Identification of Suspicious Semiconductor Devices Using Independent Component Analysis with Dimensionality Reduction

Jenny Bartholomäus

Infineon Technologies Dresden GmbH
Dresden, Germany
jenny.bartholomaeus@infineon.com

Swen Wunderlich

Infineon Technologies Dresden GmbH
Dresden, Germany
swen.wunderlich@infineon.com

Zoltán Sasvári

Technische Universität Dresden
Dresden, Germany
zoltan.sasvari@tu-dresden.de

Abstract – *In the semiconductor industry the reliability of devices is of paramount importance. Therefore, after removing the defective ones, one wants to detect irregularities in measurement data because corresponding devices have a higher risk of failure early in the product lifetime. Furthermore it would be desirable to consider multiple functional tests together due to existing dependencies. This paper presents a method to detect such suspicious devices where the screening is made on transformed measurement data. Additionally, a new dimensionality reduction is performed within the transformation so that the reduced and transformed data comprises only the informative content from the raw data. Therefore the complexity of the subsequent screening steps is simplified.*

I. INTRODUCTION

Semiconductor devices used in key product components such as motor controllers or in safety relevant applications such as airbags have to fulfill high quality requirements [1]. Therefore, many product tests at different operating conditions and temperatures are performed to check the functionality of these devices. Some of the tests have an analog output signal. The measurement data of these analog tests have to stay within predefined limits. Otherwise, corresponding devices will be scrapped since they are assumed to be defective. However, devices which pass the scrapping process are not necessarily “good” devices. Some of them are outliers with regard to the distribution of all passed devices. Those devices, called suspicious devices, are more likely to fail within the product lifetime. Hence, such devices should not be delivered to the customer.

A few tests exist which especially concern the reliability of the devices, e.g. IDDQ (Direct Drain Quiescent Current) tests. To avoid the shipping of suspicious devices several, commonly one-dimensional screening methods are applied to such tests. Nevertheless, the fact that there are tests which are dependent on each other and the fact that measurements can be superpositions of several signals (e.g. process-related, noise) necessitates advanced screening methods. In the future, the increasing complexity and the miniaturization of devices additionally complicate the screening due to an increasing amount of data which has to be handled.

Some approaches to improve the detection of suspicious devices and to detect them as early as possible are proposed in [2] – [6].

One approach is a data transformation using Independent Component Analysis (ICA) in order to facilitate the visibility of abnormalities in the measurement data. It was shown in [5] that this method improves the detection of suspicious devices. There it was assumed that the tests are not completely independent from each other and that their measurement data consist of superpositions of signals caused by the manufacturing process, noise signals and signals indicating irregularities in the data. ICA is able to separate those mixed signals into their independent source signals, making it possible to discover hidden effects. However, not all source signals have the same information content to be suitable for the detection of suspicious devices. The problem of selecting the most relevant components for outlier detection arises. This addresses the problem of a suitable dimensionality reduction during ICA which was not investigated so far in the context of semiconductor-related measurement data.

This paper proposes a method to determine the number of informative source signals which should facilitate the detection of suspicious devices. Furthermore, it shows the application of this approach combined with a one-dimensional outlier detection method and how the screening can be improved.

II. SCREENING METHODS IN SEMICONDUCTOR INDUSTRY

The main objective of screening methods in semiconductor industry is to detect irregularities in measurement data and to reject the identified suspicious devices since those are prone to failure and therefore should not be delivered to the customer. Screening methods are not only applied in order to check the actual functionality but rather to avoid lifetime failures of the product.

Fig. 1. The bathtub curve illustrating the failure rate.

The most delicate period is the early lifetime because failures occur much more frequently during this period [7, chapter 4]. The bathtub curve in Fig. 1 illustrates the failure rate of a device over its lifetime. The first period, characterized by a decreasing failure rate, represents the early lifetime. The second period, characterized by an approximately constant failure rate, represents the product lifetime assured to the customer.

