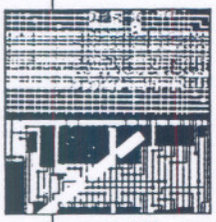
Cutting the high cost of testing

A new modeling approach to the overly long testing of analog and mixed-signal devices saves substantially on time and cost



The responsibility for production-testing a new line of low-cost 13-bit analog-to-digital converters is yours. You must devise a test plan that can correctly sort the devices into performance bins. What do you do? You test the first few to come off the assembly line extensively, examining how accurately the digital outputs correspond to the analog inputs and storing the results, which often deviate from ideal behavior.

Even though you are using the latest automatic test equipment, you notice that testing a 13-bit analog-to-digital converter (ADC) at all its possible output codes requires measuring 8192 (2^{13}) different values of input voltage—a very time-consuming task. And that is just for the room-temperature tests at nominal supply voltage. More thorough testing could take several times longer.

To keep production flowing, you realize that you may have to buy more test stations. But that capital investment would force up the price of the converters, supposedly low-cost devices. You long for a simpler test plan, one that would let you sort those converters accurately into the performance bins without increasing your costs.

Test engineers are constantly faced with that challenge: how to develop test routines that will correctly sort devices at minimum cost. There are inevitable tradeoffs between the expensiveness and thoroughness of the testing; for a given cost, the more complete the testing, the lower the throughput.

Over the last several years, a comprehensive approach that optimizes the tradeoffs associated with production testing of analog and mixed-signal electronic devices has been developed at the National Institute of Standards and Technology (NIST), Gaithersburg, Md. It is based on the fact that the behavior of many devices is governed by a relatively small set of underlying variables, which consequently determine the results of a large number of measurements. In essence, a simple linear coefficient matrix model of the device is set up to relate the (relatively large number of) measured responses to the (relatively small set of) underlying variables.

T. Michael Souders and Gerard N. Stenbakken
National Institute of Standards and Technology

This approach, an extension of the well-studied technique known as optimal design of experiments, is then coupled with the concept of empirical modeling. Although much more computationally efficient than the optimal design technique, the new approach yields nearly as good results. Early evaluations of its use in small-scale commercial experiments indicate its probable utility in situations where the candidate test space is large or otherwise expensive to test exhaustively, and where a rather few underlying parameters affect many aspects of device behavior—as is true with analog ICs.

In addition to testing converters, the approach is being applied successfully to a variety of devices and instruments, including amplifier-attenuator networks, filters, and multirange instruments.

Despite such achievements, however, this approach may not be as effective in other instances. Unless a model is already available, the method is best suited to large production runs where the cost of developing the model and selecting test points can be amortized over a large number of devices. Also, certain types of nonlinear behavior can seriously reduce the efficiency of any linear modeling approach.

LESS IS BETTER. Let us assume that the initial tests done by the test engineer on the first eight devices in our a/d converter example yield the results shown in Fig. 1. For simplicity, a fictitious 7-bit converter is illustrated with 128 (2^7) code states.

Although it is probably not known to the test engineer, and not obvious from the performance plots, the nonideal behavior of the converters is largely determined by rather few semiconductor-processing variables, here assumed to be seven [Fig. 2].

To find a solution for a system with seven variables, seven independent equations, or pieces of information, are required. In Figure 2, the seven curves on the top represent the error signatures of the seven variables.

Each variable is associated with a parameter that affects the behavior of the device in a particular way. For example, parameter a_1 causes the entire response to be offset, whereas a_2 causes a positive offset in the lower half of the response and a negative offset in the upper half.

On the bottom of Fig. 2, the productic run performance of device No. 8 is show to be a linear combination of these seven si natures; the weight of each is the value the corresponding variable. (The process i conceptually similar to the idea behind Fou rier analysis, in that a function is decom posed into a set of differently weighted stan dard functions.)

In this system, each candidate input tes condition, or test point, defines a linea equation; the total error at each point is a linear combination of the seven signatures evaluated at the same test point.

The standard way to test ADCs is to do all-codes testing—run the input over its range so that all possible output codes are generated. A 7-bit converter requires performing at least 128 tests, with 128 separate equations—one for each required value of input voltage. But, since only seven independent equations are needed to solve the system, only seven test points need to be measured to calculate the values of the seven variables. Once those variables are known, the entire behavior of the ADC can be calculated—rather than measured—at every required test point by weighting and summing the seven error signatures.

