IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 24 June 2023, accepted 10 July 2023, date of publication 13 July 2023, date of current version 21 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3295123

METHODS

A Low Cost Self-Adaptive Screening Method Based on Automatic Test Equipment for Low Dropout Voltage Regulators in Mass Production

TAOYI ZHOU AND DAVID WEI ZHANG
State Key Laboratory of ASIC and System, School of Microelectronics, Fudan University, Shanghai 200433, China
Corresponding author: David Wei Zhang (dwzhang@fudan.edu.cn)

ABSTRACT The low-dropout (LDO) regulator is an important component in power management IC (PMIC) chips and is widely utilized in electronic modules, such as mobile processors and Internet of Things (IoT) devices. To ensure the proper functionality of these modules, automatic test equipment (ATE) is typically used to test and calibrate LDO modules in the mass production stage. During the test process, the presence of pollutants in the test environment or inappropriate test settings can cause a discrepancy between the expected performance of the chip and its actual output. When conventional ATE test methods fail to identify these types of risks, this study proposes a novel self-adaptive screening method that enhances the possibility of screening out weak chips without the need for additional testing hardware or stages. This screening method accumulates and analyzes effective test data measured by ATE in real time, and derives a dynamic test limit to screen weak chips by performing systematic statistical analysis on the collected data, combined with the chip's practical application requirements and tolerance. During the experimental stage, the author selected a chip lot and applied the proposed testing method. The results showed that 90.91% of escape defective chips were successfully screened, with a reasonable yield impact of 3.38%, compared to the traditional ATE test methodology which had a low capture rate.

INDEX TERMS Automatic test equipment (ATE), low-dropout (LDO) regulator, self-adaptive method, outlier elimination.

I. INTRODUCTION

With the increasing number and diversity of electronic device applications, especially for portable and Internet of Things (IoT) devices, there has been a skyrocketing demand for Integrated Circuit chips (ICs) that are utilized in the manufacturing of these devices. As one of the important components, power management IC (PMIC) chips play a crucial role in these types of electronic devices. They can regulate the overall power consumption of the device by changing the power supply voltage provided to the other modules, allowing electronic devices to perform various tasks with optimal power consumption. PMIC chips commonly consist of buck-boost converters or low-dropout regulators (LDOs). Buck-boost converters can both increase or decrease voltage, whereas LDOs only output voltages lower than the input voltage. Although buck-boost converters provide more voltage regulation, the simple circuit structure and control logic of LDOs make them popular for use in mobile devices [1].

As is commonly known, non-ideal factors may occur during the chip manufacturing process, resulting in differences between the produced chips and the original design. Therefore, it is necessary to use automatic test equipment (ATE) to test the chips during the mass production stage to ensure that most of the delivered chips can work properly, and the overall failure rate is below 500 parts per million (PPM). The failure rate is a reasonable threshold for consumer

The associate editor coordinating the review of this manuscript and approving it for publication was Marcelo Antonio Pavanello.

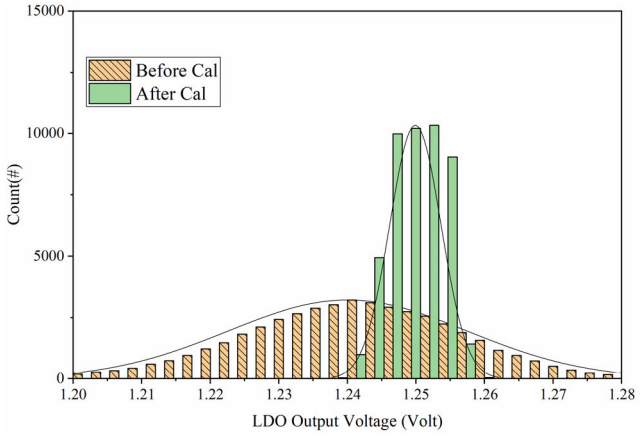


FIGURE 1. Examples of output voltage distribution of the target LDO chips before and after OTP calibration through ATE test.

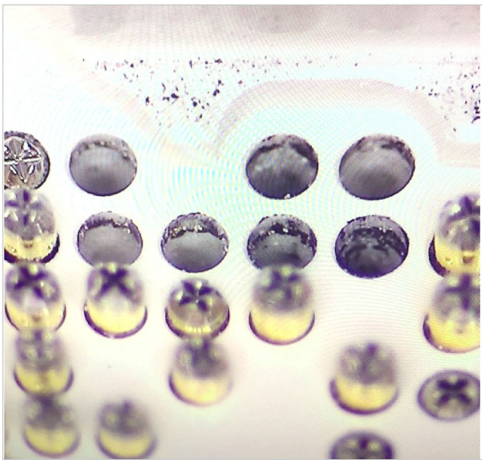


FIGURE 2. The presence of severe socket contamination can introduce abnormalities during the ATE test process, leading to a notable increase in the number of RMA chips.

electronics. Typically, we need to connect the output pins of the tested chip's LDO to the ATE testing channel, and apply a load current to the ATE testing channel to simulate the LDO under real application scenarios [2]. Furthermore, to ensure a more uniform output performance for chips that have undergone the ATE test, some LDO chip designs utilize One-Time Programmable (OTP) calibration. During the Final Test (FT), the output of the LDO under test was measured and calibrated, and the calibration value is written into the OTP area of the chip, resulting in a constant performance of the chip in practical electronic devices [3], [4]. Figure 1 shows a comparison of the output distribution of the target LDO chip SC2721G before and after calibration during the normal FT testing stage. As can be seen from the figure, the output distribution of the calibrated LDO chip was more concentrated.

However, there is a major defect in the OTP calibration method in practical tests, that is when the ATE environment itself is greatly affected (such as the socket being dirty, the chip under test not in close contact with the equipment, etc.) The trim code obtained by the test may differ from the value that the chip needs to be calibrated, causing the output voltage of the chip to deviate from the designed expectation when used in real applications. When this error is large enough, there is a possibility that a chip that passed the test will not work properly in practical applications, which may lead to chips being returned through Return Merchandise Authorization (RMA). Figure 2 displays a dirty socket during the ATE test and a significant number of RMA chips resulting from contamination.

Another scenario is that the output of the device under test (DUT) falls within the test limits set by the design but deviates from the majority of other DUTs. Such devices are known as outliers and, have the potential to cause product failures in real applications. Because the fact that the LDO output voltage of the DUT did not exceed the test limit during the FT phase, these chips were not screened during the FT test process. To minimize the probability of such occurrences, this study seeks to identify a mass production testing method suitable for ATE to effectively filter out potential risk chips.

