

國立交通大學

統計學研究所

碩士論文

考量半導體製程能力限制下

之晶圓圖隨機性辨識法及應用

Automatic Detection of Patterned Wafer Sort Maps
with Process Baseline

研究生：莊銘弘

指導教授：洪志真 博士

涂凱文 博士

中華民國九十九年六月

考量半導體製程能力限制下
之晶圓圖隨機性辨識法及應用

Automatic Detection of Patterned Wafer Sort Maps
with Process Baseline

研究生：莊銘弘

Student : Ming-Hong Chuang

指導教授：洪志真 博士

Advisor : Dr. Jyh-Jen Horng Shiau

涂凱文 博士

Dr. Kai-Wen Tu

國立交通大學
統計學研究所
碩士論文

A Thesis

Submitted to Institute of Statistics

College of Science

National Chiao Tung University

In Partial Fulfillment of the Requirements

for the Degree of Master

in

Statistics

June 2010

Hsinchu, Taiwan

中華民國九十九年六月

考量半導體製程能力限制下 之晶圓圖隨機性辨識法及應用

學生：莊銘弘

指導教授：洪志真 博士

涂凱文 博士

國立交通大學統計學研究所碩士班



半導體產業是一種高投資產業，每間公司在製程上所投資的成本都很高。由於半導體產業製程相當複雜，因此，對每間公司而言，如何使製程穩定與良率提升便是一個很重要的目標。

晶圓圖是半導體產業中偵測製程異常的重要參考依據。當非隨機性的異常晶圓圖發生時，常代表製造過程發生異常。藉由這些異常晶圓圖的圖形，也可幫助工程師找出可能發生的原因，例如機台發生問題或哪個製程步驟有異常等。對於分辨隨機與非隨機晶圓圖的議題已有很多相關研究，分別提出一些可以取代人工目視判斷方式的方法，以減少由於人為主觀的因素所導致的圖形判斷結果不一致。

然而在半導體廠製造技術水準與機台能力或產品特性等限制下，常導致晶圓圖上某些區域容易產生故障品，此種原因所形成的異常晶圓圖形並非是特殊的製程異常所引起，而是基於半導體廠製造技術水準與機台能力或產品特性的根本限制。此種「異常」和真正有歸屬原因的「異常」，一般電腦自動辨識方法是分不出來的。在本篇論文中，在分辨晶圓圖形是否隨機，我們針對此種製程能力的限制提出一個修正方法，期能正確判定真正的晶圓圖異常。此方法將可協助工程師快速掌握製程異常的發生並加以排除，使製程穩定並提升良率。

Automatic Detection of Patterned Wafer Sort Maps with Process Baseline

Student: Ming-Hong Chuang

Advisor: Dr. Jyh-Jen Horng Shiau

Dr. Kei-Wen Tu

Institute of Statistics

National Chiao Tung University



Abstract

The semiconductor industry is a very competitive and high-investment industry. The manufacturing process is one of the areas that are heavily invested. Semiconductor manufacturing are so complex that improving the process stability and yield is essential for each company to stay competitive.

The wafer sort map (WSM) is a useful tool for detecting abnormal processes. Non-random patterns on a WSM usually provides clues about process problems. With particular patterns, WSMs can help engineers to identify possible causes, such as problems in equipments or process steps. Some automatic methods were proposed to distinguish between random and non-random WSMs in the literature, attempting to replace the labor-intensive human recognition operations for cost saving as well as to reduce the inconsistency due to subjective human judgements.

In practice, however, almost all WSMs exhibit some regions of inherent failures, which generally are due to process limitations caused by, say, layouts, equipments, process technologies, etc. This kind of failures are inevitable at the present technology level; thus, in the view of statistical process control, they should be considered as caused by common causes instead of process problems caused by special causes. Thus, in fab, such failures are often accepted and referred to as the “baseline” of the process. Unfortunately, without accounting for, the baseline, most of the existing automatic methods will classify these baseline patterns as non-random patterns. In

this thesis, we propose a new approach to detecting WSMs with “genuine” non-random WSM by taking the baseline into account. Three effective schemes are developed and studied. Effective detection of patterned WSMs can help engineers to trace process failures of particular patterns back to their root causes and then improve the process stability and yield by eliminating these causes.



誌 謝

兩年的研究生生涯隨著這篇論文的結束即將告一段落，在這兩年裡，讓我成長了許多，獲得很多的經驗等等。而在這段日子裡，我要感謝我的指導教授洪志真教授與涂凱文學長。在涂凱文學長的指導下，讓我清楚知道統計應用的價值，也了解學術與實務上的差異，讓我在面對事情時多了一份不一樣思維。此外，也感謝洪志真老師的幫忙，在與老師討論的過程中，讓我學習到了做研究應有的態度與細心。再來要感謝交大的每一位老師，在求學過程中給予我很多的幫助，此外，也要特別感謝所上的郭小姐，除了常聽我們發牢騷外，還常常給予了我們許多的幫助，你就像是我們統研所的保母，真的很謝謝你。除此之外，還要感謝跟我一起度過這兩年生活的統研所同學們，能和樂的相處做學問，晚上的宵夜聊天時間，共同奮鬥與生活，讓我每一天都可以過得很開心。最後要感謝我的家人與女友，你們真的辛苦了，也因為你們的支持讓我能度過這壓力大的碩二生活，也能讓我專心的在學業上發展，謝謝你們。有了大家的支持，在即將踏出社會的我，一定會努力的，絕不會辜負大家對我的期望。

莊銘弘 謹誌于
國立交通大學統計學研究所
中華民國九十九年六月

Contents

1	Introduction	1
1.1	Background	1
1.2	Motivation	4
1.3	Overview	5
2	Introduction of Semiconductor Process and Packaging Technologies	7
2.1	Front-End Process Steps	7
2.2	Back-End Process	9
2.2.1	Wafer Sort	10
2.2.2	Assembly	10
2.2.3	Final Test	11
2.3	Advanced Packaging Technologies — System-In-Packaging	12
2.4	Wafer Sort Map Pattern	13
3	Literature Review	18
3.1	Related Studies	18
3.2	Spatial Randomness Tests	19
3.2.1	Log Odds Ratio Test	20
3.2.2	HNF Test	21
3.2.3	HNF Method Under Cluster	24
3.3	Mode Estimation Via Kernel Density Estimation	26
4	Proposed Schemes	27
4.1	Data Encoding	27
4.2	Proposed Method	27
4.2.1	Mode Odds Ratio Method	29
4.2.2	Mode HNF Method	30
4.2.3	Empirical HNF Method	31
5	Simulation Studies	34
5.1	Simulation Settings	34
5.2	Generating Patterned Wafer Sort Maps	35
5.3	Comparisons	36

5.3.1	False-alarm Rate	37
5.3.2	Detecting Power vs. Yield	38
5.3.3	Detecting Power vs. Patterned Area	39
6	Conclusion	42
	References	44
A	Table and WSMs Related to Simulation Studies	48