Therefore, the test engineer really needs to test the converters under only seven conditions to fully characterize them.

Defining terms

Error signature: the characteristic way in which an underlying variable contributes to the total error response of a device.

Integral nonlinearity (INL): a figure of merit for an analog-to-digital converter, equal to the maximum deviation from the ideal input-output curve, not counting the gain and offset errors.

Normalized prediction variance: the ratio of the variance of a prediction to the variance of the measurement noise on which it is based.

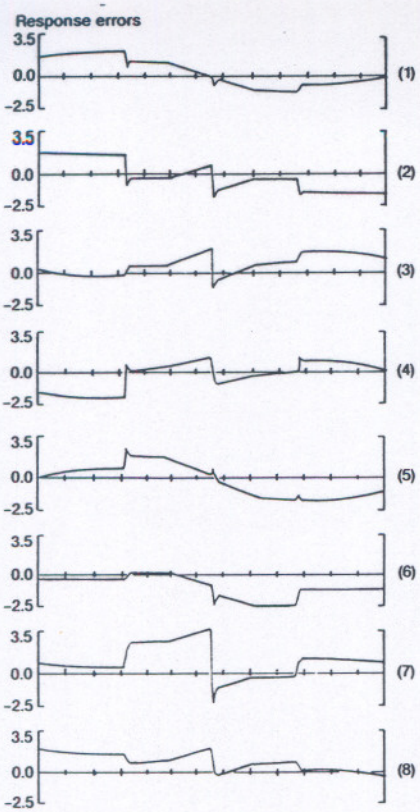
QR factorization: a standard method for factoring a matrix into a right (R) triangular matrix and an orthonormal (Q) matrix—usually done to make machine solutions less subject to computer roundoff errors.

Residual errors (residuals): the part of a device's response that is not described by the model.

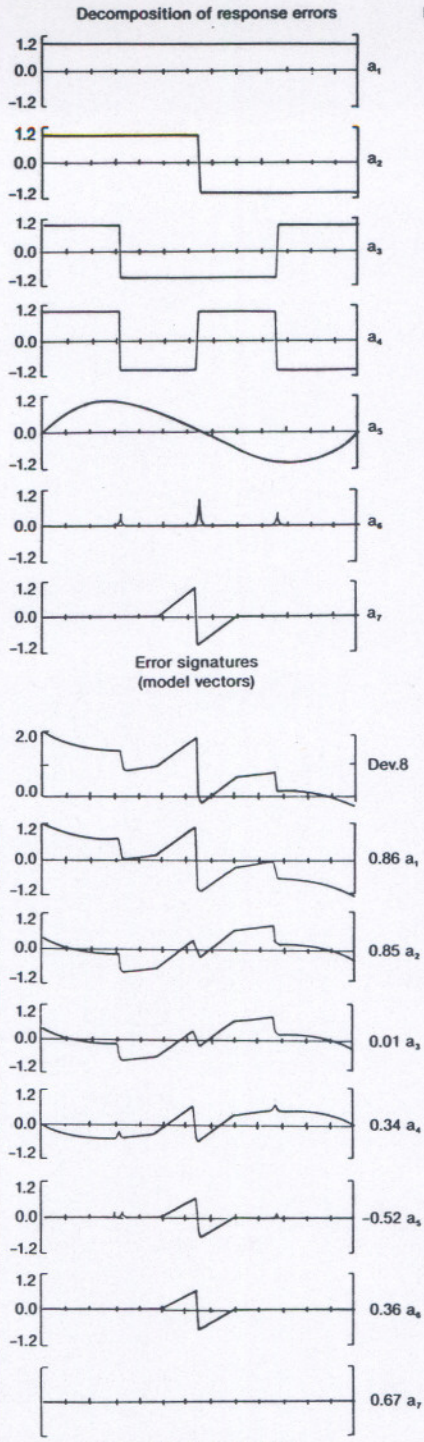
Test point: an input signal or other condition applied to a device under test, to which an ideal response can be predicted; it is also called a test condition or input condition.