II. CONVENTIONAL ATE TEST METHOD

To address the challenges identified in the previous section in mass production, researchers and engineers have proposed multiple solutions from different perspectives, including hardware and software. To improve the accuracy of the measurement hardware, a Kelvin connection structure was employed to connect the DUT and ATE testing equipment [5].

By enabling independent and closer contact with the output pins of the DUT, this approach achieved more accurate electrical performance measurements. In alternative approaches, special built-in self-test (BIST) structures were introduced into PMIC chip designs to enable self-testing, allowing these DUTs to obtain target electrical performance parameters more easily and reliably using straightforward digital protocols during the ATE test [6], [7]. All of these attempts are aimed at reducing the influence of external factors on the testing results during measurements, so that the values obtained by ATE can more accurately reflect the actual performance of the DUT.

The software method to deal with this issue is to learn and process the ATE testing results, identify chips with abnormal test results and screen them out to reduce the possibility of such chips being manufactured into products. For chips that cannot tolerate area overhead or complex test flow, this method is the only viable option, or it can serve as an effective additional check for application scenarios that prioritize chip quality. This concept evolved from rigorous quality standards required for automotive electronic chips, which spurred the development of multiple software-based methodologies such as the part averaging test (PAT), statistical bin limits (SBL) and geographical defect detection (GDD) [8]. SBL and GDD can carry out a comprehensive analysis of the DUT results without the need for an ATE before customer delivery. In contrast, the PAT method directly analyzes the raw testing

data of the DUT during the testing stage, thereby enabling the early identification of high-risk chips. Consequently, the PAT method plays a relatively more important role in ATE testing.

The first PAT method was proposed in [9]. This method screens the DUT through the following steps: (1) The initial ATE test is run and all testing values of the passed DUTs are collected. (2) The distribution range, including the majority of pass DUTs, was calculated, and the test limit was set to the same range. (3) Use the new test limit to screen the incoming fresh chips. (4) The test limit is periodically updated based on the test results on a monthly or quarterly basis. This method is also known as static PAT (SPAT) [22] because of its long update cycle and fixed test limits for each testing process.

To address this issue, researchers and engineers have shortened this feedback cycle from quarterly or monthly updates to a lot or wafer-based intervals, thus leading to the introduction of the dynamic PAT (DPAT) test method [10], [11]. A shorter update cycle enables better alignment between the ATE test limit and the DUT's actual performance, thus avoiding the inability of the static test limits to contain DUT's true performance due to manufacturing process fluctuations. Using a more robust statistical calculation method in the calculation process of the test limit for each round is called the Robust DPAT (RDPAT) method [12]. The key feature of this method is the ability to generate more effective and stable limits that can better handle test data under unexpected scenarios through high-intensity mathematical calculations, thereby reducing the probability of ATE testing errors such as test over kill or test under kill.

In recent years, owing to the rapid development of artificial intelligence technology, researches have used deep learning or machine learning algorithms to perform sample-based learning on the test results of DUTs when facing more complex test cases, and allow artificial intelligence to make a judgment on the DUT [13], [14], [15]. Some studies [16], [17], [18] utilized sequence prediction to predict the performance of the next DUT based on the collected sample sequence in the current testing environment. These methods can be used to screen problematic chips through iterative learning from the sample data and testing correction.

The ATE testing method mentioned above can be summarized as a process flowchart, as shown in Figure 3. The chart reveals that the different test methods share a similar core process. They first collected data, analyzed it, and then applied the updated test limits to achieve outlier elimination in the second test round. Advanced techniques can achieve higher efficiency in filtering or accuracy in screening.

These testing methods require two testing stages and rely on additional human intervention or data training, leading to relatively high ATE testing costs, particularly with added expenses related to test time and external equipment [19]. Although automotive electronic chips may be able to withstand the cost of such testing methods, they are too high for consumer electronics to bear. The aim of this study is to propose a cost-effective testing method that focuses on consumer electronics chips while still achieving optimal test quality.

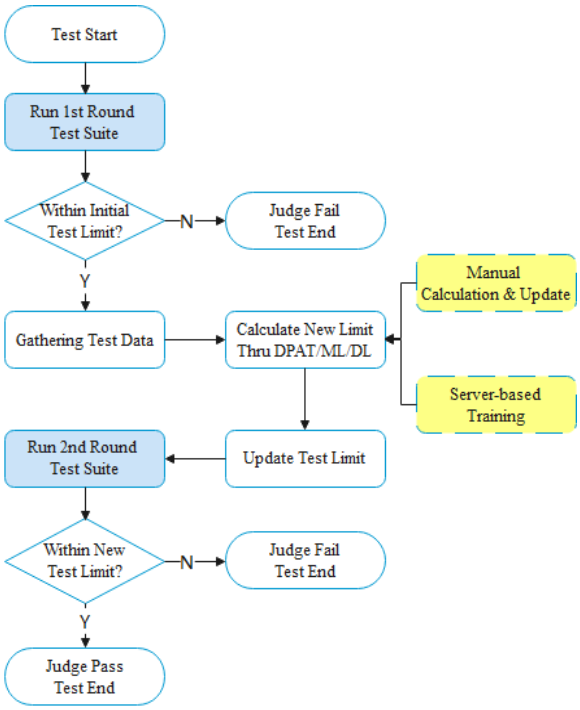


FIGURE 3. Common flow chart of conventional ATE test method. This type of flow requires two testing stage (blue blocks) and relies heavily on external resources (yellow blocks).

III. PROPOSED ATE TEST METHOD

To achieve low-cost outlier elimination for ATE testing of consumer electronic chips, we have made improvements and innovations based on the testing process and test limit generation method in traditional methods. The proposed method can be elaborated on from the following perspectives.

A. MAIN FLOW OF SELF-ADAPTIVE SCREENING METHOD

Given the high cost of ATE, adding a second round of testing to conventional methods can significantly increase the testing time and costs. Hence, it is crucial to perform risk-chip screening within a single round of testing. In addition, the processing of test data consumes additional resources: training on datasets requires external servers and updating test programs requires additional manual intervention. These steps not only require additional resources, but their implementation also presents certain difficulties. For commercial confidentiality, ATE production lines are not connected to the internet and cannot deploy servers. Furthermore, the introduction of additional human intervention during the testing process may result in unforeseen errors that can interfere with the analysis and research of chip testing results.

