Journal of Electronic Testing (2022) 38:637–651
https://doi.org/10.1007/s10836-022-06039-2

A Polynomial Transform Method for Hardware Systematic Error Identification and Correction in Semiconductor Multi-Site Testing

Praise O. Farayola[1] · Isaac Bruce[1] · Shravan K. Chaganti[2] · Abalhassan Sheikh[2] · Srivaths Ravi[2] · Degang Chen[1]

Received: 19 July 2022 / Accepted: 21 November 2022 / Published online: 13 December 2022

Abstract
Multi-site measurement (testing) increases throughput and reduces production test costs by simultaneously testing multiple chips. However, as the number of test sites is increased (to maximize throughput further), site-to-site variation in analog and mixed-signal circuits test measurement inevitably increases to levels causing mis-trim and/or misclassification of the device under test (DUT). This work proposes a practical and low-cost approach to effectively identify and correct pronounced site-to-site variation inherent in multi-site test data. Assuming the test hardware is stationary or time-invariant, the measured chip parameter at a site is modeled as a weak nonlinear function of the true parameter for that site. A polynomial transform-based method is proposed to identify this systematic nonlinearity. The approximate inverse function of the identified nonlinearity is then applied to the measurements at the issue sites to remove the effect of the induced hardware systematic errors. This approach is practical and cost-effective as it enables continued use of existing hardware, avoids expensive root-cause analysis, and re-fabrication of multi-site test boards. It improves yield by achieving more accurate chip measurements and reduces test escapes. The accuracy and robustness of the method are confirmed after application to simulated and real-world industrial test data.

Keywords Analog and Mixed-Signal Multi-site Testing · Polynomial Normal Transformation (PNT) · Site-to-Site Variation (SSV) · Yield Optimization

1 Introduction

Increasing complexity in the design and functionality of integrated circuits (ICs) has led to rising production test and measurement needs. Many of these ICs are used in applications requiring stringent requirements, hence the need for extensive testing. These test measurements consume time and increase costs, thus plaguing the semiconductor industry. One of the effective methods used to reduce test time and cost is multi-site testing [22, 25, 29].

Multi-site test, or "parallel test," is a semiconductor automatic test equipment (ATE) term that generally refers to the testing of multiple "devices" at the same time. It is now widely adopted in production wafer probes and packaged die test flows used in the semiconductor industry. The method can be argued to be one of the best techniques for reducing systems on chip (SoCs) test time and increasing throughput [17]. The method was previously dominant in digital testing but has now found a footing in parametric testing for analog and mixed-signal circuits. Multi-site testing can be used during wafer probing, final chip tests, and burn-in tests.

The number of test sites in a multi-site testing system is central to its benefits. The greater the number of sites, the higher the number of chips that can be tested simultaneously. One of the recent trends in the semiconductor industry is the migration.

of successful multi-site test boards to accommodate more test sites. While these migrations maximize throughput, massive multi-site unavoidably escalates site-to-site variations (SSV) in test measurement due to undesired interactions among analog/digital/power signals across denser dies and test hardware [14]. While multi-site testing is also used to test digital logic and memory circuits, site-to-site variation is more critical for analog and mixed-signal circuits because very small measurement errors can lead to yield loss and/or test escape for such products.

Responsible Editor: A. Ivanov

✉ Praise O. Farayola
farayola@iastate.edu

[1] Iowa State University Ames, Iowa, USA

[2] Texas Instruments Inc, Sugar Land, TX, USA

Ideally, the mean and variance of test measurements will vary slightly across test sites to reflect measurement noise, random errors, and intrinsic variations in the DUTs. However, the induced systematic errors become pronounced for some sites (issue sites) and are no longer acceptable. Figure 1 shows an example of site-to-site variations observed in the measured Offset specification of 16-bit Analog to Digital converters (ADCs) tested in a multi-site probe environment. Test data for selected 40 test sites (containing over 13,000 ADCs) were obtained from Texas Instruments, and exact measured values are excluded to protect confidentiality. By visual inspection of the boxplot visualization, Site 10 exhibits an excessive downward shift in the median and quartiles of test measurements compared to other test sites.

Measurement errors are unavoidable in most manufacturing industries and affect yield [30]. Systematic errors carry potential harm and dire consequences, especially in analog and mixed-signal circuit testing. If not carefully monitored, it could lead to yield loss (good DUTs incorrectly labeled as bad) or potential test escapes (bad DUTs incorrectly marked as good). Both mistakes are expensive and can cost semiconductor companies a lot of money. Systematic errors also pose a challenge to DUTs that need trimming after testing. Incorrect trim codes applied to such DUTs may further worsen the performance of the device. Significant site-to-site variation also makes methods developed to identify wafer outliers and monitor the manufacturing process more difficult as measurements do not reflect the true nature of the DUT [26].

Addressing issue sites often requires taking the ATE offline for diagnosis and repair. This industry solution is difficult, time-consuming, and expensive. An alternative solution is to identify the induced systematic errors and correct the test data at test sites with significant variations (issue sites).

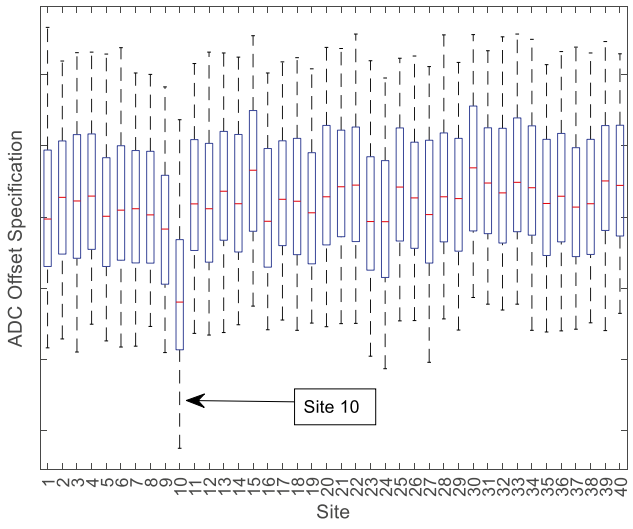

Fig. 1 Illustration of site-to-site variations in a measured chip specification, the ADC Offset specification. Site 10 is identified as an issue site

A method proposed in [11] uses an L-moment-based technique to identify and correct measurements at issue sites. While the paper demonstrated the method's potential, underlying principles, robustness, and validation were not explored in detail.

This paper further extends [11] and generalizes the polynomial transform technique to identify and correct linear and nonlinear systematic errors in multi-site testing. The main contributions in this paper can be summarized as follows:

1. We present a qualitative approach to modeling the multi-site test hardware-induced systematic errors and give reasons for this consideration (Careful evaluation of the error model).
2. We propose using a normal polynomial transform as a transfer medium to identify and correct systematic errors in multi-site test data by comparison with a reference.
3. We validate both the identification and correction methods with simulated and real test data from the semiconductor industry.
4. We further validate the method by showing that correction results for issue sites match closely with the measurements reported for the DUTs by other "good" test sites.

The proposed method accurately quantifies systematic induced errors in measurement data, making it possible to generate and apply a correction coefficient. The technique achieves a higher-order correction in addition to linear correction, increasing induced systematic error coverage. The method offers a software solution to mitigate a hardware problem, improving yield and reducing potential test escape.

The rest of this paper is organized as follows: Section II provides a brief background on hardware systematic errors in multi-site testing. Section III introduces an error model, details the proposed method, and discusses its practical implementation. Sections IV and V present the application of the method to simulation and real measurement data, respectively. Section VI concludes the paper.

2 Preliminaries

All test sites in a multi-site ATE are carefully designed to have the same accuracy and precision. By "accuracy," we mean the closeness of agreement between a measured value and the true DUT value. By "precision," we mean the closeness of agreement among measured values obtained by replicate measurements on the same DUT under specified conditions. However, the multi-site test hardware is built with imperfect elements [3, 10], which causes systematic errors and leads to variations in test measurements from site to site.

In [16], for example, the low-pass behavior of the sampling head and connectors and mismatches in the source and load ports are some of the causes of systematic errors identified in a time-domain network analyzer. Differences in tester resource arrangement, prober alignment, crosstalk, electromagnetic interference, capacitive and inductive coupling amongst analog nets are discussed as possible causes in [15].

Chip manufacturers use various board design and layout techniques to minimize site variations. In [33], ground via stitches are used to isolate all the radio frequency (RF) traces in order to reduce site to site and signal path to path interactions and crosstalk. In [27], an advanced low pin count test architecture is proposed for efficient multi-site testing.

In [21], the multi-site test board design difficulty is acknowledged. Unwanted site-specific effects are reduced by making the parasitics on sensitive signals as equal as possible. This is challenging because complete symmetry for all sites is impossible, even with computer-aided designs. Critical signal paths are also designed to be static (not switched) with power dividers, attenuators, power detectors, and matching circuits to reduce coupling interference and site dependence [24].

In addition to measurement noise, a real semiconductor test environment also includes lot-to-lot variation and wafer-to-wafer [4, 31]. Thus, the need to establish that the variations addressed in this paper are multi-site hardware-induced errors. We are even more particular about systematic errors that are more pronounced and fall outside an acceptable tolerance range, for example, Site 10 in Fig. 1.