Test space: the total range of input variables over which a device is tested.

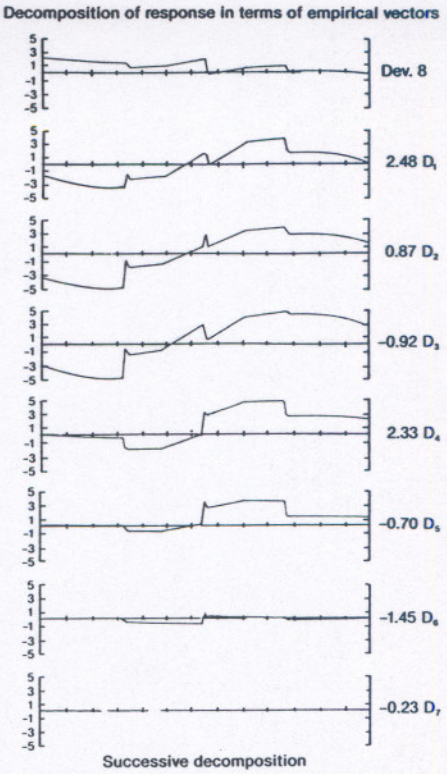
0018-9235/91/003-0048$1.00©1991 IEEE



[1] The performance of a device as a function of one input condition can be illustrated by plotting the response error as a function of the input condition. In this case, the horizontal axis represents the digital output codes of a fictitious 7-bit analog-to-digital converter and the vertical axis is the input error corresponding to that code—the difference between the actual input that produced the digital code and the input that should ideally produce it. The scales are in arbitrary units.



[2] The error response of a device can be decomposed into a weighted sum of error signatures of underlying variables. The error signatures for the analog-to-digital converter example are shown at the top, and the decomposition of the response errors of Device No. 8 of Fig. 1, in terms of those error signatures, is shown on the bottom. The topmost plot on the bottom is the response of Device 8; the other plots show the error that remains after the designated amount of each error signature is successively subtracted.



[3] In addition to calculating error signatures from a known model, one can develop a model empirically—that is, from measurement data. In the case shown, based on the measured responses of Fig. 1, the empirical vector derived from seven of the devices provides a means for describing the behavior of the eighth.

To make this method work, two things are needed: the error signatures (or a matrix model from which they can be determined) and the specific set of test points at which the measurements are to be made.

WHERE TO TEST. Many different sets of seven test points will produce seven independent equatons, but many more will not. To complicate matters further, there are degrees of independence as well. What needs to be done is to find the set of test points that is maximally independent. Discovering that set also makes the process most resistant to the corrupting effects of measurement noise.

QUICK, BUT GOOD. Finding a set of maximally independent test points is handled by the optimal design of experiments process. For large problems, this can be expensive to find; however, nearly optimal solutions can be rather cheap using a mathematical operation called QR factorization (QRF). Computationally efficient implementations of the QRF operation exist in the public domain software called Linpack by the Society of Industrial and Applied Mathematics, Philadelphia, and in its friendlier commercial descendants, such as CLAM from Scientific Computing Associates Inc., New Haven, Conn., and Matlab from Mathworks Inc., North Natick, Mass.

Those routines operate on the matrix model of a device from simple calls to the software package. They return a vector (list) of the selected test points and also provide information on the degree of independence represented by them. From the vector of test points, the system of corresponding simultaneous equations is known. That small system of test points and equations is valid for every device that is adequately described by the original model.

A measure of the prediction errors associated with the selected test points is the normalized prediction variance. This can be computed ahead of time from the original matrix model and the selected test points and then evaluated at every candidate test point given a set of selected test points.

A good selection of test points, therefore, is one that minimizes the prediction variance. If that variance is deemed too high even with a maximally independent set of test points, the error can be further reduced by adding more test points so that more measurements exist than the number of parameters to be estimated.

Of course, adding those test points will re-

sult in an overdetermined system of equations (more equations than variables), but that can be solved using standard least squares techniques. If the additional test points again constitute a "good" selection, the prediction variance will be reduced by the ratio of the number of test points to the number of variables.