Hence, to address the negative factors mentioned earlier, the proposed testing method employs a single testing round with dynamic test limits to enhance testing quality. To evaluate chips, this method uses limited-length test data

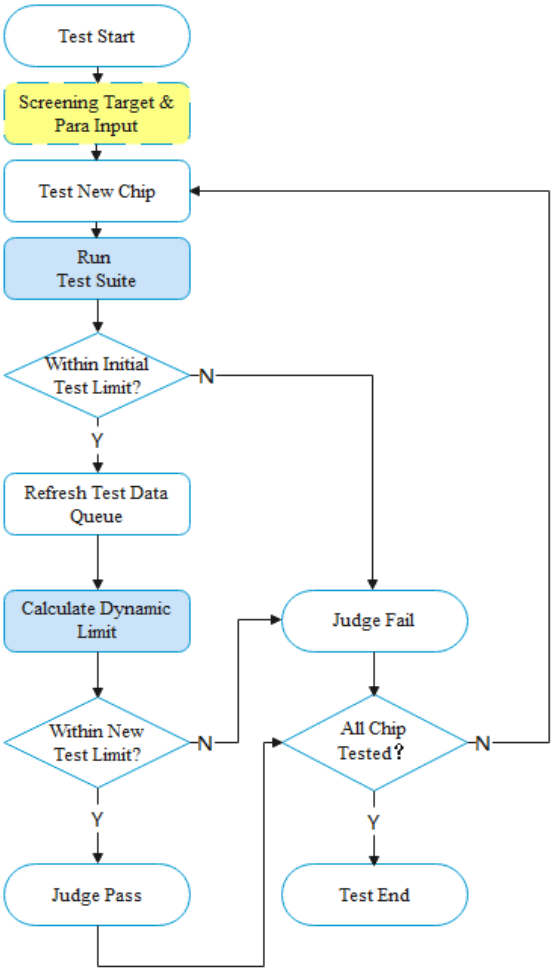


FIGURE 4. Flow chart of proposed ATE test method.

to dynamically generate two distinct test limits during a single testing run: the initial limit provided by the chip designer or data sheet and the dynamically generated limit based on the collected test data, input parameters, and target screening yield. This is accomplished using lightweight codes within the ATE program project. A flow-chart is illustrated in Figure 4.

Compared to the process illustrated in Figure 3, the proposed method tests each DUT only once, and applies dynamic test limits specific to each DUT instead of using the same test limit for the entire lot. This method relies solely on the targeted screening yield and partial system parameters as system inputs, and requires no further intervention during the testing process or external equipment such as additional servers. This method calculates a limited number of test data records for use as a database to ensure the timely convergence of the generated test limits without compromising performance.

B. SUB FLOW OF CALCULATING DYNAMIC LIMIT

The generation of the dynamic test limit is a critical aspect of outlier elimination and serves as the core module of the proposed method. Consequently, it is imperative to provide a detailed explanation of the process involved in generating the test limit.

TABLE 1. Parameters for limit generation.

Name	Description
C_{p_target}	Target C_{pk} for generated test limit
T_{queue}	Parameters to illustrate sample queue's trend
ADT	Threshold value for test limit compensation

As previously discussed, our objective is to eliminate two types of chips: those whose test values deviate significantly from the majority of chip test values and those whose calibration values deviate from the true values of the tested chip. For the first type of chip, filtering can be accomplished by computing the statistics of the test data queue and comparing the chip test values to this statistical value. By utilizing the central limit theorem, it can be assumed that chip test values follow a normal distribution with mean of μ and standard deviation of σ, as illustrated in Figure 1. According to the nature of the normal distribution, $[\mu-\sigma, \mu+\sigma]$ contains 68% of the samples and $[\mu-3\sigma, \mu+3\sigma]$ contains 99% of the samples. Consequently, we can evaluate whether the current DUT belongs to the majority by analyzing the relationship between the chip test value and the position of these ranges. This introduces the parameter C_{pl} and C_{ph} (C_{pk} to represent either of them), which are defined as follows:

$$C_{pl} = \frac{\mu - Low\ Limit}{3\sigma} \tag{1}$$

$$C_{ph} = \frac{High\ Limit - \mu}{3\sigma} \tag{2}$$

where μ and σ represent the mean and standard deviation of the test data queue obtained through the proposed test method, respectively, as illustrated in Figure 4. When $C_{pk} = 1$, it indicates that the sample mean is located in the exact center of the dynamic test limit range, and the width of this range is 6σ, thereby encompassing 99% of the pass chips. By modifying the value of C_{pk}, we can regulate the width and relative position of the dynamic test limit, which determines the percentage of chips rejected, thereby eliminating the outliers of the DUTs. Hence, C_{pk} is an essential input parameter for the test method proposed in this study.

One objective of ATE testing of consumer electronic chips is to avoid any sudden changes in the test yield. Thus, compensatory parameters must be introduced to ensure the stability of the yield performance of the proposed test method. The following parameters listed in Table 1 were introduced into the proposed test method:

Here T_{queue} will be determined by following equation:

$$T_{queue} = \frac{\sum_{i=1}^{L} (X[i] - X[i-1])}{L} \tag{3}$$

where L is the length of the test data queue, and the testing data are recorded in chronological order and in the form of

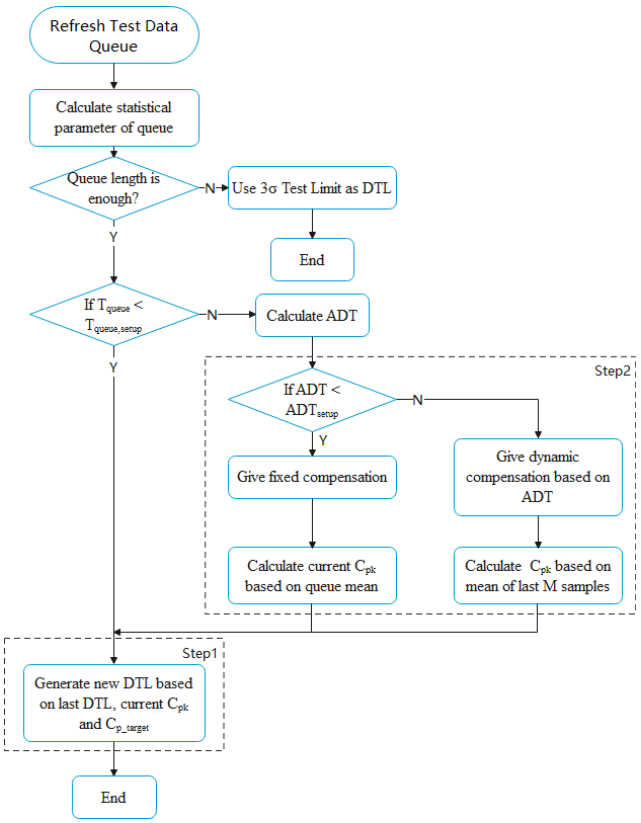


FIGURE 5. Flow chart of DTL generation.

a queue using $X[i]$. When L reaches the maximum value set in advance, the subsequent $X[i+1]$ data enters the test data queue, and $X[0]$ is removed. Using this method, the system was able to record changes in the test results over time. The system can self-correct for minor differences in the testing environment and issue a warning when significant abnormalities are detected in the testing environment.