Identification of issue sites is usually the first step. In [14], regression fitting on a quantile–quantile curve is shown to pronounce site-to-site variation inherent in test data and aid issue site identification. In [18], an issue site is considered one whose principal component values are beyond the three-sigma control limits of the principal component values of all sites. Repeatability and reproducibility studies are often conducted to correlate the measurements between test sites. However, these methods cannot effectively analyze complex multi-site testing systems that involve multiple measurements across multiple test sites [18].

While smoothing filtering can suppress random measurement errors, systematic errors need to be corrected [23]. Mechanical correction of measurement systems is expensive and time-consuming, hence the industry's search for alternative solutions. In [11], an intuitive method is discussed to adjust an issue site offset and gain error to match a good site. However, this method does not provide a quantitative analysis of the induced errors. It uses test sites' mean and standard deviation, which do not accurately describe measurement results. It is also incapable of higher-order corrections.

In a massive multi-probe test environment, we are often dealing with systematic errors that are complex. Signals pass through buffers, operational amplifiers, and other circuits capable of introducing high-order nonlinearities. A method capable of high-order systematic error identification and correction is proposed in [11]. It uses an L-moment-based method to achieve systematic error identification and calibration. L-moment is one of the tools to achieve polynomial transformation.

The polynomial transform method is a well-established method that has been used in other fields to solve similar challenges. Awareness of the method and its limitations are more easily accessible. It also lends itself to other methods like polynomial chaos, which may spawn further methods for sampling-distortion correction applications.

In [13], least square regression is used to identify and correct systematic errors. While the method showed great promise, it needs further careful investigation of its properties, robustness, applicability, and constraints. Also, the best method to identify and correct systematic errors in multi-site testing may very well depend on the type of the products (analog, mixed-signal, or RF), the nature of the measured specifications, and/or the intrinsic distributions of measurement uncertainty.

Although machine learning solutions are now widely used to solve several semiconductor manufacturing challenges [8, 9, 20], the multi-site test data is unique. Each DUT is only measured at one test site. Test data also contains variations from other sources, making it difficult for machine learning applications.

Some methods to achieve compensation and correction of measurement error in other fields are discussed in [2, 32, 34, 36]. These methods start with a careful evaluation of an error model and use this model to identify the systematic error, which is then calibrated out. We employ a similar approach in this paper.

3 Proposed Method

3.1 Error Model

Mathematical error models for measurement systems are usually related to the system's physical structure. This section discusses the difficulty of such an approach (for our case) and instead presents a black-box approach.

Multi-site test hardware is highly complex instrumentation and measurement system consisting of many signal generators, oscilloscopes, complicated printed board circuits (PCB) boards, probe head and pins, and many more parts. Figure 2 is a simplified view of the multi-site test hardware section. The probe card interfaces between the ATE and the semiconductor wafer. It is mechanically docked to the prober and electrically connected to the tester. An error model related to the system's physical structure will have to consider all these components.

The probe board alone can have over 100 layers of printed circuit boards (PCB), and on each layer, thousands of electrical components like resistors, capacitors, diodes, and inductors. Many of these components do not

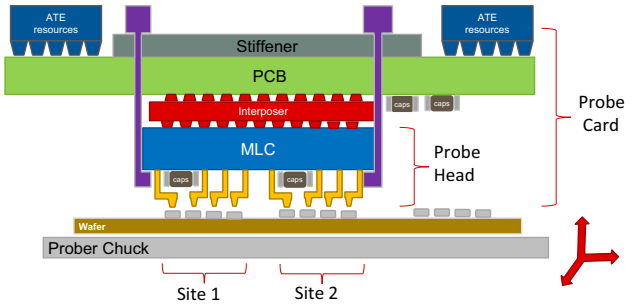

Fig. 2 Simplified view of a section of the multisite testing system showing different parts

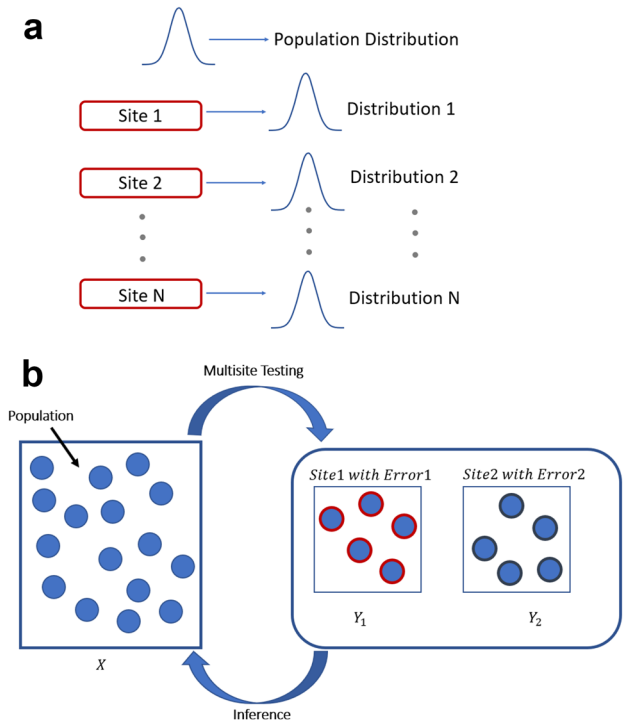

Fig. 3 a Multisite testing model without the assumption that each test site introduces errors. Each site's distribution will follow the population distribution. b Multisite testing model with the assumption that each test site introduces errors.

have error models, and vendors can only guarantee their values within a defined tolerance range. Even if we can model all these components correctly, we will have to find a way to include thermal drift, electromagnetic interference, noise, and inductive coupling in the models because all these issues come into effect when all the test sites are powered and in use.

The complexity of such a model becomes pronounced when we want to factor in the age of the components, contact resistance at the probe pins, and other building blocks. While these concerns are possible but difficult, one big challenge with a physical structure approach will be its dependence on the multi-site test hardware in question. The developed error model might not apply to test hardware from other vendors, making such model transfer difficult. We treat the entire multi-site system and its complexity as a black box to avoid these complications and other issues. This box takes in the actual value of a DUT and gives out a measured value inclusive of systematic error.

Each DUT is assigned to one of the test sites during the testing process. Hence, the multi-site test system can be modeled (with some modification) using a simple statistical sampling model. Each measured DUT specification has a true and expected distribution, which forms our population, X. Suppose multi-site testing does not introduce any errors; each site's measurement will follow the true and expected distribution, X, as shown in Fig. 3a. This model is used extensively for issue site identification. Sites with a measurement distribution different from the majority are flagged as issue sites.

However, multi-site test hardware introduces errors (assumed to be small) that cause SSV. The induced errors vary from site to site, as seen in Fig. 3b. Each test site not only samples from the population, X, but also adds its own errors.

It is not unusual for hundreds to thousands of measured specifications to be tested for each DUT. Let X be the true but unknown distribution for a particular measured specification P. Also, let Y_k be the measured distribution for site k. Ideally, the expectation is that $Y_k = X$. However, this is never true. Each measured site's distribution can be modeled as:

$$Y_k \approx \epsilon_0^k + \left(1 + \epsilon_1^k\right) * X + \epsilon_2^k * X^2 + \cdots + \epsilon_r^k * X^r + Noise_k \quad (1)$$

where $Noise_k$ represents the noise introduced by site k and where $\epsilon_0^k, \epsilon_1^k, \epsilon_2^k, \ldots, \epsilon_r^k$ are the identified induced errors. Two assumptions are made. First, noise is assumed to be uniform for each site; hence *Noise_k* is excluded in further analysis. Second, systematic errors are assumed to dominate other sources of variation during test measurements, including measurement noise.

A polynomial function is preferred because it could approximate any discontinuous or continuous function [1]. One other interesting reason for choosing a polynomial model is the interpretation of identified errors. For example, ϵ_0^k easily represents a DC shift while ϵ_1^k represents a gain shift. $\epsilon_2^k, \ldots, \epsilon_r^k$ represent higher-order nonlinearities. Other functions that could be considered required more computational time and power which would be important when this method is to be implemented in real-time.

While X could be obtained using simulations performed using IC layout and fabrication parameters, they do not contain real manufacturing and measurement errors. Silicon measurements are preferred to estimate a reference distribution. Multi-site variability-aware reference distribution estimation methods have been proposed in [5, 12]. Ordinal

optimization is used to identify high confidence good test sites. Measurements from the high confidence good sites are transformed to adjust for the test site size, mean, and standard deviation before being lumped together to form a reference.

3.2 Polynomial Transform Method

Polynomial normal transformation is often used in several probabilistic analyses, especially when multivariate nonnormal random variables are involved [6]. The technique is especially attractive when the probability density function (PDF) of the random variable is unknown, and only samples of the random variable are available. Such is the case with many DUTs measured specifications in semiconductor testing.

X can be expressed as a polynomial expression of Z, a standard normal random variable. The transformation from Z to X can be formulated as:

$$X = a_0 + a_1 Z + a_2 Z^2 + \cdots + a_r Z^r \tag{2}$$

where r is the polynomial order and $a_0, a_1, a_2, \ldots, a_r$ are the undetermined coefficients. With suitable values of $a_0, a_1, a_2, \ldots, a_r$, various probability distributions can be accurately simulated [37].