List of Tables

1	An example for wafer creation.	35
2	Counts of false alarm's of the two existing methods and three proposed methods among 150 random WSMs under various yield levels.	48
3	Counts of detection among 150 WSMs with the linear scratch pattern under various yield.	50
4	Counts of detection among 150 WSMs with the ring pattern under various yield.	52
5	Counts of detection among 150 WSMs with the bottom pattern under various yield.	54
6	Counts of detection among 150 WSMs with the moon pattern under various yield.	56
7	Counts of detection among 150 WSMs with the linear scratch pattern under various patterned areas.	58
8	Counts of detection among 150 WSMs with the ring pattern under various patterned areas.	60
9	Counts of detection among 150 WSMs with the bottom pattern under various defect areas.	62
10	Counts of detection among 150 WSMs with the moon pattern under various patterned areas.	64
11	Counts of detection among 150 WSMs with the linear scratch pattern under various patterned areas.	66
12	Counts of detection among 150 WSMs with the ring pattern under various patterned areas.	68
13	Counts of detection among 150 WSMs with the bottom pattern under various patterned areas.	70
14	Counts of detection among 150 WSMs with the moon pattern under various patterned areas.	72

List of Figures

1	System predicted yield as a function of the number of dies assembled for various die yields.	2
2	The left panel is a WSM with the code and the right panel is the corresponding visual figure of bad dies.	3
3	Some random and patterned wafer sort maps with the baseline limitation.	5
4	Flow chart for semiconductor manufacturing.	7
5	Flow chart for generic IC process sequence.	8
6	Mark the bad dies on the wafer after wafer sort.	10
7	IC assembly process sequence.	10
8	Final test process sequence.	11
9	Scope of System-in-Packaging.	12
10	The graph of SiP structure viewed by a scanning electron microscope.	13
11	Random defects of various defect level.	14
12	Examples of systematic defects.	14
13	Random defects plus systematic defects to form a mixed defects.	15
14	Bottom pattern	15
15	Ring pattern	16
16	Linear scratch pattern	16
17	Crescent moon pattern	17
18	The king-move neighborhood.	20
19	An example to depict the calculation of the four joint-count statistics.	21
20	An example to depict the calculation of the pair \mathcal{T} statistics.	22
21	Data encoding.	27
22	Generating a sequence of pseudo-inside WSMs from a WSM.	29
23	Generating a sequence of pseudo-outside WSMs from a WSM.	29
24	Mode estimation of Z_j values	30
25	Mode estimation of D_j values	31
26	The flow chart of the proposed schemes.	33
27	Simulation parameters for wafer.	34
28	Illustration for generating a linear scratch pattern.	35
29	Baseline setting.	36

30	4 different patterns with a central-disc baseline region.	36
31	False-alarm rate of the two existing method and three proposed methods for various yield level.	37
32	Detecting power for various patterns under a fixed patterned area size. . .	38
33	Detecting power for various patterns under middle yield level.	40
34	Detecting power for various patterns under high yield level.	41



1 Introduction

1.1 Background

The semiconductor industry has emerged as one of the most important industries in many countries such as the USA, Germany, South Korea, Japan, and Taiwan. Semiconductor devices are absolutely essential for almost all electronic products and systems in the sense that most of the electronic products and systems cannot be produced or operated without them. Their influences over human society are enormous. In recent years, the mainstream evolution of the electronic products, such as computers, cell phones, digital cameras/camcorders, and portable audio/video players, has continually striving people toward designing/manufacturing products to be faster in operation, smaller in size, lighter in weight, and of more value-added functionalities. Therefore, how to cope these desirable features has become major challenges in ensuing competitiveness in the semiconductor industry. Needless to say, advanced semiconductor packaging technologies play a very important role in such efforts.

One major goal in advanced semiconductor packaging technologies is to increase the density of devices in a fixed packaging size. To achieve this, the multi-chip module (MCM) was developed and a widely application of the MCM is system-in-package (SiP), which consists of multiple dies stacked vertically and connected within package. Each of these dies, such as a specialized processor, DRAM, or FLASH memory, usually has one single functionality. Dies are then combined with passive components to form a system or subsystem. Various dies can be manufactured separately in different semiconductor manufacturing companies, and then assembled together. However, the cost, as well as the quality and production yield, highly depends on the cost and the quality of the individual dies. Illustrating this with a simplified case that ignores assembly defects and different yield levels of parts in an assembly, Figure 1 shows that the predicted yield of an assembly decreases exponentially as a function of the number of devices used. A SiP consists of various types of dies packaged together in an IC. Any of these dies failed will cause the system not working properly. So, to drive up the yield of a SiP is more difficult than a traditional single-chip IC.

In the final product, if one die fails, the whole module is typically useless or so reduced in performance that it cannot be sold for its intended purpose or at a price that can cover

the cost. Thus, there is no doubt that the high quality of dies is playing an important role for an MCM in the aspect of reducing the cost and enhancing the quality/yield. This leads to the demands for high-quality dies, which are called “known good dies (KGD)” in the die market. A KGD is a bare die (i.e., without package) with high quality and reliability that can assure the functionality in the integrated circuit (IC) level. Because of that, KGD has become an imperative for component manufacturers who provide various types of dies for the multi-chip module business.

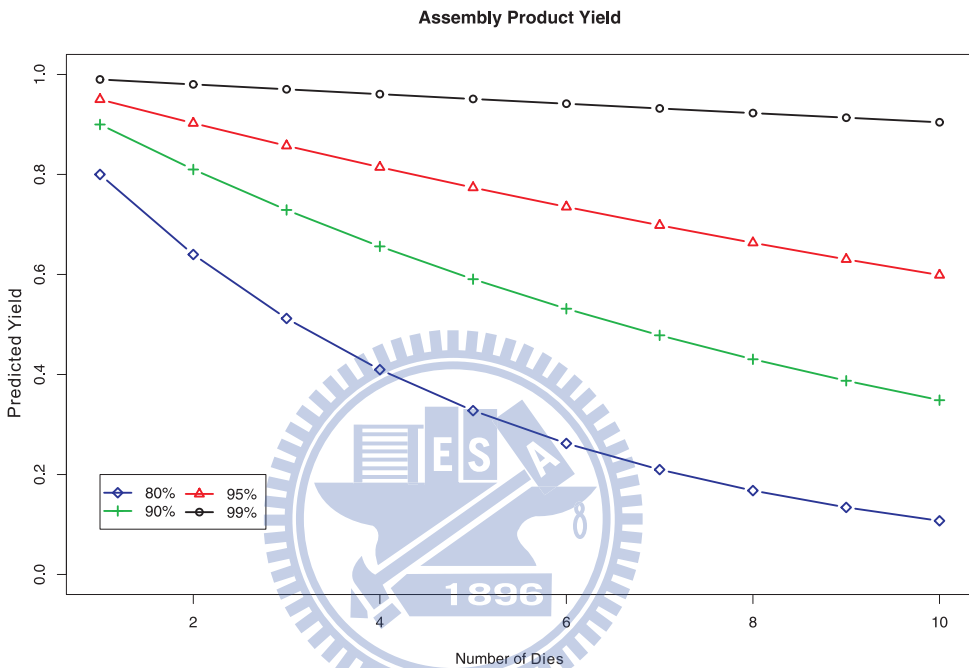


Figure 1: System predicted yield as a function of the number of dies assembled for various die yields.