Beating down the noise is not the only advantage of selecting more than the minimum number of test points; the redundancy permits model errors to be detected as well. Selecting additional test points enables a least squares solution to be found, allowing the generation of the residual errors of the solution at the measured points. Examination of those residuals can give a good indication of the accuracy of the model: a good model will produce residuals that are randomly distributed and have a standard deviation comparable to that of the measurement noise.

On the other hand, an inadequate model will cause the standard deviation of the residuals to increase, and structure to appear in the distribution.

Once measurements have been made at the selected test points, the system of seven equations is solved, again using standard matrix software routines. The solution gives the actual values of the seven variables for the specific device that was tested.

Referring back to the fictitious ADC example, the entire behavior of the converter at all of the 128 candidate test points is then easy to predict: the seven error signatures are simply weighted by the corresponding values from the solution, and then summed together. The result is the behavior for all test conditions, including the few that were actually measured, and the many that were not. For a 13-bit converter, the savings could be even greater, as will be seen later.

MODELING. Of course, the success of the method depends critically on the quality of the model. So far, it has been assumed that the error signatures for the converters were known. How they are to be determined, and how accurate they must be, are the next questions to be addressed.

In mathematical form, the model really represents the sensitivity of the converter's behavior, at each test condition, to an appropriate set of underlying variables. The variables determine the degrees of freedom available to the devices—the specific ways in which individual units can deviate from their nominal, or ideal, behavior. In some cases, the variables are known from the design and correspond to conventional modeling parameters.

> Be aware that with device complexity, model accuracy degrades

For example, if an accurate equivalent circuit is known, the error model can be computed as the partial derivatives, or sensitivities, of the output response of the circuit with respect to the component parameters, evaluated at their nominal values. (That corresponds to a first-order Taylor expansion of the circuit's response.) Versions of public domain as well as commercial software are currently available for computing such sensitivity matrices. The well-known circuit analysis program Spice 3C, for example, has that capability.

Models derived in that manner are called physical, sensitivity-based models. Their primary virtue is the direct correspondence between the model variables and measurable physical parameters, such as resistance, capacitance, transistor transconductance, and open-loop gain.

But physical modeling is not without problems. As devices get more complex, model accuracy tends to degrade and the computational burden increases. All too often, adequate modeling and computer-aided design (CAD) tools seem to lag behind the technology that requires them. In some cases, detailed design knowledge of the device may simply not be available to the test engineer. In others, a first-order Taylor expansion may be insufficient. That can occur when the device's behavior has a strongly nonlinear dependence on its components, as in testing the frequency response of a multi-pole filter, for example.

EMPIRICISM. For such situations, other modeling techniques are available. Empirical learning-based modeling, in particular, is especially attractive for performance-testing applications like the ADC example. It requires no detailed design knowledge of the device, nor is it strictly limited to linear dependence upon the variables.

In empirical modeling, the models come from the devices themselves. For that reason, they are immune to the sorts of errors that can arise from an imperfect understanding of the workings of a device. After all, what could better express the process variability of a series of widgets coming off a production line than the behavior of the widgets themselves.

If empirical modeling is applied to the converter example and a sample of the devices is fully tested—and all test points examined —the responses of the various devices will vary somewhat. This is because the values of the underlying variables differ from unit to unit. (If that were not true, testing would be unnecessary.)

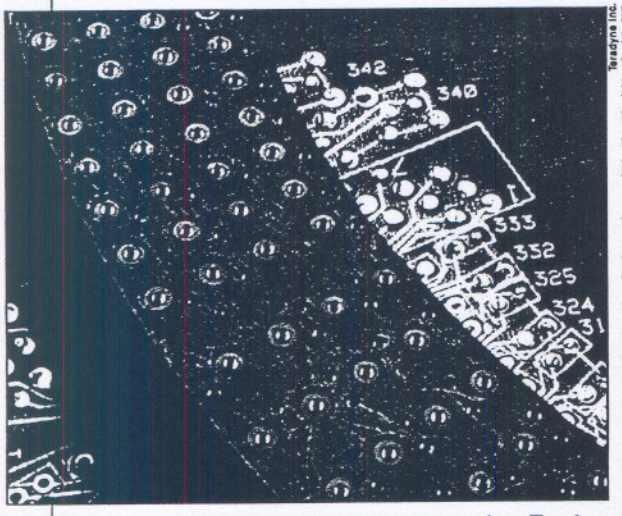
Assuming a stable manufacturing process, a reasonable statistical sampling of devices will embody all the degrees of freedom that the process allows. Using the same QRF routine employed in selecting test points, a subset is chosen of response vectors that are linearly independent. The new matrix composed of this subset will itself be a complete model useful for accurate predictions.