A burn-in is used in order to remove early failures. There the devices are subjected to defined pattern at high temperature in a so called burn-in oven in order to put the devices under stress and to “simulate” the early lifetime which leads to a cut-off of the early lifetime. During this process and afterwards the devices are tested again to identify those which have been damaged through the burn-in procedure and those whose properties have been changed significantly. The identified devices will be scrapped. Burn-in testing is a well-known and effective method to detect early failures. However, a long testing time and specific equipment is required for this procedure making the burn-in expensive. Since burn-in does not lead to the identification of all suspicious devices some of them still remain which could become failures in the product lifetime. It would be desirable to discover such devices by screening methods as soon as possible and perhaps even shortens or skips the burn-in testing procedure. Commonly used screening methods are described in the following.

A. Part Average Testing (PAT)

PAT is mainly performed on tests which are classified as reliability relevant, e.g. IDDQ. In this case not only the devices outside the predefined specification limits but also the outliers based on the distribution of the measurement data should be rejected. To determine the PAT limits usually the standard deviation σ is calculated and a normal distribution of the measurement data is assumed. Common choices of the PAT limits are $\pm 6\sigma$ [8]. All devices outside these limits are rejected (see Fig. 2). Although PAT is a frequently used screening method in semiconductor industry, it has the disadvantage that it is a one-dimensional approach. Therefore, it is not possible to detect outliers which only become apparent when taking the measurement data of two or more tests into account.

Fig. 2. Example for specification limits and PAT limits.

B. Lonely Die Screening

It is assumed that devices being surrounded by defective devices are more likely to fail early compared to those which are in a “good neighborhood”. For this reason those devices will be scrapped. There are several rules, e.g. being surrounded

by a specified number of defective devices or being in a row with defective devices, to make a decision to reject or keep the devices.

III. DATA TRANSFORMATION WITH ICA

It is assumed that the measurement data result from a mixture of signals caused by the manufacturing process, noise signals (e.g. due to the measurement equipment) and signals indicating irregularities in the data. Therefore, screening is complicated and it would be preferable to investigate the signals independently. ICA is a method of multivariate statistics which is able to decompose signals received from a linear mixing process of stochastically independent signals into their source signals [9]. This separation can be performed neither knowing the mixing process nor the original signals. The ICA model can be written as

$$\vec{x} = A \vec{s}, \quad (1)$$

where $\vec{x} = (x_1, \dots, x_m)^T$ denotes the vector of the m mixed signals and $\vec{s} = (s_1, \dots, s_n)^T$ the vector of the n source signals or sources with $m \geq n$. The $m \times n$ matrix A is called mixing matrix. Due to the fact that only the mixed signals $\vec{x}$ are known, it is impossible to calculate $\vec{s}$ or A conventionally. To tackle this problem, one has to assume that $\vec{x}$ and $\vec{s}$ are random vectors. Hence, $\vec{x}$ will turn into a $m \times T$ matrix X and $\vec{s}$ into a $n \times T$ matrix S , where T denotes the sample size. Thus, the ICA model can be written as

$$\begin{pmatrix} x_{11} & \dots & x_{1T} \\ \vdots & & \vdots \\ x_{m1} & \dots & x_{mT} \end{pmatrix} = A \begin{pmatrix} s_{11} & \dots & s_{1T} \\ \vdots & & \vdots \\ s_{n1} & \dots & s_{nT} \end{pmatrix}, \quad (2)$$

respectively

$$X = A S. \quad (3)$$

ICA separates the mixed signals by maximizing the independence of the source signals $s_1, \dots, s_n$. This can be performed by maximizing the non-Gaussianity since linear compositions of independent components are more normally distributed than the components themselves [10]. This fact is based on the central limit theorem. Therefore, ICA is not able to decompose multiple Gaussian source signals. Consequently, they will be assigned into one source signal. For further information regarding the theory of ICA estimation see [9].