However, the semiconductor manufacturing process is more complicate than those of traditional manufacturing industries. It takes about 30–60 days to complete the process of making bare silicon wafers into integrated circuits, such as microprocessors or memory chips. In general, several wafers are processed simultaneously as a “lot”, typically of 25 wafers, and the size of each wafer ranges from 3 to 12 inches in diameter. Each wafer would contain thousands of dies depending on the size of the dies being produced.

After wafer fabrication, all dies on a wafer must go through a process called the wafer sort (WS). The purpose of the wafer sort is to examine the electrical functions of each die and to prevent from packaging bad dies into a multi-chip package. The results of the wafer sort give each die a binary code that denotes it as either a good die (0) or bad

die (1). These wafer sort data then are used to generate the wafer sort map (WSM). A WSM displays the locations of bad dies on the wafer. Figure 2 presents an example of WSMs. The white squares and black squares on the map denote the good dies and bad dies, respectively.

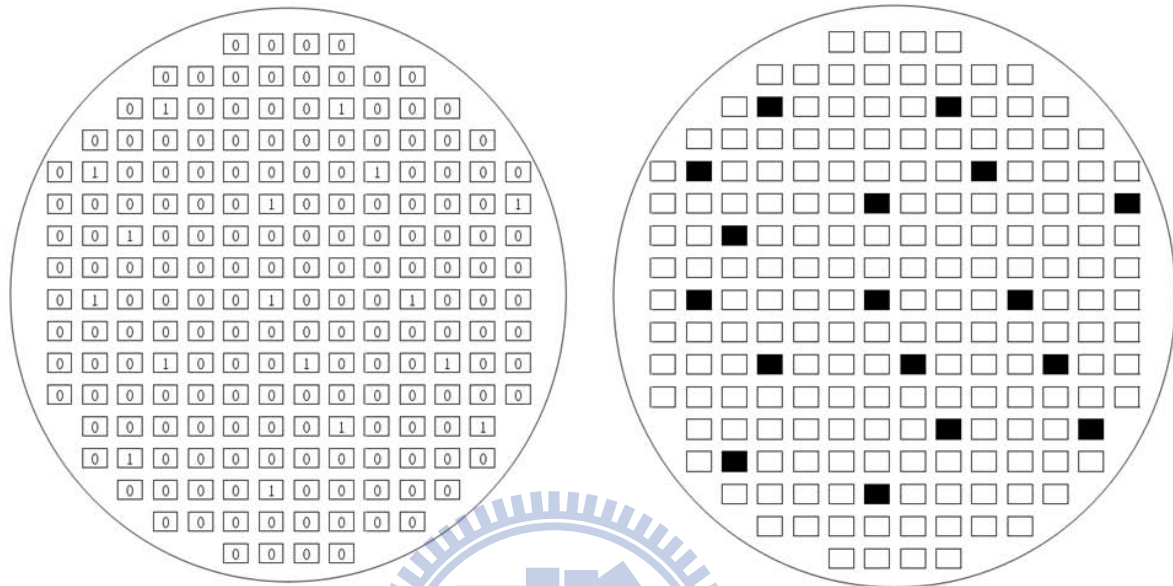


Figure 2: The left panel is a WSM with the code and the right panel is the corresponding visual figure of bad dies.

Since the occurrences of any quality excursions can induce WS failures, WSM analysis can help determine the possible causes of failures and help devise solutions to prevent such failures from reoccurring. For example, uneven temperatures or chemical aging often lead to spatial clusters on the WSM. Clustering also can be the result of crystalline nonuniformity, photo-mask misalignment, or particles caused by mechanical vibration. Improper material shipping and handling also can leave scratch lines on a WSM (see, for example, Cunningham and McKinnon [1], Hansen and Tyregod [2], Hansen *et al.* [3], and Taam and Hamada [4]).

Because WSMs contain important information that might guides quality engineers to trace back to the source of process failures, WSMs has been considered as one of the most important analysis tools in the semiconductor industry.

On the other hand, more important issues for KGD vendors to consider are (i) how to provide quality assurance to their customers and (ii) what sale strategies to take to make more profits. One possible solutions for that is to examine WSMs. If a WSM is of

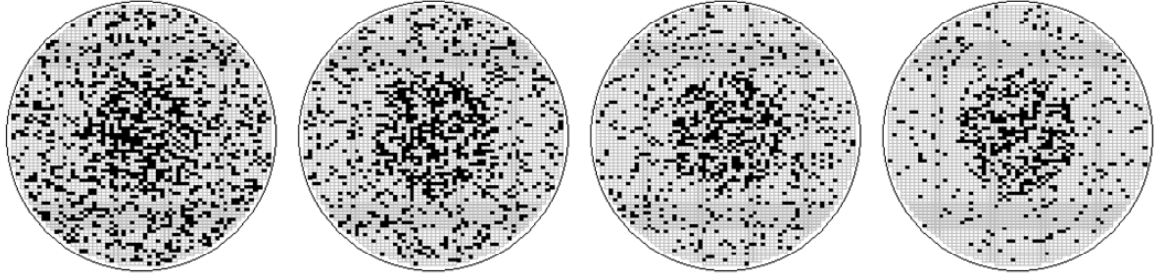
some patterns, it implies process problems exist. Furthermore, the general market has a common view that the good dies on a wafer with no patterns have better quality than the good dies on a wafer with patterns in the semiconductor manufacturing industry. To determine whether a wafer is of some patterns or not, spatial randomness tests are a useful tool. Some tests are available in the literature, see, for example, Taam and Hamada [4] and Hansen *et al.* [3].

Ideally, a completely random mechanism for dies can be defined as “bad dies on the wafer are randomly distributed under the spatial homogeneous Bernoulli process (SHBP).” This means the probability of each die being bad is the same. In practice, however, almost all WSMs exhibit some regions of inherent failures, which generally are due to process limitations caused by, say, layouts, equipments, process technologies, etc. These types of failures are inevitable at the present technology level; thus, in the view of statistical process control, they should be considered as caused by common causes instead of process problems caused by special causes. Thus, in fab, they are often accepted and referred to as the “baseline” of the process. Simply speaking, “common causes” induce usual, historical, and quantifiable variations in a system, whereas “special causes” lead to unusual, not previously observed, and non-quantifiable variations. Figure 3 presents some WSM examples: the top panel shows four “random” wafer sort maps with a baseline of a central disc pattern and the bottom panel exhibits four WSMs with patterns such as linear scratch, ring, bottom, and crescent moon patterns in addition to a central disc baseline.