The ADT can be defined by the following formula:

$$ADT = abs\left(\frac{\sum_{i=L-M}^{L}(X[i])}{M+1} - \mu\right) \tag{4}$$

where M is an integer smaller than L. This formula calculates the difference between the mean of the last M test data and the mean of all test data in the test data queue to detect sudden one-time jumps in the test data. These jumps are often caused by changes in the testing environment. When there is a jump, the dynamic test limit range must be immediately adjusted to avoid a high rate of test overkill, which can significantly affect the test yield.

We can use the parameters mentioned above to describe the process of generating the dynamic test limit (DTL), and the flow chart is shown as Figure 5.

Here C_{p_target}, T_{queue_setup} and ADT_{setup} are user defined input parameters at the beginning of the proposed test method. Every time the testing method receives new test data, it integrates the existing data in the queue with the new input data and calculates the C_{pk}, mean, standard deviation, and other statistical parameters for this dataset based on the previous round's DTL. The algorithm outputs a 6σ DTL range if there are not enough samples in the current queue, ensuring that the range covers the majority of the tested samples. Once sufficient samples are in the queue, the testing method initiates the compensation discrimination process.

The queue length at this point has a significant impact on the execution time of the entire test cycle. Because all the reference samples used in the calculation of the dynamic test limit are sourced from the queue, even for methods with O(n) complexity, an increase in n will lead to a linear increase in the execution time of the entire system. A smaller n can reduce the testing time, whereas a larger n can better reflect the overall state of the DUTs, avoiding the interference of some abnormal points with the overall evaluation of the DUTs by the system. Therefore, choosing an appropriate n according to the actual testing needs is critical.

If the T_{queue} does not exceed the preset value, this confirms that the current test fluctuation is within the range expected by the user, and the algorithm jumps to step 1. The algorithm then generates a new DTL based on the previously calculated statistical parameters, with the benchmark of making the newly generated DTL result in a C_{pk} value that is closer to the C_{p_target} set by the user. This is illustrated by the following equations:

$$DTL_{new} = (1-\alpha) \times DTL_{last} \tag{5}$$

$$\alpha = \begin{cases} -\alpha_0, & C_{pk} < C_{p_{target}} \\ \alpha_0, & C_{pk} \geq C_{p_{target}} \end{cases} \tag{6}$$

where α_0 in (6) is a non-negative variable real number. These two equations indicate that when the previous DTL range is slightly wider than expected, the algorithm adjusts the range of the DTL to decrease the C_{pk} value, and vice versa.

When T_{queue} exceeds the preset value, the proposed testing method proceeds to step 2 and performs additional compensation to avoid producing a large number of test overkills. Initially, the algorithm expands the limit range to a safety value by using the following equation:

$$DTL_{s1} = \begin{cases} (1+C_0) \cdot DTL_{last}, & ADT \leq ADT_{setup} \\ f(ADT) \cdot DTL_{last}, & ADT > ADT_{setup} \end{cases} \tag{7}$$

After expanding the previous DTL, the algorithm calculates a fresh C_{pk} based on the expanded range, with the aim of ensuring that the updated C_{pk} value reflects the most recent condition of the test dataset. To achieve this, the mean used to compute C_{pk} values will be different, and will be determined by the following equation:

$$C_{pk} = \begin{cases} \dfrac{|\mu - DTL_{s1}|}{3\sigma}, & ADT \leq ADT_{setup} \\ \dfrac{|PM - DTL_{s1}|}{3\sigma}, & ADT > ADT_{setup} \end{cases} \tag{8}$$

$$PM = \frac{\sum_{i=L-M}^{L}(X[i])}{M+1} \tag{9}$$

TABLE 2. Performance comparison of various dynamic test screening methods.

Key Parameter	This work	[21]	[12]	[10]	[14]
Base Method	DPAT	SPAT	RDPAT	DPAT	ML
Stage Count	1	1	2	3	7
Computational intensity	Low	Low	High	Medium	High
Fluctuation compensation	Yes	No	Yes	No	Yes
Capture Expectation	Normal	Low	High	Normal	Extreme

Equations (7) and (8) are used to obtain the updated DTL and C_{pk} values when the test dataset exhibits jumps. These values are then substituted into Equations (5) and (6) to calculate the most recent round of the DTL.

The second factor that may affect the execution time is the complexity of the calculation. When computing the next dynamic test limit, we can use a simple calculation formula, as described in this article, to roughly estimate the correct performance of the DUT, or we can use convolutions and matrix operations that are widely used in machine learning to enhance the confidence in the obtained data. However, using complex calculation methods will result in an additional execution time of O(n) or even O(n^2), which may not be useful for chips in consumer electronics products that do not require extremely high product quality.

If ADT or T_{queue} exceeds reasonable values by a significant margin, this implies that the environmental influence surpasses the correction limits of the algorithm. As a result, the algorithm terminates ATE testing to notify the user to inspect and fix the test environment. Given that the test limit automatically updates under specific conditions, we refer to this approach as self-adaptive.

In summary, the self-adaptive ATE test method proposed in this paper has the following significant advantages compared to other methods: (1) It requires only one testing process to achieve the goal of accurate dynamic screening. (2) The computational intensity is relatively low compared with other methods that involve a large number of square and square root operations. (3) It has a certain resistance to abnormal fluctuations because we incorporate fluctuation compensation factors into this method, which makes the dynamic test limit more responsive to actual production data. A brief comparison is presented in Table 2. The other test methods referenced in the table are all dynamic screening test methods that have been derived from base methods with improvements, and some of these test methods have been used in actual production by companies such as Ampleon or Philips.

In the software implementation process, we created a C++ class to record the test data and implement all the functions mentioned above. Because the target ATE software platform also uses a C++ architecture, our code can perfectly fit this ATE testing platform. It should be noted that, in the current ATE software architecture, different test items belong to different test classes. To achieve data interaction between test items, we set the test data queue to a static type. This ensures that all test items can access the test dataset under the necessary conditions, to calculate various compensation parameters and dynamic test limits.

The core class methods include the following: (1) a method for refreshing the static test result queue; (2) a method for calculating the statistical parameters of the test result queue and system compensation parameters; (3) a method for executing compensation and calculating dynamic test limits; and (4) a method for providing feedback on dynamic test limits to the ATE system and updating the test method's own parameter records.

It should be noted that there are several obvious factors limiting to the performance of this testing method that need to be considered.