Relating to the measurement data, m equates to the number of performed analog tests per device and T equates to the number of measured devices. Accordingly, the matrix X contains the measurement data where x_{jk} denotes the measurement value of the j th test and k th device. Since ICA transforms the data, making the new, transformed tests as independent as possible, subsequent one-dimensional outlier detection on the transformed data S is much more promising than on the original measurement data X . Furthermore, the separation can reveal characteristic features and irregularities in the data. A pattern which is caused by the manufacturing process of the devices will be called process-related signal. A signal indicating irregularities, one is interested in, will be called risk-relevant signal.

IV. DETERMINING THE NUMBER OF INFORMATIVE SOURCES

During ICA it is possible to perform a dimensionality reduction. In this case $n < m$, i.e. the number of the transformed tests is less than the number of the original tests. This is useful, for example, if similar measurement data of many dependent tests should be investigated together. Even if the number of those tests is large, the number of informative source signals, i.e. process-related and risk-relevant signals, can be comparably small, because outliers may be seen in multiple tests or their combinations and process-related signals are expected to be present in each of the tests. If so, this information should be summarized and a dimensionality reduction is strictly recommended [9, chapter 13]. Because as soon as ICA has already detected the informative source signals and a Gaussian noise component, ICA tries to find further characteristics in the measurement data if it is forced to decompose the remaining signal components. As a consequence, ICA launches into tiny effects so that spurious characteristics will be indicated. Hence, nearly every device will be suspicious for an oversized number of computed signals. However, if the dimensionality is reduced too much, there will be a loss of information and probably not all informative sources will be found. In this case, not all suspicious devices can be detected.

As the number of informative sources varies with the number of outliers, it is not known beforehand and has to be determined from the data. This is a frequently discussed problem and suitable procedures are often tailor-made to the type of data under investigation. A conventional approach is choosing the minimum number of components that explain most of the expected variation [9, chapter 13]. But in many cases such components are not related to the risk-relevant signals [11]. For this reason, other approaches were analyzed [11] – [13]. Except of [11] these methods are considered for data from other areas (e.g. human speech signals) and have not yet been applied in the context of semiconductor-related data. Furthermore, ICA is often applied only for a small number of mixed signals. However, there can be hundreds of dependent tests for semiconductor devices.

A. A new Approach for Dimensionality Reduction

As already mentioned ICA searches for characteristics in data and filters out process-related signals and risk-relevant signals. The remaining signals should be Gaussian noise. This constitutes the basic idea of the new approach. According to that, it should be possible to decide based on the Gaussianity or non-Gaussianity of a source calculated by ICA without dimensionality reduction whether this source is an informative one or not. But due to the fact that more than one Gaussian noise signal will be summarized in one signal the problem of spurious characteristics mentioned above arises and all or nearly all calculated sources will be non-Gaussian.

Therefore, an iterative method is suggested where ICA is performed with dimensionality reduction for an increasing number of source signals to be computed. The sources are tested for Gaussianity after each ICA decomposition. As soon as one of the sources is Gaussian, the algorithm terminates and returns the computed components without the Gaussian signal. Therefore, only the informative source signals are returned.

The pseudo-code is shown in Algorithm 1. The computed sources will be used for subsequent outlier detection.

Algorithm 1 Calculation of the reduced source signals

```

1  for  $n = 1$  to  $m$  do
2      perform FastICA algorithm with
        dimensionality reduction to  $n$  components
3      for  $i = 1$  to  $n$  do
4          check component  $s_i$  for Gaussianity
5          if  $s_i$  is Gaussian then
6              delete  $s_i$  from the calculated sources
7              go to line 10
8          end if
9      end for
10 end for
11 return calculated source signals, without Gaussian signal

```

B. Probability of the Error of the second kind

For the test on Gaussianity the level of significance α can be specified to control the probability of the error of the first kind, i.e. the probability of an incorrect rejection of the null hypothesis H_0 given by

$$H_0: s_i \sim \mathcal{N}, \text{ i.e. source } i \text{ is Gaussian.} \quad (4)$$

In our investigations we choose $\alpha = 1\%$. Since holding the null hypothesis interrupts the proposed algorithm, the probability of the error of the second kind denoted by β , i.e. the probability of an erroneous holding of the null hypothesis, is really crucial. Therefore, this probability should be acceptably small. Unfortunately β cannot be calculated directly in general but can be obtained by simulation studies. As β depends on the sample size and the distribution with the corresponding parameter selection, random numbers that fulfill the properties of interest have to be generated in order to determine the frequency of the occurrence of the error of the second kind. Multiple executions permit an estimation of the probability of this error and its robustness. We performed such simulations for the most popular distributions with different parameters and for various sample sizes. In particular we simulated β for the data we used in section V. Furthermore the chosen test on Gaussianity has an impact on β . Therefore, we performed our simulations for various tests.