The undetermined coefficients can be determined by the Product-moment (PM) method, the L-moments (LM) method, the Least-square (LS) method, and the Fisher-Cornish (FC) asymptotic expansion. In [6], these four methods are presented and analyzed when r is three. The performance of the four methods is investigated by comparing them with a parametric technique using the Rosenblatt transformation that preserves the marginal distribution of the nonnormal random variables.

Substituting Eq. 2 into Eq. 1 results in Eq. 3. Equation 3 is truncated based on the chosen value for r

$$\begin{aligned} Y_k \approx\ & \epsilon_0^k + \left(1 + \epsilon_1^k\right) * \left(a_o + a_1 Z + a_2 Z^2 + a_3 Z^3\right) \\ & + \epsilon_2^k * \left(a_o + a_1 Z + a_2 Z^2 + a_3 Z^3\right)^2 + \ldots \end{aligned} \tag{3}$$

Y_k can also be expressed as a polynomial transform of Z.

$$Y_k \approx a_0^k + a_1^k Z + a_2^k Z^2 + \cdots + a_r^k Z^r \tag{4}$$

where $a_0^k, a_1^k, a_2^k, \ldots, a_r^k$ are the polynomial coefficients for Y_k. Equations 3 and 4 are solved by comparing the coefficients of Z to provide the identified error coefficients.

An inverse approach to Eq. 1 can be written as

$$\begin{aligned} X \approx\ & C_0^k + \left(1 + C_1^k\right) * Y_k + C_2^k * \left(Y_k\right)^2 + \ldots \\ & + C_r^k * \left(Y_k\right)^r \end{aligned} \tag{5}$$

where $C_0^k, C_1^k, C_2^k, \ldots, C_r^k$ are polynomial corrections to site$k$. $C_0^k, C_1^k, C_2^k, \ldots, C_r^k$ are solved for using least squares and used to correct each measured result using the equation:

$$y_k^{Cal} = C_r^k \left(y_k\right)^r + \cdots + C_2^k \left(y_k\right)^2 + C_1^k y_k + C_0^k \tag{6}$$

where y_{Cal}^k is the corrected result of y_k belonging to the distribution Y_{Cal}^k.

3.3 Implementation and Discussion

A third-order polynomial is considered and applied to simulation and real test data in this paper. The L-moments method with r as three is utilized. The method is chosen as it could characterize a wider range of distributions and is more robust to the presence of outliers in the data. The third-order normal polynomial (TPNT) coefficients estimated by this method are sometimes more accurate and more computation efficient than other methods [6], hence our choice. If the first four population or sample L-moments of a random variable are known, the polynomial coefficients (a_0, a_1, a_2, a_3) be directly determined from the first four probability-weighted moments of X.

Figure 4 is a simplified flowchart of the proposed method and application using a specification-based approach. The process starts with data filtering (cleaning and processing) to remove invalid data after loading in test volume data. Invalid data are not only outliers but tester or DUT malfunctions, which do not contain any useful information. Any suitable issue site identification algorithm is used to detect the presence of issues test sites. If there are issue test sites, the proposed method is applied, and each test site's systematic error correction coefficients are stored. While the user may decide to correct all test sites, this paper only corrects test data related to identified issue sites.

New test data taken by each test site are loaded and recorded. If the measured specification has been flagged earlier to need calibration, the stored calibration coefficients are drawn and used to correct the issue sites before data output (offline).

The Big-O complexity notation to investigate the time and space complexity of the method is irrelevant because the number of test sites is not large enough to be a concern for now. It is a fixed amount of time for each chip correction. Currently, it takes less than 1 ms in MATLAB® to identify induced errors in more than 10,000 DUTs and correct them. In the future, when the technique is moved online and adaptive (in real-time as testing of DUTs is in the process), computation performance and time will be critical.

We assume that only a few test sites have unacceptable variations. This assumption is valid because not only is the multi-site test system well designed; it is also rigorously tested and sometimes calibrated (hardware) during the setup before use. Another assumption is the neglect of noise in our derivations. In [16], the authors found that the noise magnitude in their system can be described in the time domain by a Gaussian distribution with a standard deviation. We

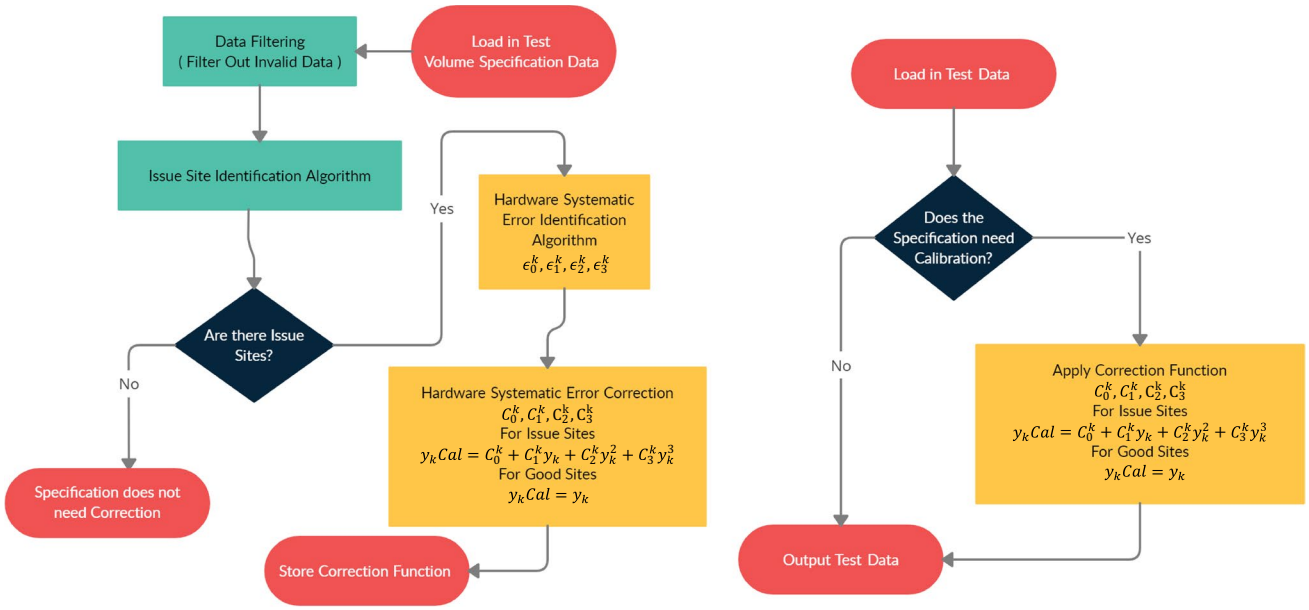

Fig. 4 Simplified flowchart of the proposed method and application

assume that this is similar for test sites and that systematic errors dominate and mask them.

A multivariate approach to the proposed method and equations appears to give a more complete picture of the problem; however, it does not necessarily realize the answers on a per-specification basis that test engineers seek. To elaborate further, multi-site hardware systematic errors do not uniformly affect measured specifications of DUTs. Looking at real multiprobe test data of production chips, we have seen that signal interference on a specific site might affect specification A positively (making measured data look better), affect specification B negatively (making measured data look worse), and exhibit no effect on specification C. A multivariate approach will combine all measured specifications together, and the investigated systematic error might be lost in the process. Test engineers are also most interested in investigating particular specifications showing noticeable yield loss.

The normal polynomial transformation has been applied in several fields where there is a need to represent statistical data of a Fig. 4. Simplified flowchart of the proposed method and application.

variable when the underlying distribution is unknown [37]. In [28], it is used to characterize multiple rainfall random variables that have a mixture of nonnormal distributions for which a joint distribution function is difficult to establish. In [7], it was used to establish a relationship between wind speed distribution and standard normal when historical data is given. While these papers used a third-order normal transformation because it was good enough [5], a higher-order transformation could be considered.

In [19], a fifth-order polynomial transformation (FPNT) is considered, allowing additional control of the fifth and sixth moments. The ability to control higher moments increases the precision of nonnormal distribution approximations. A ninth-order polynomial transformation (NPNT) was implemented in [37] and shown to accommodate many distributions that are difficult for the TPNT/FPNT [37]. A TPNT is considered sufficient in this paper because most measured DUT specifications are gaussian and symmetric. A few are skewed (ordered statistics), like the maximum and minimum values of some measured specifications. Very few have other distributions.

Note that the TPNT does not guarantee compliance with the monotonicity condition [28]. The boundary condition in Eq. 6 bounds the L-moment method. The method discusses the $X - Z$ transformation. However, in practical analysis, the inverse transformation $Z - X$ is also essential. The $X - Z$ transformation is essentially the solution of the cubic equation of the standard random variable as shown in Eq. 5. Generally, there are three real or complex roots. This means that one value of X may correspond to multiple possible values of Z. This is a critical problem to determine the suitable root and its range under the circumstance of the complete monotonic requirement of the transformation. This is not usually a problem in multi-site test data because test measurements are already bounded by the DUT design. Also, outlier data are filtered out.

In [35], the authors propose a complete monotonic expression of the TPNT with different combinations of the first four L-moments. The paper differentiates the applicable boundaries and monotonic regions of each expression. The accuracy of the transformation is examined, and the statistical uncertainty of the TPNT based on the first four L-moments is investigated. This expression can be applied if monotonicity issues are inherent in test data.