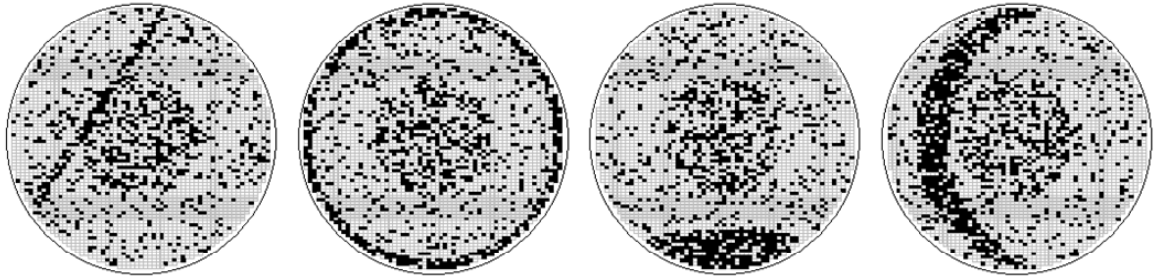
However, the currently available spatial randomness tests can not distinguish between the wafer patterns induced by common causes or special causes. Because of this reason, semiconductor companies usually hire engineers to classify these WSMs visually based on their own experiences. This approach is subjective as well as time-consuming. It would be desirable to develop an automatic detecting procedure for classifying the wafer patterns induced by common causes or special causes and this is the main objective of this study.

1.2 Motivation

As the technology evolves and advances, KGDs start playing an important role in advanced packaging technologies such as SiP. Because the yield of SiP devices highly depends on the internal dies, KGDs are widely used to obtain higher yield of SiP devices.



(a) Random wafer sort maps with a baseline of a central disc.



(b) Patterned wafer sort maps with a baseline of a central disc.

Figure 3: Some random and patterned wafer sort maps with the baseline limitation.

Furthermore, for KGD suppliers, one of the quality assurance is from inspecting WSMs. In general, KGDs on WSMs with common-cause patterns are better than those on WSMs with special-cause patterns. So, KGD suppliers do classify wafers into random (i.e., better grade) wafers or patterned (i.e., lower grade) wafers after wafers pass through the wafer sort.

However, to classify wafers, many companies still rely on the visual inspection of experienced engineers. The manual sorting is not only time-consuming but also full of inconsistency, restricted by the capability of human recognition. By providing a randomness test that takes the baseline of the process into account, this study will help fab engineers to automatically determine whether a WSM is random or not. For simplicity, in this paper, we assume the baseline regions are pre-determined by engineers based on their experiences.

1.3 Overview

The remainder of this thesis is organized as follows. In Section 2, we give a brief overview on the semiconductor manufacturing process and advance packaging technologies. Moreover, some WSM patterns that often occur in the process with their probable causes

are reviewed. In Section 3, we give a literature review including some related studies in semiconductor manufacturing, spatial randomness tests, and a tool we use for mode estimation. In Section 4, we present the proposed schemes, which extend three existing spatial randomness tests to cope with baseline regions. Our proposed schemes can easily adapt to any kind of baselines set by engineers. In Section 5, we evaluate and compare the three proposed schemes via simulation studies. Finally, we conclude the thesis in Section 6.



2 Introduction of Semiconductor Process and Packaging Technologies

The processes involved in semiconductor manufacturing have been grouped by many researchers into four stages, which are termed wafer fabrication, wafer sorting, device assembly, and final testing. Wafer fabrication is an extremely sophisticated and complex process that manufactures silicon dies. Assembly is a highly precise and automated process that packages silicon dies for protection and providing users a practical way of handling. For detailed descriptions of semiconductor manufacturing processes, see, for example, Uzsoy *et al.* [5] and Knutson *et al.* [6]. The following descriptions are taken from May and Spanos [7].

The four stages of semiconductor manufacturing system are carried out separately in different work centers. These four stages are further grouped into two categories: front-end manufacturing operations consist of wafer fabrication and back-end manufacturing operations include wafer sorting, device assembly, and final testing (see Figure 4). We will discuss these two categories in the following subsections.

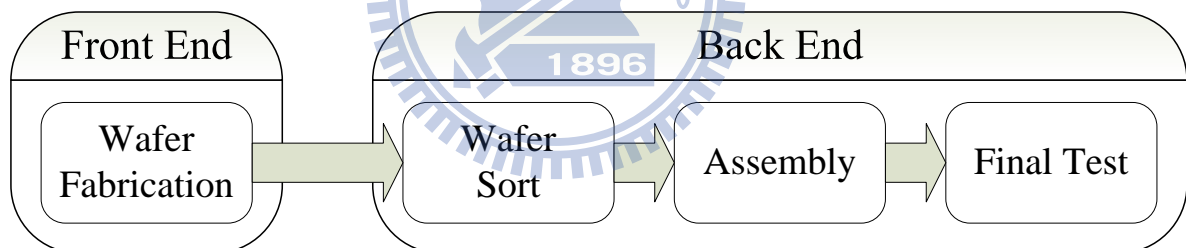


Figure 4: Flow chart for semiconductor manufacturing.

2.1 Front-End Process Steps

Semiconductor manufacturing consists of a series of sequential process steps including crystal preparing, wafer preparation, lithograph, etching, ion implantation, metallization, and cleaning to produce ICs. Figure 5 illustrates the interrelationship between the major process steps in IC fabrication. The main steps are summarized as follows.

- Crystal Preparing

The first stage of the semiconductor manufacturing process involves growing a single

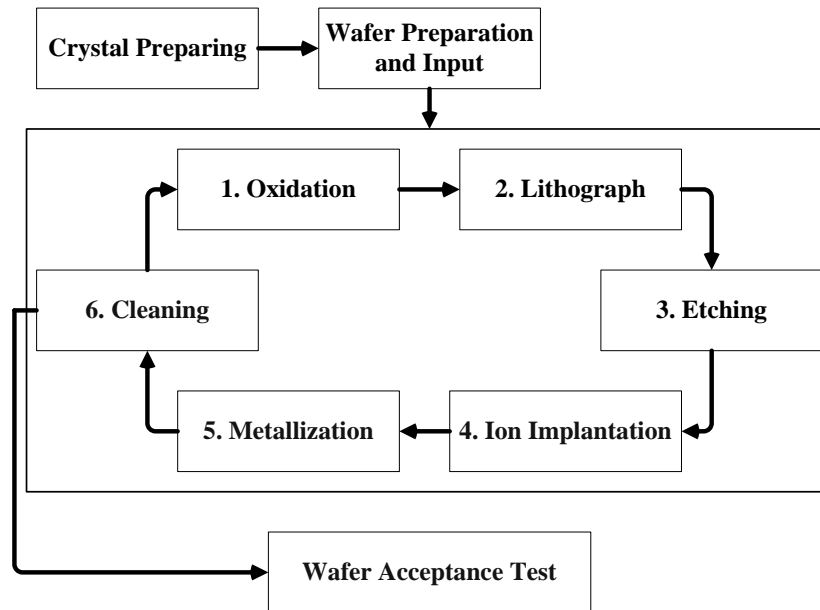


Figure 5: Flow chart for generic IC process sequence.

crystal of silicon into a solid cylindrical shape. The silicon is first purified and heated into a bath of molten liquid, into which a small crystal of silicon, called a seed, is dipped. As the seed is slowly withdrawn, the surface tension between the seed and the molten silicon causes some liquid to rise with the seed. The liquid solidifies around the seed to form a single crystal ingot.