V. APPLICATION ON SEMICONDUCTOR DATA

For our investigations we only considered analog tests which are dependent and have similar measurement data. We investigated wafers separately in order to eliminate influences of wafer-to-wafer variations. All computations were made in MATLAB. The dimensionality reduction was performed by Algorithm 1 which was programmed by us using the FastICA algorithm of Aapo Hyvärinen, which can be downloaded from [14]. For the test on Gaussianity we used either the one-sample Kolmogorov-Smirnov test (KS-Test) or the Shapiro-Wilk test (SW-Test) provided by MATLAB. The choice depended on the data under investigation. In general, the probability β was smaller but the number of computed sources was higher for the SW-Test. If β was acceptably small for both tests, the KS-Test to receive the smallest possible number of sources was chosen.

A. Examination of IDDQ Data: Use Case I

In the first example, IDDQ (Direct Drain Quiescent Current) data from wafer level testing of a complex micro-controller were used. The devices that constituted the sample set passed the whole wafer level testing process. Therefore, they were classified as “good devices”. For the chosen product $m = 97$ wafer level IDDQ tests were performed and the wafer under investigation had $T = 1296$ passed devices that should be tested for suspicious ones. Hence the matrix X of the observations was a 97×1296 matrix.

As it is known from device physics that IDDQ data are log-normally distributed [15], we simulated β for log-normally distributed random numbers using the parameters μ , σ and their variations estimated from the IDDQ measurement data. Using the SW-Test β was between 0% and 0.4% and its median was zero for $T = 1200$ devices. Therefore, β was acceptably small. This leads to a high power $1 - \beta$ of the test, i.e. to a high probability that the test correctly rejects the null hypothesis assuming that the data are log-normally distributed. This was not the case for the KS-Test.

The proposed algorithm using the SW-Test detected 11 transformed tests or rather informative source signals which are shown in Fig. 3. One of these source signals is clearly different from the other ones. The visualization of this signal by means of the wafer map in Fig. 4 shows a pattern which should correspond to a process-related signal, since it can be seen less evident in each of the original IDDQ tests. If there is a process-related problem during manufacturing which causes a specific pattern on the wafer, this could also be filtered out by ICA. The other 10 source signals are either approximately Gaussian with one or only a few outliers or signals showing a weak structure only in their probability plots. The sources, which are nearly Gaussian, are expected to be signals indicating irregularities in the measurement data because there are no similarly distributed tests in the raw data. The probability plot for such a source can be seen in Fig. 5.

Fig. 3. All new, transformed tests or rather source signals.

To detect suspicious devices we performed subsequent one-dimensional outlier detection on each of the reduced and via ICA transformed source signals. The outlier detection was performed by the PAT method using a $\pm 8\sigma$ limit. Normally a $\pm 6\sigma$ limit is used [8] but the suspicious devices clearly differ from the other ones after ICA enabling a wider limit to be sufficient for screening the outliers in this study.

Fig. 4. Data of a transformed test or rather process-related signal. The value of each device is depicted by colors and the devices are arranged due to their location on the wafer. This source shows a distinctive pattern.

Fig. 5. Data of a transformed test or rather risk-relevant signal. The outlier corresponds to a suspicious device.

Since burn-in is an effective method to screen early failures, the pass-fail-information resulting from the post burn-in package level IDDQ testing was used to verify the identified outliers. Table I. shows the identified suspicious devices and their validation regarding the package level IDDQ pass-fail-information. As the table shows, 4 of 5 package level fails were detected correctly. Another 3 devices, characterized as suspicious ones, passed the package level testing process. Nevertheless those devices could fail within product lifetime with a higher probability. Compared to the new method, commonly used screening methods wouldn't have detected the 4 package level fails.