4 Simulation Results

We created MATLAB® models to simulate and generate volume test data. MATLAB® R2021a version was used. Each ADC is modeled to have an intrinsic error similar to what is observed in reality. Each test site is programmed to test between 2000 and 3000 ADCs across multiple simulated wafers based on a modeled test site location. In all, the presented 40 test sites measure the Gain Error specification of over 90,000 ADCs.

Figure 5 presents the simulated Gain Error true values. The variations inherent in this data are only due to intrinsic variations in the DUTs, hence minimal. In reality, the true specification values are unknown. What is reported by test sites are measured values.

Figure 6 presents the simulated Gain Error measured values (reported by test sites). Although the SSV is more than the SSV shown in Fig. 5, the variations are within a tolerance range, and test quality is still guaranteed. This data is closer to what is obtainable in the semiconductor testing industry.

In our simulation experiment, we simulate error sources to cover some possible faults evident in test data, like linear errors, higher-order nonlinearities, changes in distribution, and the addition of another distribution. In [11], high-order nonlinearities were included in addition to the modeled clock signal coupling. This paper explored other functions like the sigmoid function and sine wave as induced additions to critical nets. Figure 7 presents the simulated Gain Error measured values. Sites 2, 11, and 26 have been randomly selected as issue sites. Site 2 has a range similar to other sites, but its median and quartiles appear to be shifted downward due to the induced errors. Site 11 has its range affected, and quartiles noticeably moved upward. Although not visually evident, site 26 has an extended interquartile range (IQR) compared to other sites. We controlled the level of induced errors in simulated issue sites to be similar to what is obtained in real semiconductor testing. These sites are also confirmed as issue sites by the issue site identification algorithm presented in [14].

4.1 Robustness

We investigate the robustness of the method to test data discrepancies. Systematic errors induced by issue test sites should be consistent irrespective of the DUTs tested by such sites. Each test site in Fig. 7 measured between 2000 and 3000 DUTs based on the test site location on the probe card. We randomly select 600 DUTs from each test site and apply the proposed identification algorithm to each random selection.

Figure 8 shows the identified induced errors for ten random selections of 600 DUTs for each test site. The identified induced errors reported for the issue sites are consistent and within an error bound for each selection. Reported systematic errors for 'good' sites are within a tolerance range across each selection. These results confirm the consistency of the method.

We observed a smaller selection error bound for ϵ_0 and ϵ_1 in Fig. 8. This is expected because linear errors are more

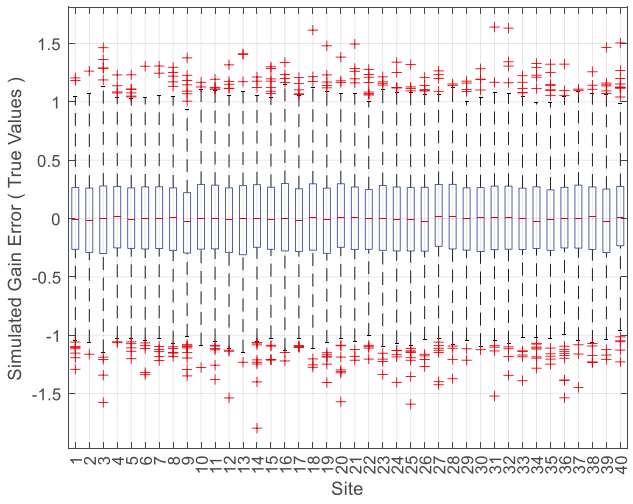

Fig. 5 Simulated ADC gain error true values. The variations inherent in this data are only due to intrinsic variations in the DUTs

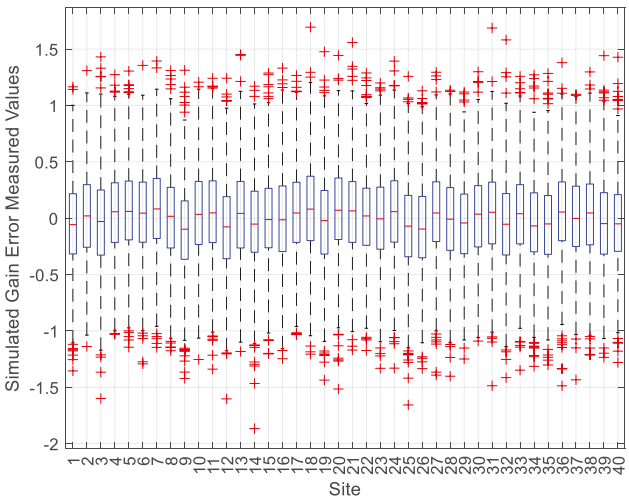

Fig. 6 Simulated ADC gain error measured values. (No issue sites). The variations inherent in this data are due to intrinsic variations in the DUTs and also measurement errors from the measuring instruments

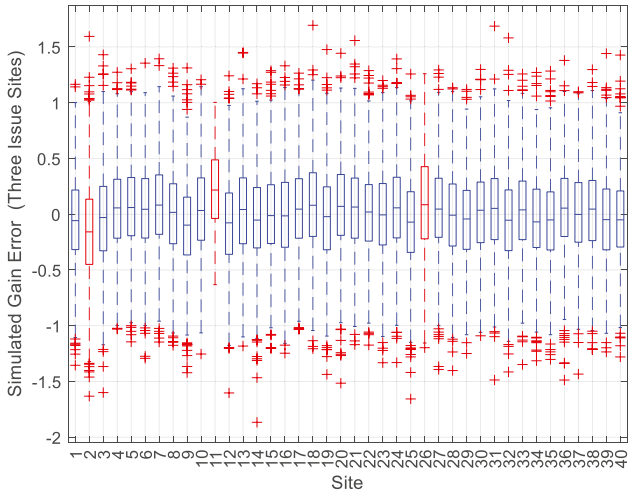

Fig. 7 Simulated ADC gain error measured values (Sites 2, 11 and 26 are issue sites)

dominant and consistent. Also, issue sites with high order nonlinearities; for example, site 11 remained an outlier in all selections for ϵ_3.

4.2 Accuracy

We use test data from one of the ten selections in subsection A to generate and store our correction function. This correction function is applied to the test data presented in Fig. 7, and the results are presented in Fig. 9. It is observed that the shift, gain, and higher-order nonlinearity evident in Fig. 8 have been corrected. While this correction is visually evident in the box plot summary of the issue site, we proceed to compare DUT measurements before and after correction against the true DUT values.

In Fig. 10, measured results for site 11 deviate from the true value line. Our proposed method correction method accurately corrected the nonlinearities introduced by site 11 and brought the corrected measurements closer to the true value slope line.

Similar results are obtained for site 2, as presented in Fig. 11. Our corrected results are closer to the true value slope line for the issues sites. They are not exactly the true values because of the measurement noise included in the simulation.