- Wafer Preparation

The solid cylinder of silicon, typically 150-mm (6 in.), 200-mm (8 in.), and 300-mm (12 in.) in diameter, is then cut with a diamond saw blade to wafers about 0.5–0.7 mm thick. Each wafer is cleaned, smoothed, and polished through a series of machines.

- Oxidation

Generally, the step in semiconductor device fabrication involves the oxidation of the wafer surface in order to grow a thin layer of silicon dioxide (SiO_2). This oxide is used to provide insulating and passivation layers.

- Lithograph

The lithography is the process for pattern definition by applying a thin uniform layer of viscous liquid (photo-resist) on the wafer surface. The photo-resist is hardened

by baking and then selectively removed by projecting light through a reticle that contains mask information.

- Etching

This step selectively removes unwanted material from the surface of the wafer. The pattern of the photo-resist is transferred to the wafer by means of etching agents.

- Ion Implantation

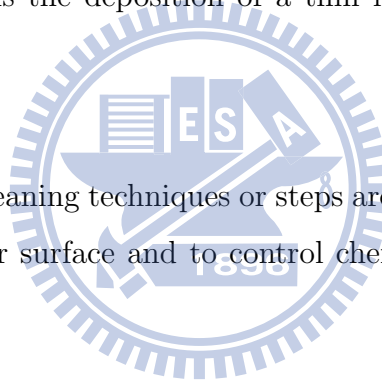
Ion implantation is a material-engineering process, by which ions of a material can be implanted into another solid, thereby changing the physical properties of the solid. Ion implantation is used in semiconductor device fabrication and in metal finishing, as well as in various applications of materials science research.

- Metallization

Wafer metallization is the deposition of a thin film of conductive metal onto the wafer surface.

- Cleaning

A number of wafer cleaning techniques or steps are employed to remove the contaminants from the wafer surface and to control chemically grown oxide on the wafer surface.



At the end of these process steps, several test sites located on the fixed locations of each wafer are selected to perform the wafer acceptance test (WAT) with 100-500 electrical test items sequentially. The objective of the WAT is to perform the device characteristic analysis.

2.2 Back-End Process

After the wafer fabrication is completed, the wafers go through the back-end process, which includes wafer sort, assembly, and final test. During the back-end process, the good dies are put through a series of processes to create the electrical connections necessary for device to function. The following is a brief description of the three back-end stages.

2.2.1 Wafer Sort

After wafer fabrication is completed, each finished wafer undergoes the wafer sort process. The wafer sort test with 50–100 test items is performed sequentially to each die. The objective of WS is to perform the die functionality sorting. Then, the bad dies are automatically marked with a black dot so they can be separated from the good dies after the wafer is cut (see Figure 6).

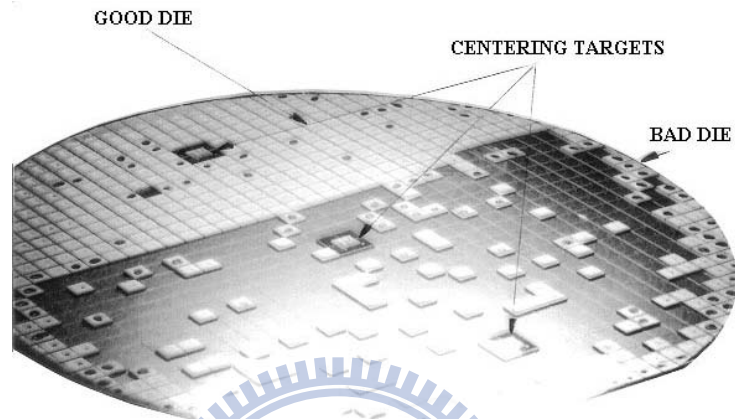


Figure 6: Mark the bad dies on the wafer after wafer sort.

2.2.2 Assembly

Assembly is the process after wafer sort that enables dies to be packaged for system use. Each good die which goes through the process is usually called an integrated circuit. The following are the main steps of the assembly process (see Figure 7).

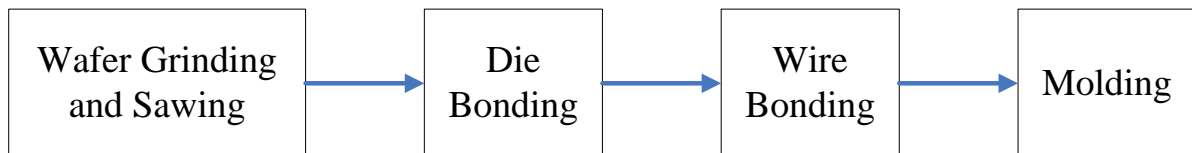


Figure 7: IC assembly process sequence.

- Wafer Grinding and Sawing

Wafer Grinding and Sawing is the process to saw the wafer to dies.

- Die Bonding

Die Bonding is the process of attaching the die either to its package or to some substrate such as tape carrier for tape automated bonding. The die is first picked from a separated wafer or from a waffle tray, aligned to a target pad on the carrier or substrate, and then permanently attached, usually by means of a solder or epoxy bond.

- Wire Bonding

Wire Bonding is the process of providing electrical connections between the die and the external leads of the semiconductor device using very fine bonding wires.

- Molding

Molding is the process of encapsulating the device in plastic material. Transfer molding is one of the most widely used molding processes in the semiconductor industry because of its capability to mold small parts with complex features.

2.2.3 Final Test

In this stage, final test is the final arbiter of process quality and yield at the completion of manufacturing. A description is in the following (see Figure 8).



Figure 8: Final test process sequence.

- Final Test

During the packaging process, dies may be damaged or packaging may not be correctly performed. Final test is a 100% test performed on each packaged IC prior to shipment to insure that any ICs improperly packaged are not shipped. The purpose of final testing is to ensure that the product performs to the specifications it was designed for.

- Appearance Check

Appearance check is to examine the surface for defects.

- Packing

Packing is to stamp QC label and seal the product in vacuum bags.

2.3 Advanced Packaging Technologies — System-In-Packaging

As the semiconductor technology advances, the process technology improves from .35um to 90nm and further to 45nm. Thus, in recent year, the electronics industry has experienced a great enhancement in development of new materials and processes to support the demand for “smaller, lighter, faster, and better” products. At the same time, it is seen that packaging technologies are also shifted toward achieving these requirements through the use of high-density packaging. One of these high-density packaging, system-in-packaging, is an ideal solution in markets to support the demand. By stacking multiple silicon dies vertically inside the same package, the technology effectively increases the functionality and capacity of the electronic devices. (See Figures 9 and 10.)