First, compared with the rigorous mathematical calculations used in machine learning, the screening efficiency of the simple algebraic method used in this study is not as good as that of machine learning. This means that in scenarios that demand strict quality requirements for industrial electronic products, the proposed method would result in a serious overkill if we want to achieve the same screening results.

The second limitation is that the current system still requires users to input reliable initial conditions for the system to operate normally, and the calculation of these parameters must be externally calculated and input by people. However, the system cannot automatically calculate them based on historical samples.

Third, the current testing method adopts the same testing process for all DUTs without selecting characteristics, whereas other multistage testing methods in previous research have selected the testing process to some extent for the DUTs being tested. This means that if the proportion of abnormal chips in the target lot is very low, the testing time loss of the current process will be significantly greater than that of multistage processes. This characteristic makes the method proposed in this study more suitable for consumer electronics projects with shorter flows and certain quality requirements for chips. For chips with long testing flows and stricter quality requirements, this method may consume unexpected overhead test times and fail to achieve desired results.

C. DETECTION OF LDO CALIBRATION OFFSET

As previously mentioned, another goal of the method proposed in this paper is to effectively screen chips with a potential risk for LDO calibration anomalies. The previous analysis indicated that dirty test sockets are a significant cause of calibration anomalies. Testing the same lot of chips on two the ATE machines shows that some chips may have lower test values when the testing environment is abnormal; however, the overall distribution of each test run still indicates a normal distribution, as shown in Figure 6.

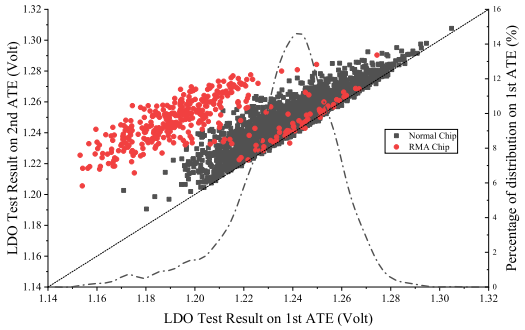


FIGURE 6. Result of LDO test on different ATE and the distribution of 1st ATE test (with anomalies).

Ideally, the test results should be the same for two ATE testing sessions, and the line formed by plotting the scatter of the two test results should follow the formula $y = x$. However, the figure above shows that the distance between the scatter points of most RMA chips and the $y = x$ line is noticeably larger than that of normal chips. This observation could serve as a valuable starting point for further analysis.

The voltage offset suggests a voltage drop in the testing circuit, which occurs because of the accumulation of dirt inside the socket. Thus, there exists an opportunity to detect sockets that are significantly dirty during the testing process, provided we are able to capture this voltage drop in ATE testing. According to the characteristics of the LDO, increasing the load current within the designated range hardly affects the output voltage of the LDO [20]. As the voltage drop generated by resistive dirt significantly increases with increasing current, we aim to capture this information by separately measuring the voltage output of the LDO under half and full-loads during testing.

Assuming that the LDO output voltage measured by ATE is V_1 under load I_1 and V_2 under load I_2 when the resistance R of the resistive dirt is constant, we can obtain the following equation:

$$\Delta V = -\Delta I \times R \tag{10}$$

Therefore, the contact resistance R can be obtained indirectly from the difference between the two measurements. According to our experiment and the raw data required, Figure 7 shows that the two measurements of the output voltage at half-load and full-load for a given LDO can be approximated to follow a linear relationship of $y = kx + b$.

If we can assume that the output voltage of the LDO between two current-load conditions is negligible, and that the output voltage under load I_1 is V_1 and under load I_2 is V_2, then the following equation is valid:

$$\begin{cases} V_1 = V_{out} - I_1 R \\ V_2 = V_{out} - I_2 R \end{cases} \tag{11}$$

$$\begin{cases} I_2 = 2I_1 \\ V_2 = kV_1 + b \end{cases} \tag{12}$$

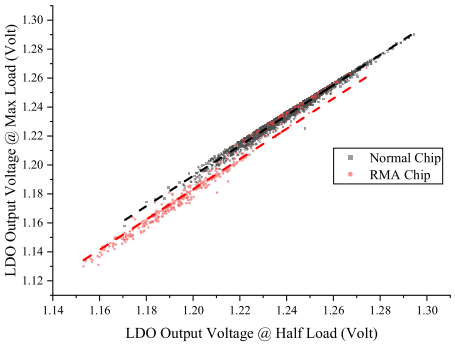


FIGURE 7. Result of LDO output voltage under half load and max load, and the linear regression of these two results.

TABLE 3. Parameters for linear regression.

Parameter	Normal Line	RMA Line
k	1.0442	1.0442
b	-0.0698	-0.0606
R^2	0.9709	0.9890

When Equations (11) and (12) are combined, the relationship between R and b is given by Equation (13):

$$b = (1-k)\,V_{out} - \left(1 - \frac{1}{2}k\right) I_2 R \tag{13}$$

Equation (13) can be rewritten in the following format because V_{out} can be considered a constant:

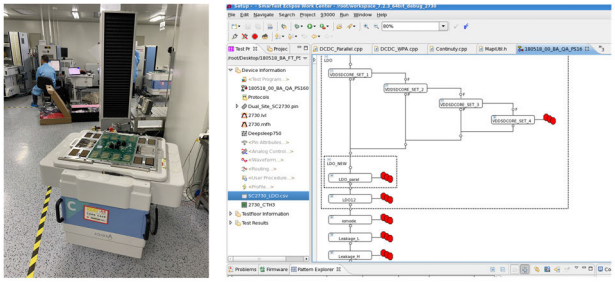
$$\Delta b = \left(\frac{k}{2} - 1\right) \cdot I_2 \Delta R \tag{14}$$

The correlation between b and R is linearly related, and this enables us to identify abnormal R values by constructing parameter b and applying the testing method proposed in this paper. From Figure 7, it is evident that a discrepancy exists between the intercepts of the linear fits of the RMA chips and those of normal chips, when the regression slope is kept constant. Further details regarding this observation are provided in Table 3.

IV. EXPERIMENTAL RESULT

During the experimental validation stage, we selected the PMIC chip SC2721G for production verification, and the corresponding experimental setup is shown in Figure 8. The figure indicates that the on-site production line does not require any ancillary equipment, and the test method presented in this paper can be easily synthesized into the tester software as well. Consequently, additional test hardware or software is rendered unnecessary.

In the original test, due to testing environment factors such as dirty sockets, the tested chip lot did not fully reflect the actual output performance of the chip in terms of LDO output voltage. We reconstructed the LDO output voltage during both the ATE testing and actual application stages based on the Unique Identification (UID) information of the chip recorded on the chip OTP. The results are shown in Figure 9.