TABLE I. VALIDATION OF THE DETECTED SUSPICIOUS DEVICES

		pass-fail-information based on package-level IDDQ fails	
		pass	fail
classification via method	not suspicious	1120	1
	suspicious	3	4

B. Examination of IDDQ Data: Use Case II

In the second example, again IDDQ data from wafer level testing of a complex micro-controller but of a different product were examined. For this product $m = 200$ wafer level IDDQ tests were performed and the wafer under investigation had $T = 460$ passed devices.

Like the previous example we simulated the probability β for log-normally distributed random numbers using the

parameters estimated from the IDDQ measurement data. The boxplots in Fig. 6 depict the results of the simulations for various sample sizes using the KS-Test. As the maximum of β is 0.2% for $T = 450$ devices, β is acceptably small.

Fig. 6. Simulated β for log-normally distributed random numbers with $\mu \cong -10.6$ and $\sigma \cong 0.6$. For each sample size 500 times 500 samples were generated to estimate 500 β -values illustrated in the related boxplot.

The proposed algorithm using the KS-Test detected 17 informative source signals which are shown in Fig. 7. Again one signal (in red) is set apart from the other ones. The probability plot of this process-related signal is shown on the left hand side in Fig. 8. The other 16 sources are approximately Gaussian with one or only a few outliers. The probability plot for such a risk-relevant source can be seen on the right hand side in Fig. 8.

Fig. 7. All transformed tests or rather source signals. Process-related signal in red. Risk-relevant signals in blue.

Fig. 8. Probability plot of the process-related signal on the left hand side. Probability plot of a risk-relevant signal on the right hand side.

We performed subsequent one-dimensional outlier detection as described in the previous example. Again, the

pass-fail-information resulting from package level IDDQ testing was used to verify the identified outliers. Table II. shows that 4 package level fails were detected correctly. Moreover, these fails were the only package level IDDQ fails on the wafer under investigation. Furthermore, 8 devices were identified as suspicious ones but passed the package level testing.

TABLE II. VALIDATION OF THE DETECTED SUSPICIOUS DEVICES

classification via method		pass-fail-information based on package-level IDDQ fails	
		pass	fail
	not suspicious	444	0
	suspicious	8	4

C. Examination of IDDQ Data: Use Case III

In the third example, the potential of the proposed algorithm combined with one-dimensional outlier detection will be demonstrated by analyzing an early failure, returned from the customer, which passed the whole testing process. Analyzing the measurement data of tests which concern the reliability of the device, more precisely IDDQ data, showed that it would have been possible to detect the returned device using the proposed method. The corresponding wafer had $T = 989$ passed devices. For this product 15 IDDQ tests were performed.

The proposed algorithm using the KS-Test calculated 4 informative source signals, which are shown in Fig. 9. It can be seen that there are most likely 3 process-related signals (source 1, 2 and 4) and 1 risk-relevant signal (source 3). The mean of each source is close to 0, as the data have been centered when removing site effects from the raw data. Nevertheless, the mean of source 1, 2 and 4 is much bigger compared to the mean of source 3. The risk-relevant signal, which is shown in Fig. 10, is Gaussian with one outlier. This outlier is precisely the returned device. An outlier detection performing the screening only on the risk-relevant source would detect this early failure without rejecting further devices. Furthermore, examining the data for each IDDQ test individually or analyzing the ΔIDDQ values, which is a common approach for IDDQ failure analysis, would not identify the returning device. The probability plot of the ΔIDDQ values can be seen in Fig. 11. The returned device is highlighted in red.

Fig. 9. Boxplots of the calculated transformed tests or rather source signals.

Fig. 10. The risk-relevant signal whose outlier is the device returned by the customer.

Fig. 11. Probability plot of the maximal ΔIDDQ value of each device. The returned device is highlighted in red.