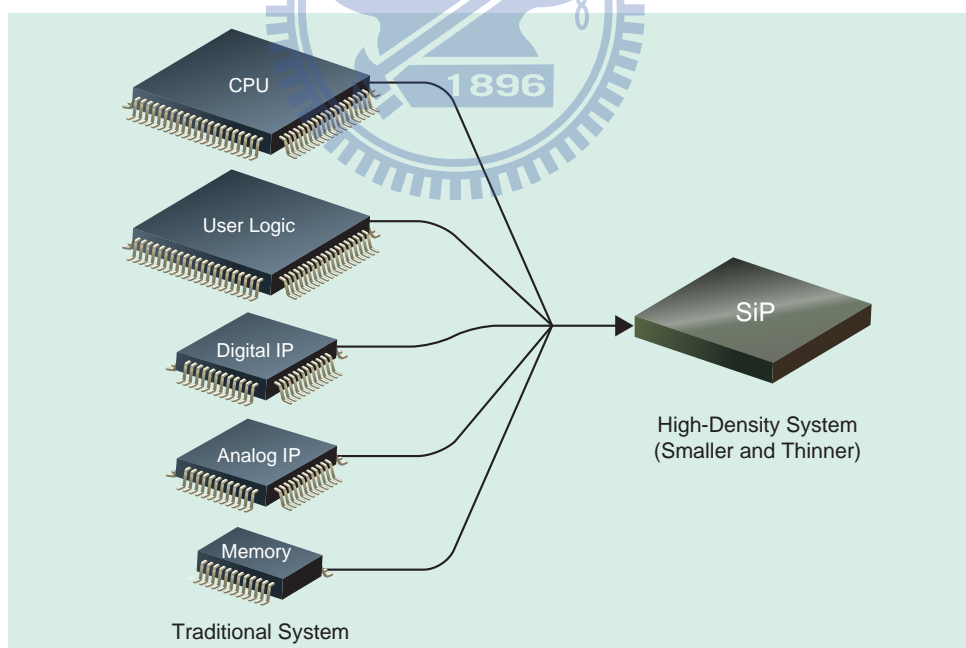
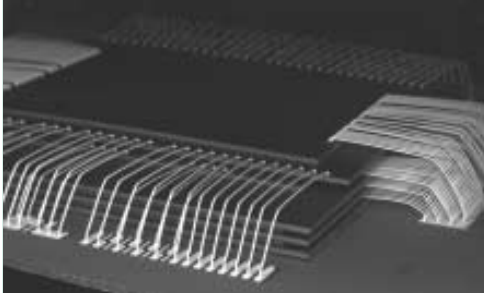
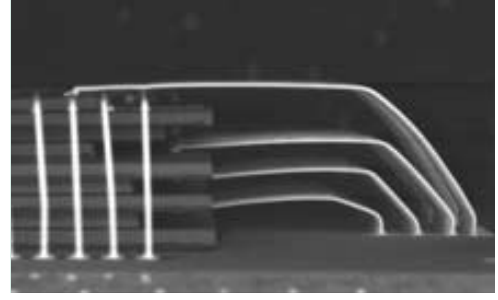


Figure 9: Scope of System-in-Packaging.

There are several reasons why the market demand seems to be growing strongly for SiP solutions. These include:



(a) 9-layer 5 dies stacked (Front View)



(b) 9-layer 5 dies stacked (Lateral View)

Figure 10: The graph of SiP structure viewed by a scanning electron microscope.

1. Size: The size of sub-system can be reduced by integrating multiple dies and other components in a SiP.
2. Performance advantages: Power reduction by minimizing line lengths (capacitive loads) between dies in a SiP.
3. Lower System Cost: An optimized SiP solution usually results in an overall system cost reduction compared to discrete IC packages.
4. SiP solutions reduce the complexity of the motherboard by moving the routing complexity to the package substrate. This often results in a reduced layer count in the motherboard and simplifies the product design.

More advantages were discussed in Scanlan and Karim [8], Lin [9], Buck [10], and Aguirre [11].

2.4 Wafer Sort Map Pattern

As mentioned early, the wafer sort is a process that examines the functionalities of each die by specific test conditions. One can color each die black (fail) or white (pass) from the test results for a wafer, and the resulting map is called a “wafer sort map”. By incorporating various plotting features into these maps, spatial patterns for passing and failing dies become readily apparent. Since the occurrence of any quality inferiority can usually be attributed to some specific causes, WSM analysis can help determine the possible causes of process failures and help devise solutions to prevent the reoccurrence of these failures.

Usually the failure patterns of WSM can be classified into three major categories as follows: (Kaempf [12])

1. Random patterns:

No spatial clustering and pattern exist, and the defective dies randomly distributed in the two-dimensional map. Figure 11 is an example. Random defects are usually caused by manufacturing environment factors. Even in a near sterile environment, particles cannot be removed completely. Nevertheless, reducing the level of random defects can improve the overall productivity of wafer fabrication.

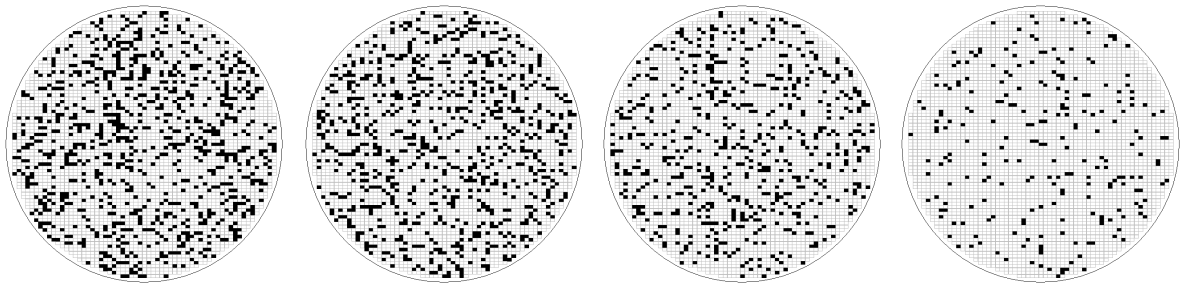


Figure 11: Random defects of various defect level.

2. Systematic patterns:

As examples, Figure 12 shows some patterns of systematic defects. The positions of defective dies in the wafer show the spatial correlation. Therefore, one may be able to trace back to the assignable cause from the problematic process steps or mechanism by analyzing the spatial distribution of failed dies. Systematic defects usually give an analyst some clues to find problematic steps and ways to eliminate them.

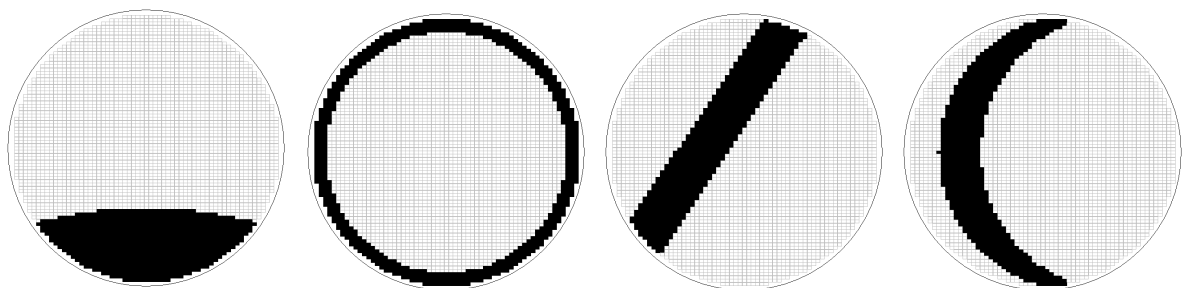


Figure 12: Examples of systematic defects.