(a) Tester Hardware (b) Tester Software

FIGURE 8. Experimental environment hardware and software.

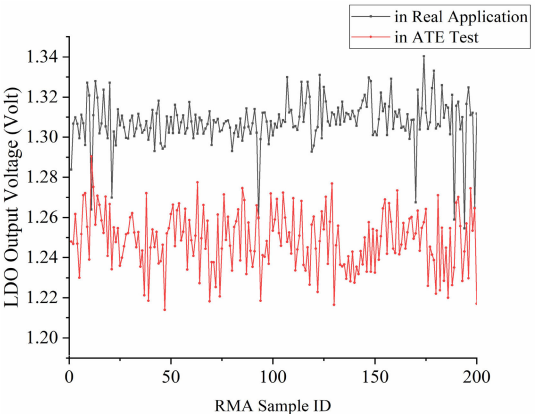


FIGURE 9. Relationship between ATE test result and real application output for same RMA chips.

The above figure demonstrates that the LDOs of most RMA chips have a higher output voltage at the actual application end, which is caused by the underestimation of their output voltage during ATE testing. During testing, an additional compensation voltage was applied to the chips, which turned out to be too large because the actual output voltage of the chips was higher than that obtained during testing. The initial purpose of this compensation voltage is to make the output of the chips more uniform. As a result, these chips output a voltage higher than that originally expected for actual use. According to observations, when the output voltage of the chip increases anomalously by approximately 50mV, the actual product encounters usage issues. However, the 50mV output deviation cannot be detected by the initial ATE test program because the screening range of the SPAT test limit is far greater than this value. The abnormal performance of the entire chip lot is illustrated in Figure 10.

In order to detect and screen out these potentially risky abnormal chips during the ATE stage, the testing method proposed in this paper has been integrated into the current testing program. All the important parameter values used in this method can be obtained from Table 4. We use this parameter set and the flow illustrated as Figure 5 to generate the first dynamic screening result.

After testing the chips of the sample lot using the testing method proposed in this paper, we obtained the chip test value, as well as the dynamic low limit (DLL) and dynamic

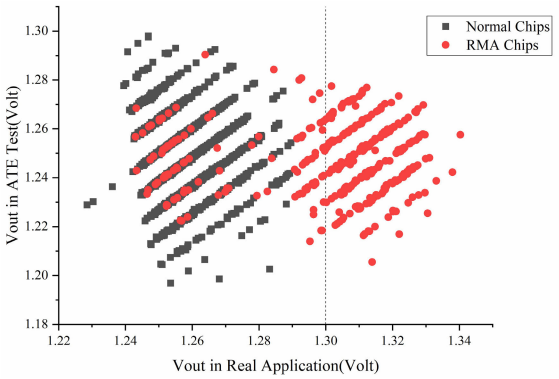


FIGURE 10. Overall performance of normal chips and RMA chips in target chip lot.

TABLE 4. Key parameters used in proposed test method.

Parameter	Description	Value
Queue Length	Number of samples kept in queue	50
Queue Valid	Valid queue length to go into DTL	25
C_{ph}	Target C_{ph} compared to high limit	0.67
C_{pl}	Target C_{pl} compared to low limit	0.67
ADT_{setup}	Threshold value for compensation	0.15

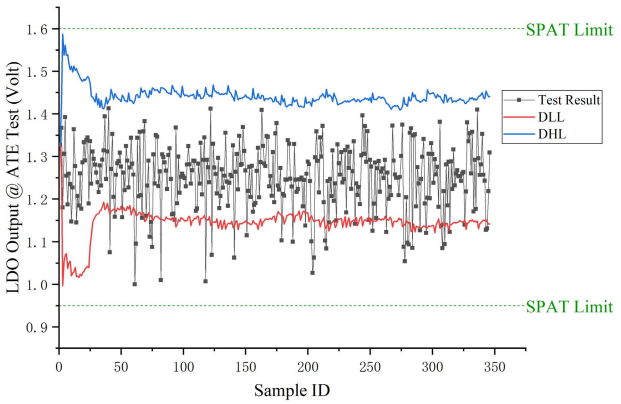


FIGURE 11. Relationship between ATE test result and dynamic test limits.

high limit (DHL) generated based on the chip test value, as shown in Figure 11.

As depicted in the figure, the DTL range was relatively wide during testing of the initial few DUTs. As the number of samples in the test data queue became sufficient, DTL started to converge rapidly for dynamic screening. In this particular case, because failure was more likely to result in a lower LDO output voltage, the C_{pl} was set smaller than the C_{ph} when configuring the system parameters. This is reflected in the figure, where the DLL is closer to the test data than the DHL is.

This sample lot was selected from an abnormal lot that had undergone routine ATE testing, indicating that the conventional ATE testing program failed to screen out potentially defective chips. On the other hand, our proposed testing method successfully identified some abnormalities.

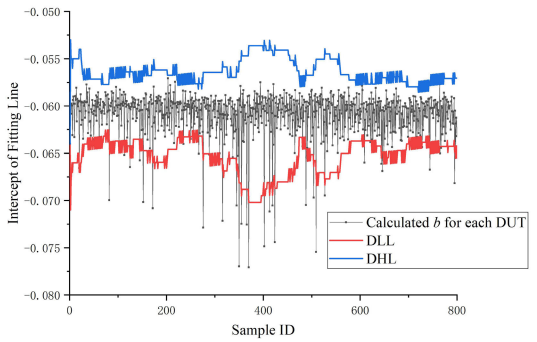


FIGURE 12. Result of sample lot based on proposed test method, $k = 1.0442$.

TABLE 5. Performance of proposed test method based on Figure 12.

Parameter	Value	Parameter	Value
Total Tested	1090	Pass DUT	1031
Overkill DUT	37	Miss DUT	4
Catch DUT	18	Catch Rate	81.82%

It can be observed from Figure 6 that utilizing the proposed testing method presented in this paper on the unprocessed output voltage data of the LDO may lead to considerable test over kill or test under kill. This is because the test data of the RMA chips is highly comparable to the mean test data of the regular chips.

Thus, during the experimental phase, we employed the testing method proposed in this study by selecting a specific fitting line slope based on Equation (14), with intercept b as the target of analysis. In this experimental process, the fitting line slope k, together with C_{p_target}, T_{queue} and other related parameters, served as an initial input parameter. The result are shown in Figure 12.

It can be seen that there is a lot of small jitter in the generated DTL at this parameter setting. This is because, during the calculation process, the sample value happens to be within the loop of Judgment Step 2 in Figure 5, and the system switches between dynamic compensation and fixed compensation. The specific performance of the testing system proposed in this study under this parameter is presented in Table 5.