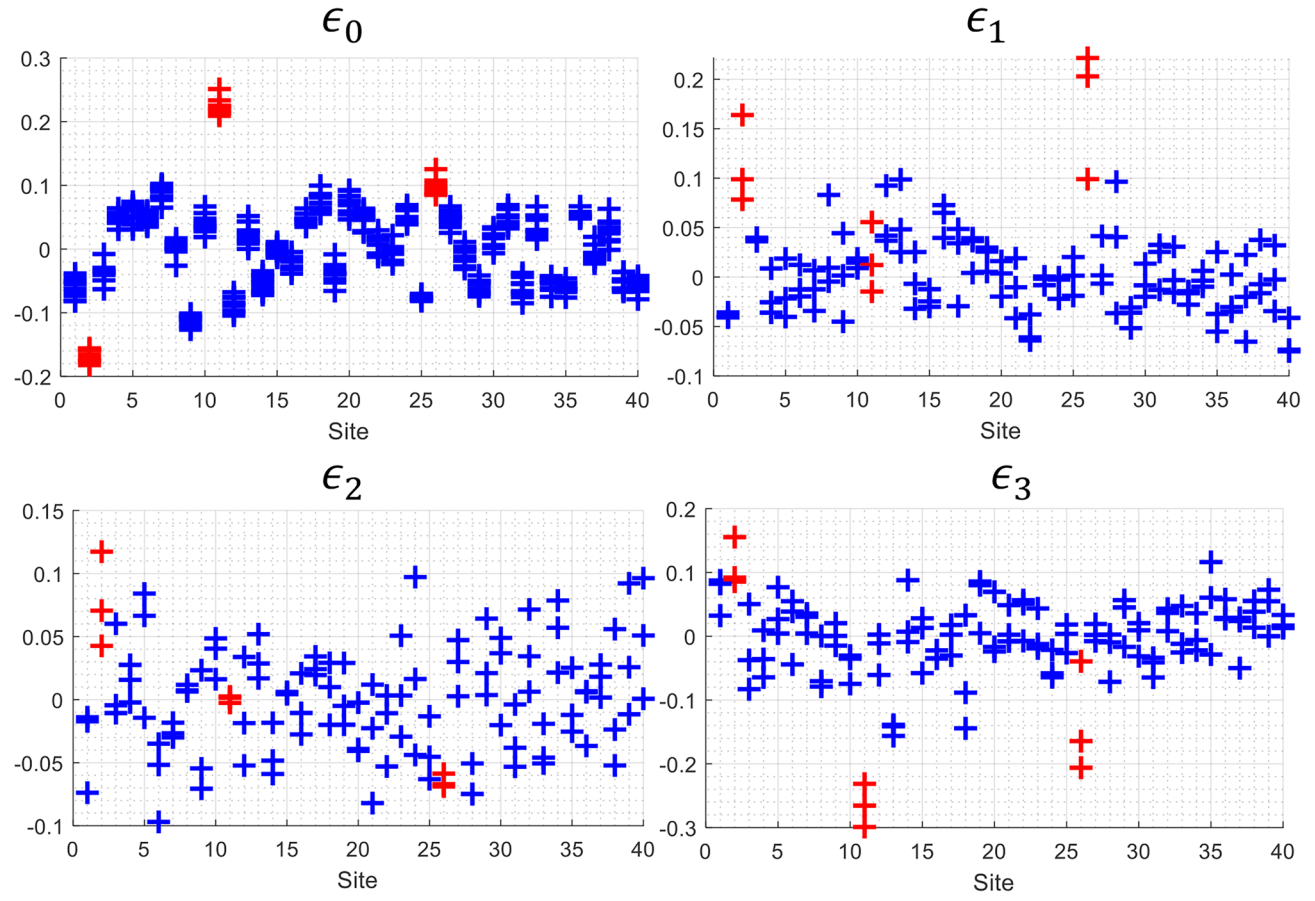

Fig. 8 Identified hardware systematic errors for 10 random selections of 600 DUTs for each site (Test sites with red boxplots are identified issue test sites)

5 Silicon Results

To further investigate the effectiveness of the proposed method, we applied the proposed method to test data obtained from Texas Instruments. This volume data was procured from one of their multi-site test hardware, where site-to-site variations have been observed. For data confidentiality reasons, we exclude the measured parameter values. We present data for a subset of 40 test sites.

We perform a series of experiments, as described in three parts.

Part A: We apply the proposed method to volume data, present the identified systematic errors, and correct the volume test data. In [11], the nature of the induced error was linear in nature. In this paper, we consider real test data with more variability and higher-order coefficients.

Part B: In real test measurements, we do not have the true DUT values to compare our corrected results with. A normalized histogram approach is used to compare the measurement distributions of test sites before and after corrections.

Part C: To further validate the proposed method, we proceed to take more measurements. We test a particular DUT at multiple sites (both good and issue sites) using the same test hardware and ATE. We show that the results from corrected issue sites match closely with multiple measurements from good sites. This validation method is labor-intensive and manually generated. We explain why results from a few DUTs are used.

5.1 Part A

Figure 12 presents the measured ADC Gain Error parameter for 40 selected test sites. The volume test data contains more than 171,000 ADC measurements. To reduce other sources of variation, only data from one probe card is analyzed and presented.

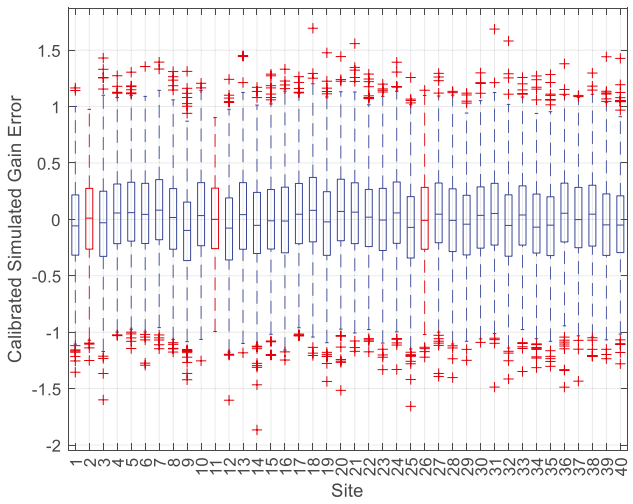

Fig. 9 Correction results for the simulated ADC gain error parameter. (Test sites with red boxplots are identified issue test sites)

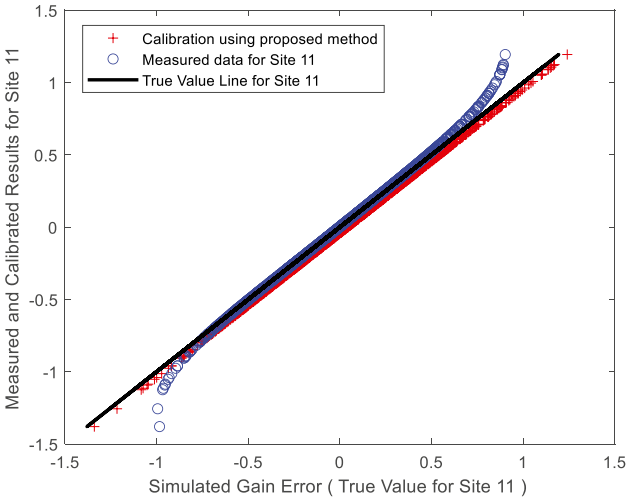

Fig. 10 Measured and corrected data for Site 11. Our correction result is closer to the true value slope line

As seen in Fig. 12, sites A, B, and C exhibit a downward shift compared to other sites. The issue site identification algorithm discussed in [16] is implemented in MATLAB® and confirms our visual inspection. Figure 12 established unacceptable pronounced variations in test measurements across sites for this measured specification.

The proposed method is implemented in MATLAB® and applied. The identified systematic errors are presented in Fig. 13. The ϵ_0^k reported for the three identified issue sites can be visually confirmed as outliers when compared with

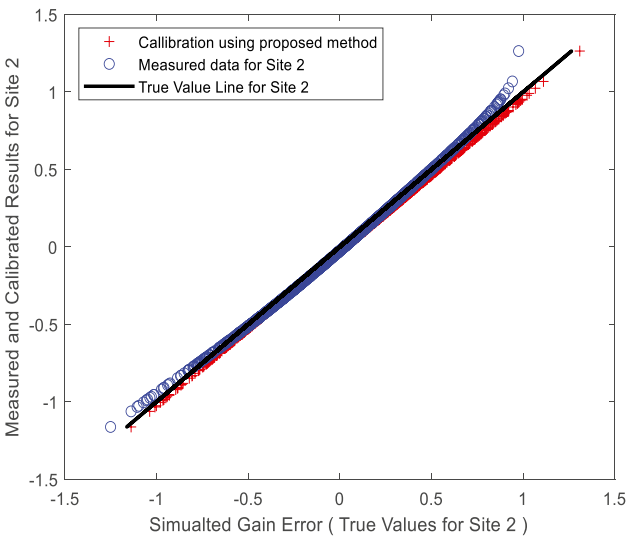

Fig. 11 Measured and corrected data for site 2. Our correction result is closer to the true value slope line

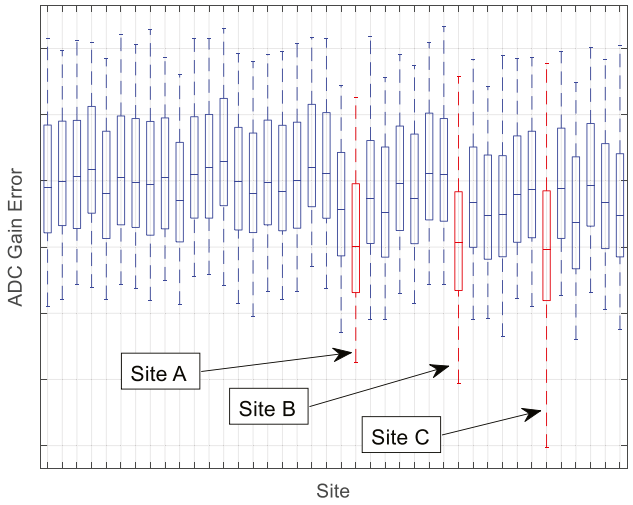

Fig. 12 Measured ADC Gain Error parameter (Sites with red boxplots are identified issue test sites, Sites A, B, and C)

other sites. Site B is the only outlier for ϵ_1^k. Site C is the only outlier for ϵ_2^k. Site B and site C exhibit higher-order error coefficients, ϵ_3^k even though they are not obvious from the box plot visualization.

A correction function is derived from the identified systematic errors and applied to the issue sites. It is evident in Fig. 14 that the induced errors in the three issue sites have been corrected. This does not confirm that the algorithm is correct or provide any evidence. It only shows the effectiveness of the correction by looking at the boxplot summary of the corrected issue sites.

5.2 Part B

To further validate the proposed method with real test data, we compared the distribution of measured parameters before and after correction using a normalized histogram approach. Each site's histogram is normalized to cover the same area. Due to systematic errors, the normalized histogram of issue sites deviates from the expected reference. When it is corrected, we expect that the histogram of the corrected measurement data will be closer to this reference.

Figure 15 shows the normalized histogram curve comparing the before and after correction of sites A, B, and C with a reference distribution.

The massive multi-site variability-aware method in [12] is used to obtain this reference distribution. Ordinal optimization is used to identify high confidence good test sites. Measurements from the high confidence good sites are transformed to adjust for test site size, mean, and standard deviation before they are lumped together to form a reference distribution.