3. Mixed pattern:

A mixed pattern is most common for a WSM, which consists of random defects and

systematic defects in one map. Figure 13 gives an example.

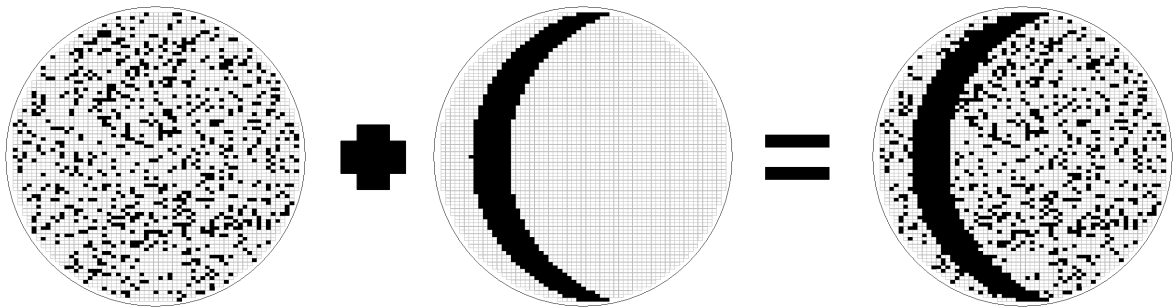


Figure 13: Random defects plus systematic defects to form a mixed defects.

The following are four patterns occurring most often and some potential causes associated with them. A baseline of a central disc is included as an example in the illustrative figures.

- Bottom

The bottom pattern (Figure 14) could be the result of uneven heating during a diffusion process or the probe card itself failed.

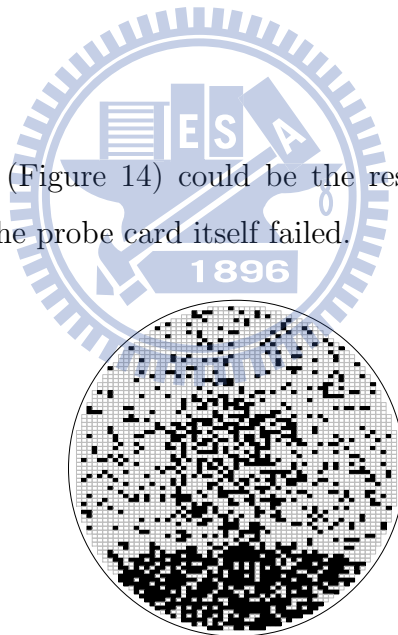


Figure 14: Bottom pattern

- Ring

The ring pattern (Figure 15) appears on the WSM as a result of non-uniformities created in the thin film deposition process or an uneven temperature distribution during the rapid thermal annealing process.

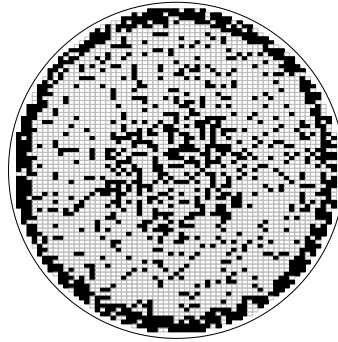


Figure 15: Ring pattern

- Linear Scratch

The linear scratch pattern (Figure 16) on the WSM could be a result of material shipping and handling during the manufacturing or testing.

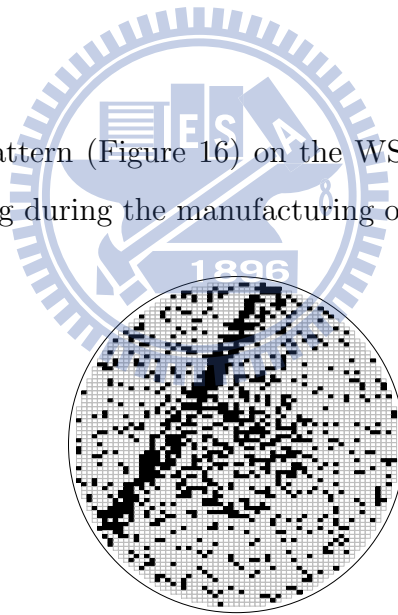


Figure 16: Linear scratch pattern

- Crescent Moon

The crescent moon pattern (Figure 17) could be the result of defective wafer materials or adverse processes. For example, a fab engineer who notices this pattern might decide to immediately look at the rapid thermal anneal (RTA) process.

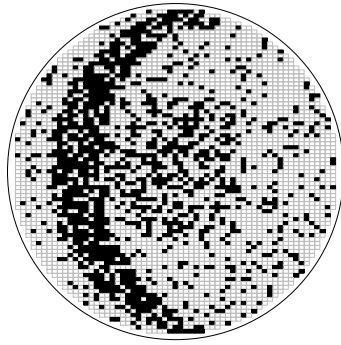


Figure 17: Crescent moon pattern



3 Literature Review

In this section, we review some areas of research related to wafer maps, two spatial randomness tests for testing the spatial randomness of wafer maps, and the kernel density estimation that we use as a tool in our proposed schemes.

3.1 Related Studies

These areas in semiconductor manufacturing related to wafer maps are briefly described here, including decision systems, yield models, and pattern recognition.

1. Decision System

Decision system is namely the application of various kinds of knowledge systems to failure analysis. It integrates parameter analysis and engineering experiments to help engineers effectively find the assignable causes and make decisions.

Two early examples of knowledge systems used in semiconductor manufacturing are PIES (Pan and Tenenbaum [13]) and SMART (Mary [14]) that diagnose problems in semiconductor fabrication processes by analyzing parametric test data. Maly *et al.* [15] recommended using a hierarchical methodology for the interpretation of test data. Methods such as CART (Breiman *et al.* [16]) and decision tree (Venkat [17]) would be useful in developing a decision system.

2. Yield Models

Yield in many ways is the most important financial factor in producing ICs. This is because yield is inversely proportional to the total manufacturing cost. The higher the yield is, the lower is the cost.

A yield model that provides good estimates of manufacturing yield can help predict product cost, determine optimum equipment utilization, or be used as a metric for supporting decisions involving new technologies and the identification of problematic products or processes. Cunningham [18] provided a good historical review of yield models.

Yield prediction of semiconductor dies can be used to:

- determine the cost of a new chip before fabrication,

- identify the cost of defect types for a particular chip or a range of chips,
- estimate the number of wafer starts required,
- show which defect types accounted for the most yield loss,
- identify when a fabrication process is not performing as expected,
- determine the extent of parametric problems (in both design and process),
- monitor the fabrication process.

3. Pattern Recognition

Since wafer maps contain important information that could be used to trace the process failures back to their root causes for quality engineers, how to recognize wafer patterns is an important issue.