VI. CONCLUSION & OUTLOOK

We proposed a method which is able to reveal suspicious devices having a higher risk to fail early in the product lifetime based on the processing of several hundreds of similar, dependent tests. This is done by a data transformation called ICA and an embedded dimensionality reduction. The data transformation and reduction was combined with one-dimensional outlier detection in order to identify the suspicious devices. The presented method was applied on semiconductor measurement data, in particular IDDQ data, and verified by reasonable criteria. This method detects burn-in or package level failures earlier and thereby reduces the test costs. Furthermore, the method can be helpful in order to identify devices that are predestinated to be customer returns.

The method is not limited to IDDQ or rather log-normally distributed data. However, it must be noted that the power of the chosen test for data reduction has to be acceptably large to ensure a reasonable data transformation, reduction and thereby a successful subsequent screening of suspicious devices.

For future research it would be useful to work on an automatic grouping of the sources to separate the risk-relevant signals and thereby reduce the overscreening. Alternatively, one can use screening methods taking the distribution of each source signal into account. Another focus of future research could be the identification of groups of similar, dependent test for which it is beneficial to apply the presented method.

ACKNOWLEDGMENT

The authors would like to thank Dr. Stefan Trost who provided expertise in the field of IDDQ testing that greatly assisted this research. We are also thankful to Dr. Torsten Obier for his general support for this work at Infineon Technologies Dresden.

REFERENCES

- [1] "AEC Documents". Available: <http://www.aecouncil.com/AECDocuments.html>
- [2] W. R. Daasch, K. Cota, and J. McNamers, "Neighbor selection for variance reduction in IDDQ and other parametric data," in Proceedings of the International Test Conference, pp. 92–100, IEEE, 2001.
- [3] R. Turakhia, B. Benware, R. Madge, T. Shannon, and R. Daasch, "Defect screening using independent component analysis on IDDQ," in Proceedings of the 23rd VLSI Test Symposium, pp. 427–432, IEEE, 2005.
- [4] L. Rao, M. L. Bushnell, and V. D. Agrawal, "Graphical IDDQ signatures reduce defect level and yield loss," Transactions on Very Large Scale Integration (VLSI) Systems, vol. 15, no. 11, pp. 1245–1255, IEEE, 2007.
- [5] A. Zernig, O. Bluder, J. Pilz, and A. Kästner, "Device level maverick screening: detection of risk devices through independent component analysis," in Proceedings of the Winter Simulation Conference, pp. 2661–2670, IEEE, 2014.
- [6] A. Zernig, O. Bluder, J. Pilz, A. Kästner, and A. Krauth, "Identification of risk devices using independent component analysis for semiconductor measurement data," International Journal of Industrial Engineering, vol. 23, no. 5, 2016.
- [7] J. Lienig and H. Brümmer, *Fundamentals of Electronic Systems Design*, Springer, 2017.
- [8] "AEC_Q001_Rev_D," 12/09/2011. Available: <http://www.aecouncil.com/AECDocuments.html>
- [9] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.
- [10] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," Neural Networks, vol. 13, no. 4-5, pp. 411–430, 2000.
- [11] P. M. O'Neill, "Production multivariate outlier detection using principal components," International Test Conference, pp. 1-10, IEEE, 2008.
- [12] H. Sawada, R. Mukai, S. Araki, and S. Makino, "Estimating the number of sources using independent component analysis," Acoustical Science and Technology, vol. 26, no. 5, pp. 450–452, 2005.
- [13] A. Kassouf, D. J. R. Bouveresse, and D. N. Rutledge, "Determination of the optimal number of components in independent components analysis," Talanta, 179, pp. 538-545, 2018.
- [14] A. Hyvärinen, "FastICA 2.5," 10/19/2005. Available: <http://research.ics.aalto.fi/ica/fastica/code/dlcode.shtml>
- [15] S. Sabade and D. Walker, "Improved wafer-level spatial analysis for IDDQ setting," in Proceedings of the International Test Conference, pp. 82–91, IEEE, 2001.