To evaluate the effectiveness of this testing method, we present two parameters, namely, the Hit Rate (HR) and Actual Yield (AY), which are defined as follows:

$$HR = \frac{Captured\ Fail}{Actual\ Fail} \tag{15}$$

$$AY = 1 - \frac{Overkilled\ Fail}{Total\ Count - Actual\ Fail} \tag{16}$$

"Captured Fail" refers to the DUTs that truly failed and were identified by the proposed test method, "Actual Fail" refers to the DUTs that truly failed and were tested in the laboratory, and "Overkilled Fail" refers to the DUTs that were screened out by the proposed test method but were actually good dies. Hence, HR reflects the capability of the testing method to capture abnormal chips, whereas AY represents the efficiency of the testing method in capturing such chips. To comprehensively evaluate both these abilities, we introduce parameter M as an additional evaluation metric, which is defined as follows:

$$M = HR \times AY \times 100 \tag{17}$$

In an ideal scenario, HR could achieve 100%, indicating that all failed chips can be captured by the screening method, and AY would also be 100%, indicating that all failures are true fail DUTs, and no "false" errors have been mistakenly captured. In the example shown in Figure 12, the HR is 81.82%, AY is 96.54%, and the value of M is 78.98. Because C_{p_target} determines the statistical properties of the chips to be retained, it cannot be adjusted, whereas T_{queue} has been found to have a limited impact on M in other experiments. Therefore, for the current testing scenario, the k value is a crucial metric. Table 6 lists the performance metrics of the proposed testing method for different k values:

Based on the experimental results shown above, it can be seen that the system has an HR of 81.82% when k equals 1.04 or 1.06; however, when k is further increased, the HR decreases to 77.27%. This implies that an excessively large k value cannot effectively screen potential abnormal chips. Neglecting the negative intercept compensation, a larger value of k implies that the DUT has a higher full-load output voltage at the same half-load output voltage, clearly deviating from the normal LDO output characteristic curve. Thus, it is logical that an excessively large k value results in a decrease in the M value of the testing system, according to both the experimental results and physical theory.

TABLE 6. Performance of proposed test method based on different k value.

k value	HR	AY	M
1.06	81.82%	96.91%	79.29
1.08	77.27%	97.00%	74.96
1.10	77.27%	95.22%	73.58
1.02	90.91%	96.62%	87.84
1.00	90.91%	95.97%	87.25

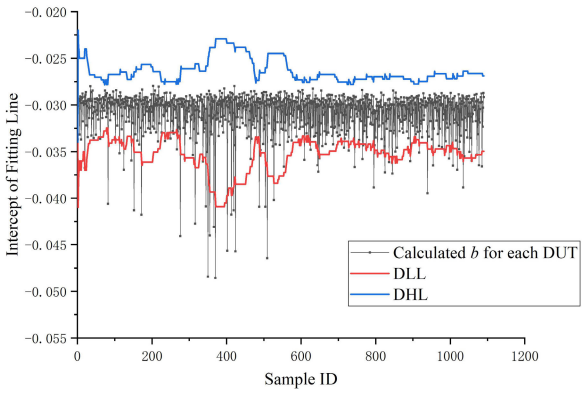


FIGURE 13. Result of sample lot based on proposed test method, k = 1.02.

TABLE 7. Performance of proposed test method based on Figure 13.

Parameter	Value	Parameter	Value
Total Tested	1090	Pass DUT	1032
Overkill DUT	36	Miss DUT	2
Catch DUT	20	Catch Rate	90.91%

TABLE 8. Performance comparison of different testing method.

Parameter	This work	Original	[12]	[10]
Capture Rate	90.91%	0%	34.6%	82.7%
Yield Loss	3.30%	0%	8.90%	N/A
Overhead Time*	10%	0%	N/A	133%

Selecting a k value of 1.02 or 1 indicates that the output characteristic curve of the LDO is considered the primary factor, whereas the uncertain additional factor is the contact resistance introduced by dirty socket contact. Experimental results indicate that both k value have a high HR but a k value of 1 has a higher overkill rate than a k value of 1.02. This is due to the fact that in the measured LDO half-load to full-load characteristic curve, the intercept is always negative. Therefore, a k value of 1 implies that even with an ideal socket contact, the full-load output of the LDO decreases significantly compared to the half-load output, which deviates from the actual test conditions.

After considering various experiments, we ultimately selected a k value of 1.02 in the actual production of the 2721G, achieving a capture rate of 90.91% of potential risk chips while ensuring that the proposed testing method has an impact on yield of less than 3.38%, compared to the zero capture rate of the risk chips produced by the SPAT ATE test method in previous production. The results are shown in Figure 13. The specific performance of the testing system proposed in this study under this parameter is presented in Table 7.

Compared to the initial testing method and some of the referenced testing methods that listed detailed testing performance, the current testing method showed competitive testing performance for the cases encountered in this chapter. The relevant comparisons are shown in Table 8.

It should be noted that for overhead time, because the testing method designed in this study only performs one test for all DUTs, using this system will lead to an overall increase in testing time. For the testing method described in [10], three tests were performed only when abnormal chips were detected. The overhead time presented in the table represents the ratio of the testing time for abnormal chips to that for normal chips.

In general, there are two common ways to utilize chips screened by the dynamic test method. The first is to conduct retesting, where the details are carefully checked and the test environment is finely set up to recycle chips that can work properly. The second is to sell these degraded chips to customers who require less stringent quality standards, to improve the yield and maximize the utilization of the finished chip products.

V. CONCLUSION

The basic ATE testing method of screening chips using fixed test limits is no longer sufficient to meet the demands of chip production as the demand for chip quality continues to rise. To address this problem, various solutions have been proposed using traditional testing methods. Unfavorable factors, such as excessive test time overhead, the need for additional hardware devices, or complex software implantations, hinder the implementation of methods such as RDPAT and machine learning, especially in the mass production of consumer electronics.

By performing laboratory analysis on the failed samples, we found that it was difficult to completely screen out abnormal chips. However, if we only aim to eliminate 90% of them, we can avoid using current complex ML algorithms. This is an acceptable RMA PPM level for consumer electronic products. The testing method proposed in this study can effectively eliminate potential risk chips during the ATE testing stage with lower test cost overhead, thereby significantly enhancing the product quality of consumer electronic chips.

The key advantage of this method lies in its simple mathematical model which can be directly integrated into the ATE tester software and hardware. This allows rapid deployment and has parameters that can be manually adjusted within a certain range to optimize the system performance, which can be used to adjust the dynamic test limit in real time to achieve fine-grained screening at the per DUT level.

