}} = 43$ mLSB, which is significantly lower than the value of 67 mLSB we obtained without whitening the noise.

On the production line, a device measurement is performed, denoted by $\hat{b}$. The model parameter vector $\tilde{x}$ is estimated by solving

$$ S \cdot [(\hat{b} - b_0) - D(D^T(\hat{b} - b_0))] \approx \overline{A} \tilde{x} $$

in the least-squares sense. Then, the predicted device response $\hat{b} = \overline{A} \tilde{x} + b_0$ is computed.

Since the noise-induced uncertainty of the estimated model parameter vector $\tilde{x}$ is reduced, the prediction error reduces to $\tilde{\epsilon}_{\text{max}} = 64$ mLSB, which constitutes a 20% improvement compared to the 81 mLSB for the prediction error without whitening the measurement noise.

**IV. IMPLEMENTATION IN PRODUCTION TEST**

**A. Test Program Implementation**

For the 12-b ADC, a test program was available that performed the linearity measurement based on histogramming 22 periods of a low-frequency sinewave. Based on this measured INL characteristic, denoted by a vector $\overline{b} \in \mathbb{R}^{1024}$, the predicted device characteristic $\tilde{b}$ is computed. Then, the worst-case INL value, i.e., $\|\overline{b}\|_{\text{max}}$, is determined and, in order to pass or fail the device under test, the test condition (2) is applied as follows:

$$ \|\overline{b}\|_{\text{max}} + \Delta < b_{\pm} $$

with $\Delta = \tilde{\epsilon}_{\text{max}} = 64$ mLSB accounting for the prediction error.

The computations required to derive $\tilde{b}$ from a measurement $\hat{b}$ add, in our implementation, half the time required for taking the measurement $\hat{b}$. Equating the measurement time to $T$, the test time using the model-based approach is $1.5T$, which achieves a reduction of the test uncertainties to $\tilde{\epsilon}_{\text{max}} = 64$ mLSB.

When targeting the same reduction of the test uncertainty by accumulating more samples in the histogram, the measurement time would increase by a factor of $(260/64)^2 \approx 17$, thus resulting in a total test time of $17T$. With our implementation of model-based testing, we obtain a speed-up by a factor of $17T/1.5T \approx 11$ for the same test accuracy of $\Delta = 64$ mLSB.

**B. Verification of Model Performance**

We have applied the modified test program to two wafer lots of approximately 8000 devices each. In each device lot, we found six devices whose measured characteristic $\overline{b}$ violated the customer INL specification limits $b_{\pm} = 1.5$ LSB by more than $\tilde{\epsilon}_{\text{max}} = 260$ mLSB, thus failing the test even before the linear model was applied.

Excluding devices that grossly violate the INL specifications, we show in Fig. 5 histograms of the probability density versus the exhibited most positive and negative INL values. We can compare these histograms to the pdf we obtain for the modeling set by overlaying this pdf as a solid curve in the figure. This comparison indicates that wafer 1 has an INL performance spread closer to the modeling set, while the histogram for wafer 2 indicates a significantly wider spread of the exhibited INL values.

That wafer 2 is “different” from wafer 1 is also indicated by the results for another test that was performed on the devices. In contrast
A linear test performed under dynamic conditions, i.e., with a high-frequency sinusoidal input.

The question arises whether the manufacturing process excursion experienced by wafer 2 degrades the model performance we derived earlier for the modeling set of devices. Such a degradation can be caused by error mechanisms that were not present or were of minor significance in the set of devices used for model building [11].

In order to verify that model performance is sustained, we took a subset of 203 devices from wafer 2 and measured the INL characteristics in the set of devices used for model building [11].

In order to estimate the magnitude of this increase, we determined in Section III-C that $\hat{c}_{\text{max}} = 64$ mLSB. From the measurements of the 203 devices from wafer 2, we determine this prediction error to be 85 mLSB. The gap between these two results can be explained by the uncertainties of the measurements used as the reference $b$.

Recall that, in the model building stage, the reference measurement $b$ was obtained by histogramming 4400 periods of the sinusoidal waveform, i.e., 200 times as many periods as used in production to obtain $b$. Since the measurements for the 203 devices from wafer 2 were taken during the production test run, measurement time was limited, and only 550 sinusoidal wave periods were histogrammed. The smaller number of histogrammed samples leads to significant noise-induced uncertainties in the reference $b$ which tends to increase the numbers we obtain for the lack-of-fit $\hat{c}_{\text{max}}$ and the prediction error $\hat{c}_{\text{max}}$, respectively.

In order to estimate the magnitude of this increase, we determined “reference” characteristics for the modeling set of devices by histogramming 550 periods only. Recomputing the prediction error $\hat{c}_{\text{max}}$ for the model yields 82 mLSB. This value is comparable to the 85 mLSB we obtain for the 203 devices from wafer 2 and, therefore, the results in Table I do not suggest a degradation of the model performance for wafer 2. This sustained model performance is achieved despite the fact that wafer 2 experienced an abnormal manufacturing process excursion that led to a 50% drop in yield.

### Table I

| LACK-OF-FIT $\hat{c}_{\text{max}}$ AND PREDICTION ERROR $\hat{c}_{\text{max}}$ FOR A MODEL OF ORDER $n = 46$ AND REFERENCE MEASUREMENTS |
|---|---|---|
| modeling set $p = 4400$ | modeling set $p = 550$ | wafer 2 subset $p = 550$ |
| $\hat{c}_{\text{max}}$ | $\hat{c}_{\text{max}}$ | $\hat{c}_{\text{max}}$ |
| 49 mLSB | 72 mLSB | 68 mLSB |
| 64 mLSB | 82 mLSB | 85 mLSB |

V. CONCLUSION

For the example of a 12-b Nyquist-rate ADC, we establish that the uncertainty of the test outcome can be reduced by model-based testing. We implemented the test procedure in an industrial test environment and achieve this reduction in one eleventh of the test time it takes with the traditional approach of increasing the number of acquired converter samples. Furthermore, by applying the model-based technique to two wafer lots of devices, we established that the model is robust to significant manufacturing process excursions.

### REFERENCES