Mathematically speaking, the new matrix spans the space of possible responses, but these in turn are constrained by the variables of the manufacturing process. Figure 3 illustrates how the first seven of the response vectors of Fig. 1 also make up a model that fully describes the response of the other devices (in this case, the eighth) from the production run.

The empirical approach not only eliminates the need for detailed design knowledge of the device under test, but also minimizes the number of variables required. For example, the error signatures of many components of a device will be negligibly small and therefore need not be considered. Other components may have signatures that are identical to each other, such as the components of cascaded gain stages. And still others can have error signatures that are different but always track each other.

That last circumstance is common in ICs. A single processing variable, such as dopant level or exposure time during metalization, affects many of the components equally, causing them all to vary in fixed proportion.

MIX AND MATCH. Interestingly, it is not only possible, but often desirable, to combine the physical and empirical modeling approaches and benefit from the best features of each.



Teradyne Inc.

The power of high-throughput test systems such as Teradyne Inc.'s A500 can be increased manyfold by the application of modeling techniques for designing test procedures.

Perhaps that is best accomplished by starting with the physical parameters that are known to be important and are perhaps trimmable. It is often rather simple to compute the sensitivity to those parameters, even when the overall circuit is unknown.

Next, the sensitivity matrix of the physical model is augmented with empirical vectors. The result will be a model that can be useful, not only for making accurate predictions, but also for estimating the actual values of critical parameters, which can then be trimmed to achieve compliance.

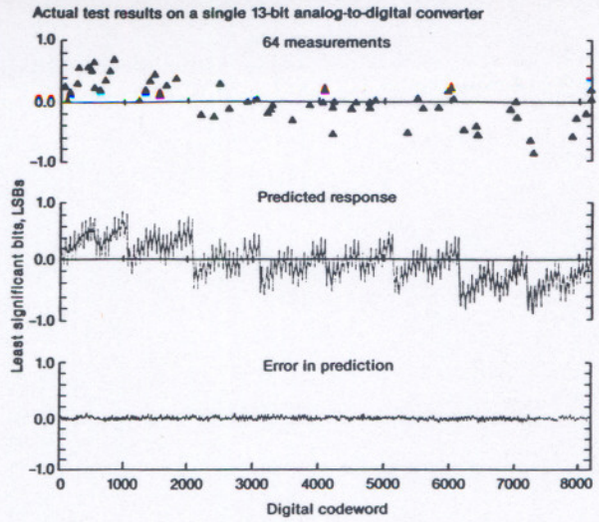
One trial NIST testing program performed in collaboration with Teradyne Inc., a manufacturer of automatic test equipment for the analog and mixed-signal IC industry, combined QR factorization and physical plus empirical modeling. The test was applied to the measurement of integral nonlinearity (INL) for a batch of 127 commercial 13-bit ADCs, all of the same model type. As indicated in our earlier example, the common industry practice for determining the INL of an ADC is exhaustive testing—measuring each of the possible code states to determine the largest error. Even using the fastest available test equipment, that practice adds US $1 or more to the cost of a part that typically sells for only $15. No wonder the industry is looking for a less expensive testing methodology.

For the NIST-Teradyne study, an 18-parameter model of the 13-bit ADC was developed using a combination of physical and empirical modeling techniques. The empirical modeling was based on exhaustive test data obtained from the first 50 devices, which revealed that an 18-vector model sufficed to represent the error space with suitable accuracy. Using that model, 18 test points were chosen, and 46 others were added to obtain redundancy—a total of 64 test points.