It is observed by visual inspection of Fig. 15 that for sites A, B, and C, the distribution after correction is closer to the reference distribution. Several distance functions can be used to measure the differences between two histograms. Some examples include the Hellinger distance, Euclidean Distance, Chebyshev Distance, etc. We decided to use the Manhattan or City Block distance defined as

$$D_s = \sum_i \left| h_r(i) - h_k(i) \right| \tag{7}$$

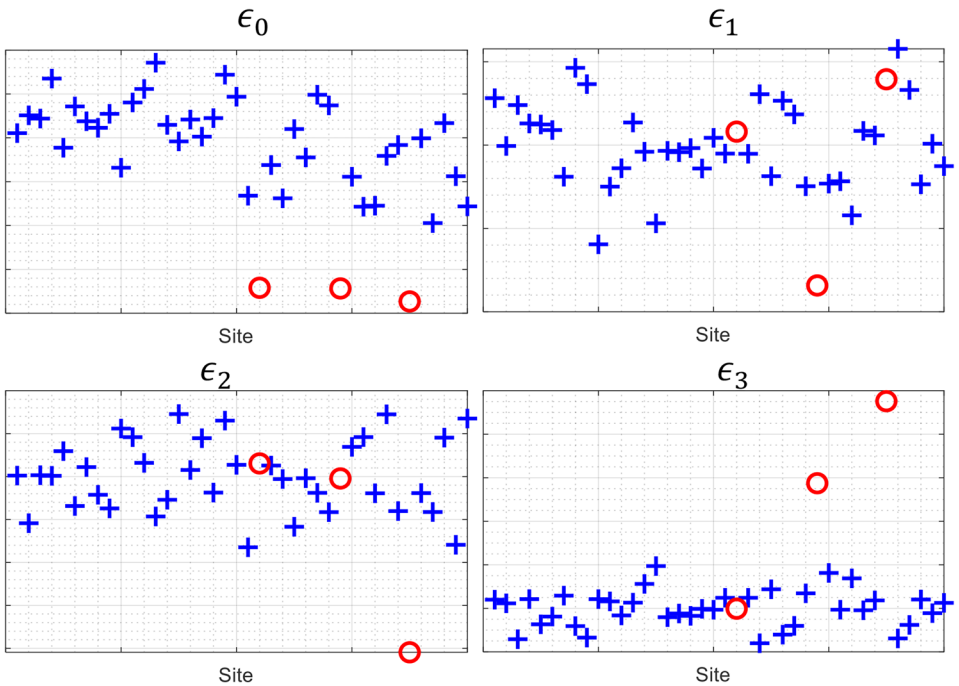

Fig. 13 Identified errors after applying the proposed method to the test data presented in Fig. 12. (Sites with red circles are previously identified issue sites.)

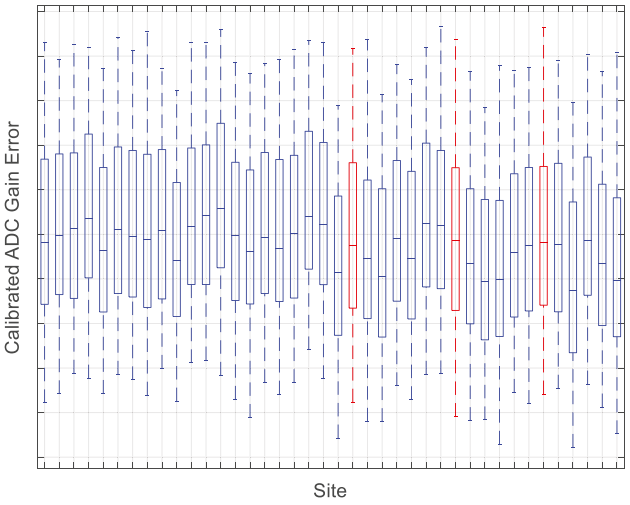

Fig. 14 Corrected ADC gain error parameter (Test sites with red boxplots are identified issue test sites)

where $h_r(i)$ is the histogram counts of the reference distribution, and $h_k(i)$ is the histogram count for site k distribution. D_k is the difference metric for site k.

Table 1 presents the distance metric for the measured and corrected distribution for sites A, B, and C to a reference distribution as shown in Fig. 15. For Site A, the proposed method reduced the distance metric from 65.28 to 5.2. The distance metric for sites B and C are reduced as well after correction.

5.3 Part C

To validate our correction algorithm with real test data, we manually took repeated measurements of some selected ADCs. Suppose we test an ADC at multiple sites that include both issue sites and good sites. We can determine if the corrected results of the issue sites match closely with the measurement results reported at other good sites.

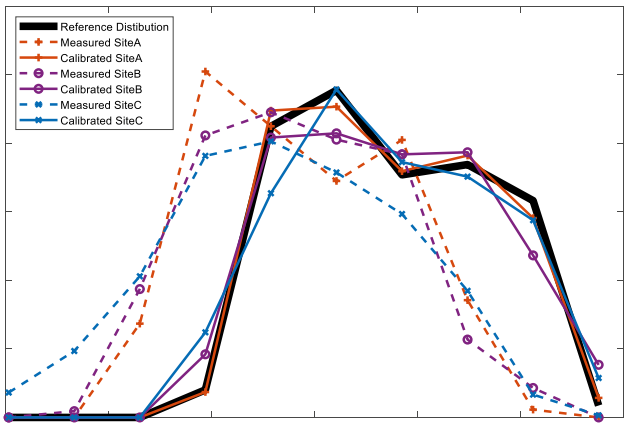

Fig. 15 The normalized histogram curve shows that for sites A, B and C, the distribution after correction is closer to the reference distribution

Multiple repeated testing of the same DUT has its demerits. It is destructive to the probe pins and the DUT pins. Also, contact resistance becomes an issue after a series of touchdowns, limiting the number of possible probe touchdowns on the same DUTs on a wafer.

For each touchdown, we take ten measurements to even out the measurement noise at each site. In total, 209 selected ADCs on a wafer were tested multiple times. A fewer number of dies were used because the method is labor-intensive, manual, and carefully planned to ensure ADCs are measured by both issue and good sites. Still, fewer ADCs reported being measured by at least four good sites and one issue site during six touchdowns.

Figure 16 presents the measurement and correction results from five different sites for the same ADC. Each box plot contains all the measurements taken during the six touchdowns on the selected ADC by that site. GS1, GS2, GS3, and GS4 are the measurements reported by the good sites, while Cal. GS1, Cal. GS2, Cal. GS3, and Cal. GS4 are the corrected result. They are the same because only the measurements of issue sites are corrected. We observe that the corrected results for site C, Cal. Site C moved the measurement result of the issue site, site C, closer to the reported measurement by the good sites.

Figure 16b is a better representation of the result presented in Fig. 16a. In Fig. 16b, we lump all the test data for the good sites together into one (Good Sites). This aids in comparison with the correction results for issue sites. This method is used to present data for other issue sites and different ADCs, as shown in Fig. 17. For each of the issue sites, the correction results are closer to the measurements reported by the good sites.

Reported measurement by good sites (which is what we have and use for comparison) varies slightly because each site introduces systematic errors (though acceptable), further making validation tricky.

In summary, we use repeated data to validate the proposed correction algorithm. On average, the correction algorithm moved the measurement of issue sites closer to the measurement reported by good sites.

Table 1 Distance metric showing that for sites A, B and C, the distribution after correction is closer to the reference distribution

	Site A D_A	Site B D_B	Site C D_C
Measured	65.2821	61.8426	68.0891
Corrected	5.2649	15.9124	14.3951

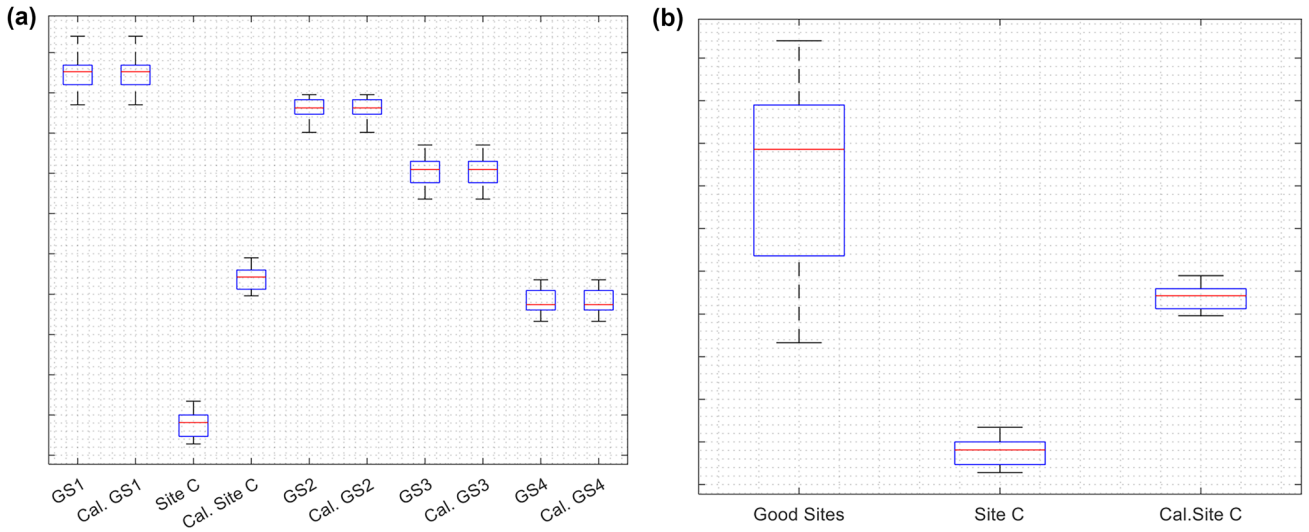

Fig. 16 **a** All measurement and corrected measurement for site C. **b** Better presentation of the results in Fig. 16a