Through the analysis of the characteristics of LDO chips, we established a simple physical model to create an effective physical quantity for identifying potentially risky chips. With fine-tuning and experimental testing, we were able to eliminate 90.91% of the potential risk chips with minimal test cost overhead.

This method is not limited to the testing of LDO chips, as the ATE testing method proposed in this study for generating dynamic test limits can effectively screen potential risk chips for other parameter types. For instance, it can be applied to the RF output power of RF chips and the run time frequency of digital ASIC chips, among others. Given its potential to test different types of chips, this testing method can serve as a versatile and cost-effective approach for enhancing testing quality, thus making it ideal for the mass production of diverse types of consumer-grade chips.

REFERENCES

[1] J. Park, B. Lee, and S.-W. Hong, "An output capacitorless low-dropout regulator with a low-VDD inverting buffer for the mobile application," *IEEE Trans. Ind. Electron.*, vol. 67, no. 10, pp. 8931–8935, Oct. 2020.

[2] M. Burns, G. W. Roberts, and F. Taenzler, *An Introduction to Mixed-Signal IC Test and Measurement*. New York, NY, USA: Oxford Univ. Press, 2001.

[3] G. Croce, "Non volatile memory in advanced smart power technology: Product requirements and integration solutions," in *Proc. IEEE Int. Memory Workshop (IMW)*, May 2022, pp. 1–4.

[4] L. Li, P. Xiao, Z. Lu, Y. Wang, Y. Hu, Y. Guan, K. Wang, D. Li, and F. Zhang, "Testing and trimming methods of high-resolution and large swing for ADC based on ATE," in *Proc. IEEE 14th Int. Conf. ASIC (ASICON)*, Oct. 2021, pp. 1–4.
[5] K. Luehman, "Applying Kelvin measurements to PMIC testing on ATE," *Eval. Eng.*, vol. 46, no. 11, pp. 32–37, 2007.
[6] A. Tulsiram, "Development of power management IC test solutions for ultra low defects," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Florida, Gainesville, FL, USA, 2021.
[7] A. Tulsiram and W. R. Eisenstadt, "Design for testability of low dropout regulators," in *Proc. IEEE 39th VLSI Test Symp. (VTS)*, Apr. 2021, pp. 1–7.
[8] B. Gessner, "How to ensure zero defects from the beginning with semiconductor test methods," in *Proc. IEEE Int. Test Conf.*, Oct. 2007, pp. 1–2.
[9] *Guidelines for Part Average Testing*, document AEC-Q001 Rev-C, Council, Automotive Electronic, 2003.
[10] W. Dobbelaere, R. Vanhooren, W. De Man, K. Matthijs, A. Coyette, B. Esen, and G. Gielen, "Analog fault coverage improvement using final-test dynamic part average testing," in *Proc. IEEE Int. Test Conf. (ITC)*, Nov. 2016, pp. 1–9.
[11] T. Sakamoto, K. Yofu, and T. Kyuho, "New method of screening out outlier; expanded part average testing during package level test," *IEEE Trans. Semicond. Manuf.*, vol. 30, no. 4, pp. 351–356, Nov. 2017.
[12] M. Moreno-Lizaranzu and F. Cuesta, "Improving electronic sensor reliability by robust outlier screening," *Sensors*, vol. 13, no. 10, pp. 13521–13542, Oct. 2013.
[13] C.-H. Chuang, K.-W. Hou, C.-W. Wu, M. Lee, C.-H. Tsai, H. Chen, and M.-J. Wang, "A deep learning-based screening method for improving the quality and reliability of integrated passive devices," in *Proc. IEEE Int. Test Conf. (ITC)*, Nov. 2020, pp. 1–9.
[14] N. Xama, M. Andraud, J. Gomez, B. Esen, W. Dobbelaere, R. Vanhooren, A. Coyette, and G. Gielen, "Machine learning-based defect coverage boosting of analog circuits under measurement variations," *ACM Trans. Des. Autom. Electron. Syst.*, vol. 25, no. 5, pp. 1–27, Sep. 2020.
[15] T.-H. Tsai, Y.-C. Lee, and C.-Y. Hsieh, "Enhancing the data analysis in IC testing by machine learning techniques," in *Proc. 14th Int. Microsyst., Packag., Assembly Circuits Technol. Conf. (IMPACT)*, Oct. 2019, pp. 183–186.
[16] A. Namdari and Z. S. Li, "An entropy-based approach for modeling lithium-ion battery capacity fade," in *Proc. Annu. Rel. Maintainability Symp. (RAMS)*, Jan. 2020, pp. 1–7.
[17] J. Lee and M. Mitici, "Deep reinforcement learning for predictive aircraft maintenance using probabilistic remaining-useful-life prognostics," *Rel. Eng. Syst. Saf.*, vol. 230, Feb. 2023, Art. no. 108908.
[18] A. Namdari and Z. S. Li, "A multiscale entropy-based long short term memory model for lithium-ion battery prognostics," in *Proc. IEEE Int. Conf. Prognostics Health Manage. (ICPHM)*, Jun. 2021, pp. 1–6.
[19] C. Streitwieser, "Real-time test data acquisition and data processing enabling closed loop control systems for adaptive test," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2016, pp. 902–905.
[20] N. Kularatna, *Modern Component Families and Circuit Block Design*. Amsterdam, The Netherlands: Elsevier, 2000.
[21] K. K. S. Lim, M. E. B. Francisco, and G. L. R. F. Abug, "Moving limits: A more effective approach in outlier screening at final test," in *Proc. IEEE 24th Electron. Packag. Technol. Conf. (EPTC)*, Dec. 2022, pp. 806–810.
[22] L. Fang, M. Lemnawar, and Y. Xing, "Cost effective outliers screening with moving limits and correlation testing for analogue ICs," in *Proc. IEEE Int. Test Conf.*, Oct. 2006, pp. 1–10.

TAOYI ZHOU received the B.S. degree in microelectronics from Shanghai Jiaotong University, Shanghai, China, in 2013, and the M.S. degree in electronic circuit and system from the University of California at San Diego, San Diego, CA, USA, in 2015. He is currently pursuing the Ph.D. degree in microelectronics with Fudan University, Shanghai. He is also with UNISOC as an ATE Test Engineer, from 2015 to 2021.

DAVID WEI ZHANG received the B.S., M.S., and Ph.D. degrees in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 1988, 1991, and 1995, respectively. He was an Associate Professor with Fudan University, Shanghai, China, in 1997, where he has been a Full Professor, since 1999. He is currently the Chairperson of the Department of Microelectronics and the Director of the Fudan-Novellus Interconnect Research Center.