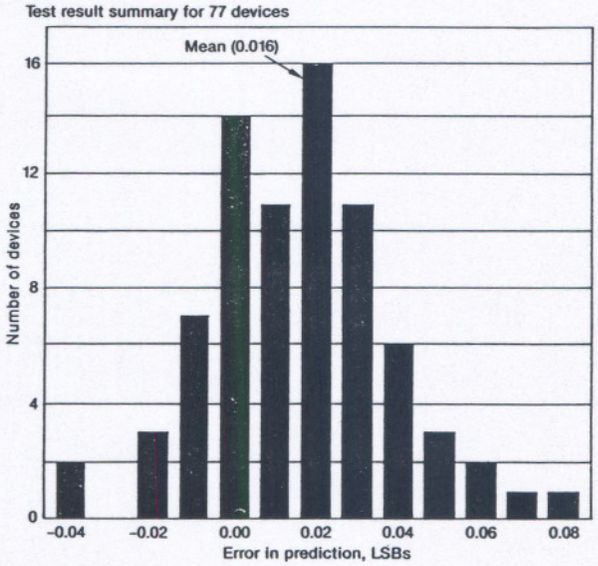
IT WORKS! Measurements at only those 64 (out of a possible 8192) test points were used to predict the overall response of each of the remaining 77 devices. To evaluate the success of the method, the predictions were compared with the results obtained from exhaustive testing [Fig. 4]. The root-mean-square value of the differences was 0.024 least-significant bit (0.0003 percent of full scale), where one LSB is 2^{-13} (0.012 percent of full scale).

Converters such as these are typically sorted according to their maximum INL. The error in predicting that for the 77 devices was also computed [Fig. 5]. (A positive error indicates the predicted maximum is smaller than the measured maximum.) For comparison, the effective noise level in the measurement process obtained by taking the standard deviation of repeated measurements of the same device, was 0.02 LSB.

Since the standard deviation of the predictions is not much greater than that of the measurements, and both are very small, the sorting error rate based on the limited, 64-point test would be similar to that achieved using conventional all-codes testing involving 8192 measurements. With an array processor to speed up the computations, the computational overhead can be kept below 1 second per device. So the test time, which is reduced by a factor of 128, becomes negligible.



[4] Modeling works, as these results for a single analog-to-digital converter make clear. The top plot shows the integral nonlinearity measured at the 64 selected codes. The middle plot gives the predicted errors at all 8192 codes based on the 64 measurements, while the bottom plot gives the error in the predictions—the difference between the measured errors and the predicted errors at all codes. Vertical scales are in least-significant bits.



[5] The value of the new approach is best illustrated by this histogram, which illustrates the differences between the measured and predicted maximum integral nonlinearity for the 77 devices tested. For comparison, the standard deviation of the measurement process—the repeatability of each measurement—is 0.02 least-significant bit.

TO PROBE FURTHER. Various facets of analog and mixed-signal testing strategy have been discussed by the authors and their colleagues. Good surveys of the work are: G.N. Stenbakken and T.M. Souders, "Test Point Selection and Testability Measures Via QR Factorization of Linear Models," *IEEE Transactions on Instrumentation and Measurement*, Vol. IM-36, No. 2, June 1987; and T. M. Souders and G. N. Stenbakken, "A Comprehensive Approach for Modeling and Testing Analog and Mixed-Signal Devices," in the *1990 Proceedings of the International Test Conference*, IEEE Computer Society Press, Los Alamitos, Calif., September 1990.

A three-day workshop is being offered on April 2–4 at the National Institute of Standards and Technology (NIST), Gaithersburg, Md., to provide more in-depth training in the techniques discussed in this article. For more information, call NIST, 301-975-2406.

ABOUT THE AUTHORS. T. Michael Souders (SM) is a career employee of the National Institute of Standards and Technology (NIST), Gaithersburg, Md., where he is concerned with efficient testing strategies for complex systems, among other activities. Souders has a B.S. in physics from the Johns Hopkins University, Baltimore, Md.

Gerard N. Stenbakken (M) has been with NIST for 21 years. He is currently working on test strategies for complex electronic systems and the design of wideband, high-accuracy sampling wattmeters. He has a B.S. in physics from the University of Minnesota in Minneapolis and an M.S. in electrical engineering from the University of Maryland, College Park.

Certain commercial products are identified in this article to illustrate the experimental procedure. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology; nor does it imply that the products identified are necessarily the best available for the purpose.