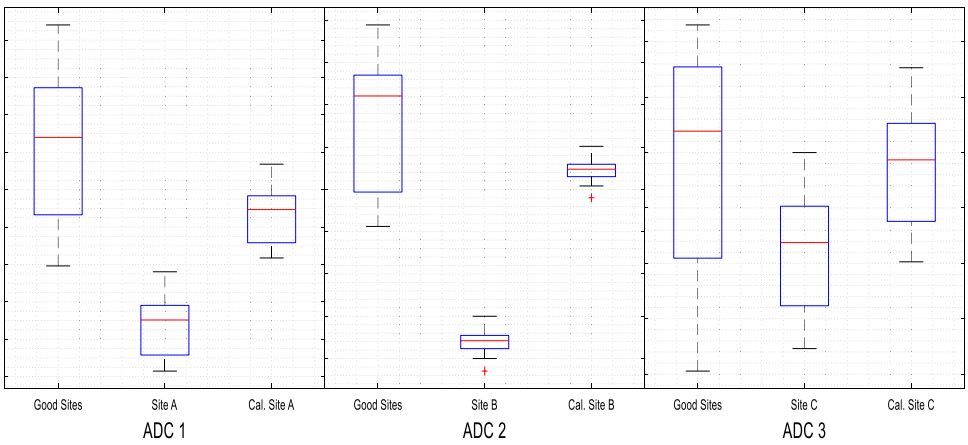

Fig. 17 Extra measurement with chip showing good site, issue site and corrected site

6 Conclusion

The multi-site testing system is modeled as a sampling process with systematic errors introduced by each test site. An error model is presented, and a polynomial transformation method to solve and correct for induced nonlinearity is discussed. Identified systematic errors are used to correct test data avoiding expensive and time-consuming hardware solutions.

We validate the method by applying it to simulated and real test data. The method significantly reduced the level of systematic error observed in test measurements for issue sites. We further validate the method with additional DUT measurement by 'good' sites. The corrected DUT values agree with multiple measurements of the same DUT by good test sites. We discuss the assumptions, limitations, and potential of the method.

While the detection of site-to-site variations in measurement data has been receiving attention recently, identifying and correcting the root causes remains challenging and impractical given the complexity, long design cycle, and high cost of the multi-site test hardware. This paper introduces a practical and low-cost approach to effectively identify and correct site-to-site variations inherent in multi-site measurement data. MATLAB® codes were developed to implement the proposed method. Simulation results demonstrate the accuracy and robustness of the proposed method. Application to real production multi-site measurement data also verifies the effectiveness of the method.

The proposed method will further guarantee test quality while reducing the risk of yield loss and possible test escapes. The method also ensures continuous use of the multi-site test hardware until a mechanical solution is

available. Future work is to make the method online and adaptive so that the correction coefficients are updated as more site measurements are available.

Acknowledgment The authors will like to thank Texas Instruments Incorporated and Semiconductor Research Corporation (SRC) for their support, collaboration, and sponsorship.

Availability of Data The data that support the findings of this study are available from Texas Instruments Inc, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however, available from the authors upon reasonable request and with permission of Texas Instruments Inc.

Declarations

Conflict of Interest I certify no actual or potential conflict of interest about this article The authors declare that they have no known competing financial interests or personal relationships that could influence the work reported in this paper.

References

1. Alici G, Shirinzadeh B (2005) A systematic technique to estimate positioning errors for robot accuracy improvement using laser interferometry based sensing. Mech Mach Theory 40(8):879–906. https://doi.org/10.1016/j.mechmachtheory.2004.12.012
2. Arpaia P, Cennamo F, Daponte P (1999) An error correction technique for scan conversion-based transient digitizers. IEEE Trans Instrum Meas 48(4):785–792. https://doi.org/10.1109/19.779172
3. Bajurko PR (2014) Calibration of the measurement system using picoseconds pulses. In Proc. 20th International Conference on Microwaves, Radar and Wireless Communications (MIKON) 1–4. https://doi.org/10.1109/MIKON.2014.6899857
4. Boning DS et al (2008) Variation. IEEE Trans Semicond Manuf 21(1):63–71. https://doi.org/10.1109/TSM.2007.913194
5. Bruce I, Farayola PO, Chaganti SK, Obaidi AO, Sheikh A, Ravi S, Chen D (2021) An ordinal optimization-based approach to die distribution estimation for massive multi-site testing validation: a case study. In Proc. IEEE European Test Symposium (ETS) 1–4. https://doi.org/10.1109/ETS50041.2021.9465402
6. Chen X, Tung Y-K (2003) Investigation of polynomial normal transform. Struct Saf 25(4):423–445. https://doi.org/10.1016/S0167-4730(03)00019-5
7. Cheng H, Zhang L, Yan T, Zhou Q, He H, Zeng P (2015) Reliability evaluation of wind integrated composite power system based on third-order polynomial normal transformation. In Proc. 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT) 472–477. https://doi.org/10.1109/DRPT.2015.7432307
8. Cheng KC-C et al (2021) Machine Learning-Based Detection Method for Wafer Test Induced Defects. IEEE Trans Semicond Manuf 34(2):161–167. https://doi.org/10.1109/TSM.2021.3065405
9. Cheon S, Lee H, Kim CO, Lee SH (2019) Convolutional Neural Network for Wafer Surface Defect Classification and the Detection of Unknown Defect Class. IEEE Trans Semicond Manuf 32(2):163–170. https://doi.org/10.1109/TSM.2019.2902657
10. Ciganda Brasca LM et al (2011) A parallel tester architecture for accelerometer and gyroscope MEMS calibration and test. J Electron Test 27(3):389–402. https://doi.org/10.1007/s10836-011-5210-2
11. Farayola PO, Bruce I, Chaganti SK, Obaidi AO, Sheikh A, Ravi S, Chen D (2021) Systematic Hardware Error Identification and Calibration for Massive Multisite Testing. Proc. IEEE International Test Conference (ITC) 2021:304–308. https://doi.org/10.1109/ITC50571.2021.00042
12. Farayola PO, Bruce I, Chaganti SK, Sheikh A, Ravi S, Chen D (2021) Massive multi-site variability-aware die distribution estimation for analog/mixed-signal circuits test validation. In Proc. 16th International Conference on Design Technology of Integrated Systems in Nanoscale Era (DTIS) 1–6. https://doi.org/10.1109/DTIS53253.2021.9505144
13. Farayola PO, Bruce I, Chaganti SK, Sheikh A, Ravi S, Chen D (2022) The least-squares approach to systematic error identification and calibration in semiconductor multisite testing. In Proc. IEEE 40th VLSI Test Symposium (VTS). pp. 1–7. https://doi.org/10.1109/VTS52500.2021.9794216
14. Farayola PO, Chaganti SK, Obaidi AO, Sheikh A, Ravi S, Chen D (2020) Quantile – quantile fitting approach to detect site to site variations in massive multi-site testing. In Proc. IEEE 38th VLSI Test Symposium (VTS) 1–6. https://doi.org/10.1109/VTS48691.2020.9107616
15. Farayola PO, Chaganti SK, Obaidi AO, Sheikh A, Ravi S, Chen D (2021) Detection of Site to Site Variations From Volume Measurement Data in Multi-site Semiconductor Testing. IEEE Trans Instrum Meas 70:1–12. https://doi.org/10.1109/TIM.2021.3051666
16. Ferrari P, Angenieux G (2000) Calibration of a time-domain network analyzer: a new approach. IEEE Trans Instrum Meas 49(1):178–187. https://doi.org/10.1109/19.836331
17. Hashempour H, Meyer FJ, Lombardi F (2005) Analysis and evaluation of multi-site testing for VLSI. IEEE Trans Instrum Meas 54(5):1770–1778. https://doi.org/10.1109/TIM.2005.855099
18. He S-G, Wang GA, Cook DF (2011) Multivariate measurement system analysis in multi-site testing: An online technique using principal component analysis. Expert Syst Appl 38(12):14602–14608. https://doi.org/10.1016/j.eswa.2011.05.022
19. Headrick TC (2002) Fast fifth-order polynomial transforms for generating univariate and multivariate nonnormal distributions. Comput Stat Data Anal 40(4):685–711. https://doi.org/10.1016/S0167-9473(02)00072-5
20. Imoto K, Nakai T, Ike T, Haruki K, Sato Y (2019) A CNN-Based Transfer Learning Method for Defect Classification in Semiconductor Manufacturing. IEEE Trans Semicond Manuf 32(4):455–459. https://doi.org/10.1109/TSM.2019.2941752
21. Kore I, Schuffenhauer B, Demmerle F, Neugebauer F, Pfahl G, Rautmann D (2011) Multi-site test of RF transceivers on low-cost digital ATE. In 2011 IEEE International Test Conference 1–10. https://doi.org/10.1109/TEST.2011.6139142
22. Kulovic K, Margala M (2012) Time-Based Embedded Test Instrument with Concurrent Voltage Measurement Capability. J Electron Test 28(5):653–671. https://doi.org/10.1007/s10836-012-5299-y
23. Lan L, Guo K, Shen Y, Yang H (2021) Research and implementation of two-axis antenna shafting error calibration based on calibration satellite. In Proc. 13th International Conference on Communication Software and Networks (ICCSN) 150–154. https://doi.org/10.1109/ICCSN52437.2021.9463617
24. Lehner T, Kuhr A, Wahl M, Brück R (2014) Site dependencies in a multi-site testing environment. In Proc. 19th IEEE European Test Symposium (ETS) 1–6. https://doi.org/10.1109/ETS.2014.6847808
25. Ma X, Lombardi F (2006) Multi-site and multi-probe substrate testing on an ATE. In Proc. 21st IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems 495–506. https://doi.org/10.1109/DFT.2006.45
26. O'Donoghue G, Gomez-Uribe CA (2008) A Statistical Analysis of the Number of Failing Chips Distribution. IEEE Trans Semicond Manuf 21(3):342–351. https://doi.org/10.1109/TSM.2008.2001208

27. Seo S, Lee Y-W, Lim H, Kang S (2020) Advanced Low Pin Count Test Architecture for Efficient Multi-Site Testing. IEEE Trans Semicond Manuf 33(3):391–403. https://doi.org/10.1109/TSM.2020.2994182
28. Tung YK, You L, Yoo C (2019) Third-order polynomial normal transform applied to multivariate hydrologic extremes. Water 11(3). https://doi.org/10.3390/w11030490
29. Volkerink EH, Khoche A, Rivoir J, Hilliges KD (2003) Modern Test Techniques: Tradeoffs, Synergies, and Scalable Benefits. J Electron Test 19(2):125–135. https://doi.org/10.1023/A:1022829321216
30. Wang F-K (2008) Process Yield With Measurement Errors in Semiconductor Manufacturing. IEEE Trans Semicond Manuf 21(2):279–284. https://doi.org/10.1109/TSM.2008.2000270
31. Wang R, Zhang L, Chen N (2019) Spatial Correlated Data Monitoring in Semiconductor Manufacturing Using Gaussian Process Model. IEEE Trans Semicond Manuf 32(1):104–111. https://doi.org/10.1109/TSM.2018.2883763
32. Yu H, Ye L, Guo Y, Su S (2021) An Effective In-Field Calibration Method for Triaxial Magnetometers Based on Local Magnetic Inclination. IEEE Trans Instrum Meas 70:1–9. https://doi.org/10.1109/TIM.2020.3010671
33. Zhang L, Heaton D, Largey H (2005) Low cost multi-site testing of quadruple band GSM transceivers. In Proc. IEEE International Conference on Test 7–411. https://doi.org/10.1109/TEST.2005.1583999
34. Zhang Z-Q, Yang G-Z (2015) Micromagnetometer Calibration for Accurate Orientation Estimation. IEEE Trans Biomed Eng 62(2):553–560. https://doi.org/10.1109/TBME.2014.2360335
35. Zhao Y-G, Tong M-N, Lu Z-H, Xu J (2020) Monotonic Expression of Polynomial Normal Transformation Based on the First Four L-Moments. J Eng Mech 146(7):06020003. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001787
36. Zhikun S, Zurong Q, Chenglin W, Xinghua L (2015) A New Method for Circular Grating's Eccentricity Identification and Error Compensation. In Proc. Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC) 360–363. https://doi.org/10.1109/IMCCC.2015.83
37. Zou B, Xiao Q (2014) Solving Probabilistic Optimal Power Flow Problem Using Quasi Monte Carlo Method and Ninth-Order Polynomial Normal Transformation. IEEE Trans Power Syst 29(1):300–306. https://doi.org/10.1109/TPWRS.2013.2278986

Praise O. Farayola (Student Member, IEEE) received the B.Sc. degree in electrical and electronics engineering from the University of Ibadan, Ibadan, Nigeria, in 2015. He is currently pursuing the Ph.D. degree in electrical engineering with Iowa State University, Ames, IA, USA. In 2020 (summer), he was an Analog Circuit Design Intern with Texas Instruments, Dallas, TX, USA. He worked majorly on current feedback amplifier designs and verification with the High-Speed Operational Amplifier Team, Tucson, Arizona. His research interest includes multisite testing, test hardware error identification and calibration, low-cost high-precision ADC testing, analog circuit design, and fault coverage analysis for analog circuit test.

Isaac Bruce is a first-year student at Iowa State University's College of Engineering where he is pursuing a doctorate degree in Electrical and Computer Engineering with emphasis on Analog and Mixed Signal Circuits design and testing. He is currently focused on solving issues Multisite testing, a relatively novel technique in Analog circuit testing that drastically reduces test time and consequently test cost. Mr. Bruce has interests in analog circuit design and testing, RF circuit design and machine learning.

Shravan K. Chaganti (Member, IEEE) received the B.Tech. degree in electronics and communication engineering from IIT Roorkee, Roorkee, India, in 2014, and the Ph.D. degree in electrical engineering from Iowa State University, Ames, IA, USA in 2020. In 2016, he was an Intern with Skyworks Solutions, Cedar Rapids, IA, where he was involved in analog circuit design, verification, and test. In 2018, he interned at Texas Instruments, Sugar Land, TX, USA, where he worked on data converter linearity test time reduction. He is currently working with the C2000 MCU Team on new product development and test cost reduction. His research interests include low-cost high-precision SoC testing, analog circuit design-for-test, functional safety for automotive, and analog fault coverage simulation and improvement. Dr. Chaganti received the IEEE Ned Kornfield Best Paper Award in 2018.

Abalhassan Sheikh received the bachelor's and M.Sc. degrees in electrical engineering from the University of Houston, Houston, TX, USA, in 2003 and 2007, respectively. In 2007, he joined Texas Instruments, working on product and test engineering of SOC's targeted for the Automotive markets. In 2010, he joined the C2000 32-bit MCU Team working on new product development of SOC's comprising numerous IP's including ADC and DAC where his area of interests are in test cost reduction and data analysis. Mr. Sheikh was elected Member Group of Technical Staff in 2016. He received the IEEE Ned Kornfield Best Paper Award in 2018.

Srivaths Ravi (Senior Member, IEEE) received the B.Tech. degree in electrical and electronics engineering from IIT Madras, Chennai, India, and the M.A. and Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ, USA. He is the Hardware Development Manager for the C2000 microcontrollers business unit with Texas Instruments. He currently manages C2000's global RESEARCH AND DEVELOPMENT organization responsible for end-to-end silicon development— product strategy and roadmap, IP/SoC definition and architecture/specifications, new product silicon design, product/test engineering, as well as sustaining engineering activities. At TI, he has also played multiple leadership roles prior, including managing SoC platform, design, design-for-test (DFT), and customer-facing system application teams. He has also been a Senior Member of TI's Technical Staff (SMTS). He was a Research Staff Member with NEC Laboratories America, Princeton, where he worked in the areas of low power EDA and embedded security. He co-architected and designed the mobile security processor MOSES used in NEC's cellphone chips, which was also awarded the NEC Laboratories Technology Commercialization Award in 2005. He also held a visiting research collaborator position at Princeton University. He has over hundred publications in the form of conferences/journal contributions and book chapters, and has received 7 best paper awards and 13 granted patents. He has served on the organizing/program committees of various leading conferences, including as the General Chair of VLSI Design Conference 2010, Program Chair of Asian Test Symposium 2011, Program Chair of VLSI Test Symposium 2016–2017, and General Chair of TI Technical Conference in 2018.

Degang Chen (Fellow, IEEE) received the B.S. degree in instrumentation and automation from Tsinghua University, Beijing, China, in 1984, and the M.S. and Ph.D. degree in electrical and computer engineering

from the University of California at Santa Barbara, Santa Barbara, CA, USA, in 1988 and 1992, respectively. His industry experience includes the Beijing Institute of Control Engineering, Beijing, China, from 1984 to 1986; Faculty Fellow with the Boeing Company, Seattle, WA, USA, in 1999; Maxim Integrated, Dallas, TX, USA, in 2001; and Texas Instruments (TI), Tucson, AZ, USA, in 2011, 2012, and 2014. His academia experience includes the John R. Pierce Instructor of electrical engineering with the California Institute of Technology, Pasadena, CA, in 1992. Since 1992, he has been with Iowa State University, Ames, IA, USA, where he is currently a Professor of Electrical and Computer Engineering and holds the Jerry Junkins Chair with the College of Engineering. He has authored and coauthored over 300 refereed journal articles, conference papers, and patents. His current research interests include analog and mixed-signal integrated circuit design and testing, integrated circuit sensor design for life-time reliability and functional safety, and high-accuracy test without requiring high-accuracy instrumentation. Dr. Chen and his students' work has received many best paper awards and other honors, including the IEEE Ned Kornfield Best Paper Awards in 2013, 2014, and 2019, and the IEEE EJ McCluskey Best Ph.D. Dissertation Awards in 2017 and 2019, in recent years.