Gleason *et al.* [19] employed an automated clustering algorithm using artificial intelligence. Chen and Liu [20] and Liu *et al.* [21] used neural networks for pattern recognition. Lee *et al.* [22] adopted a self-organized feature map for advanced process controls. Chao and Tong [23] used multi-class support vector machines with a novel defect cluster index for pattern recognition. The above methods need a large number of good training samples in order to successfully recognize defect patterns.

In the process of pattern recognition, performing a spatial randomness test is an important step to classify raw WSMs into two categories, patterned or random. If the spatial randomness test is too sensitive, the frequency of false alarms would be large. Conversely, if the spatial randomness test is not powerful enough, process failures may not be detected and opportunities of quality improvement are lost. In the following, we describe two existing spatial randomness tests.

3.2 Spatial Randomness Tests

In this subsection, we review two existing spatial randomness tests, respectively, proposed by Tamm and Hamada [4] and Hansen, Nair, and Friedman [3], for semiconductor manufacturing applications. Both methods use the joint-count statistics to measure the adjacencies between good and bad dies (Weszka *et al.* [24]). The basic idea is based on

comparing the number of good dies around a bad die and the number of bad dies around a good die.

3.2.1 Log Odds Ratio Test

Given a WSM and its site map, let $Y_i = 0$ and $Y_i = 1$ denote the die at site i being good or bad, respectively.

A neighboring relationship is formed when two dies are located in the neighborhood of each other. In general, the king-move neighborhood that consists of a central die and its eight surrounding neighbors, as depicted in Figure 18, is the most popular construction rule.

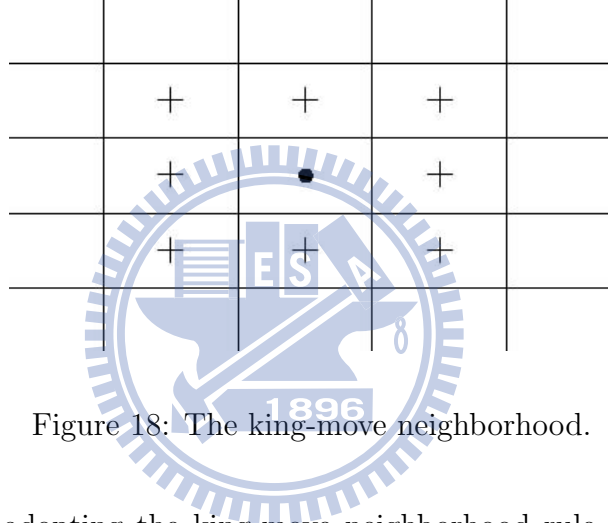


Figure 18: The king-move neighborhood.

In this study, we adopting the king-move neighborhood rule. Denote the set of all neighboring relationships of a wafer by \mathbf{H} , that is, the notation $(i, j) \in \mathbf{H}$ implies that two dies i and j are neighbors. Let N denote the total number of dies per wafer. Then, the following four statistics can be computed:

$$N_{GG} = \sum \sum_{i < j} \delta_{ij} (1 - Y_i) (1 - Y_j), \quad (3.2.1a)$$

$$N_{GB} = \sum \sum_{i < j} \delta_{ij} (1 - Y_i) Y_j, \quad (3.2.1b)$$

$$N_{BG} = \sum \sum_{i < j} \delta_{ij} Y_i (1 - Y_j), \quad (3.2.1c)$$

$$N_{BB} = \sum \sum_{i < j} \delta_{ij} Y_i Y_j, \quad (3.2.1d)$$

where

$$\delta_{ij} = \begin{cases} 1, & (i, j) \in \mathbf{H}, \\ 0, & \text{otherwise.} \end{cases}$$

Take Figure 19 as an illustrative example. If the i th die is a good die, then the summand is 0 for N_{BG} and N_{BB} , 1 for N_{GG} , and 3 for N_{GB} , corresponding to the king-move neighborhood rule and (3.2.1).

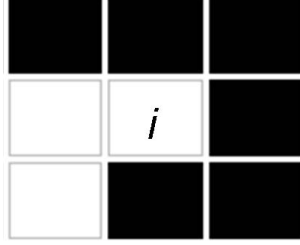


Figure 19: An example to depict the calculation of the four joint-count statistics.

To measure spatially associative effects on the WSM, Tamm and Hamada [4] proposed the following log odds ratio

$$\hat{\theta} = \log\left(\frac{(N_{GG} + 0.5)(N_{BB} + 0.5)}{(N_{GB} + 0.5)(N_{BG} + 0.5)}\right), \quad (3.2.2)$$

by employing the king-move neighborhood rule.

When the total number of dies on a wafer N is large, $\hat{\theta}$ is approximately normal distributed with mean 0 and variance (Agresti [25])

$$\sigma^2 = \left(\frac{1}{N_{GG} + 0.5} + \frac{1}{N_{BG} + 0.5} + \frac{1}{N_{GB} + 0.5} + \frac{1}{N_{BB} + 0.5}\right). \quad (3.2.3)$$

Thus, when $|\hat{\theta}|$ is greater than the critical point determined by the above approximate normal distribution, there is significant evidence to claim that the WSM is non-random; otherwise, claim that the WSM is random.

3.2.2 HNF Test

For convenience, the test developed by Hansen *et al.* [3] is referred to as the HNF test here after. In the test, two weighted “joint-count” statistics are computed, one counts the number of bad neighbors of bad dies and the other counts the number of good neighbors of good dies. To make this intuitive formulation mathematically precise, we introduce some notation.

Let \mathcal{N} represent the collection of die locations on a given wafer and N denote the total number of dies on a wafer. And let $\mathcal{N}_1 \subset \mathcal{N}$ denote the positions of the bad dies on

a wafer and $N_1 = |\mathcal{N}_1|$, the number of bad dies. Similarly, let $\mathcal{N}_0 = \mathcal{N} \setminus \mathcal{N}_1$, the locations of the good dies and $N_0 = |\mathcal{N}_0|$. Hence, $N = N_1 + N_0$. Next, for $i \in \mathcal{N}$, let $I_{\mathcal{N}_1}(i)$ be the indicator function of the event that the die at site i is bad, and set $I_{\mathcal{N}_0}(i) = 1 - I_{\mathcal{N}_1}(i)$. Finally, we let $p_i = EI_{\mathcal{N}_1}(i)$, the probability that the die at site i is bad, and set $q_i = 1 - p_i$, the probability that the same die is good.

In the absence of any non-random patterns, assume that bad dies are distributed randomly across the wafer. More specifically, we assume that the variables $I_{\mathcal{N}_1}(i)$'s are independent and identically distributed as Bernoulli random variables with constant probability $p_i = p, i \in \mathcal{N}$. In other words, we have an SHBP over the die locations in \mathcal{N} (Cliff and Ord [26]).

After the wafer sort, the proportion of bad dies on a wafer, \hat{p} , can be computed. Consider the statistic $\mathcal{T} = (T_{\mathcal{N}_0}, T_{\mathcal{N}_1})'$ to test the null hypothesis that there is no non-random pattern, where

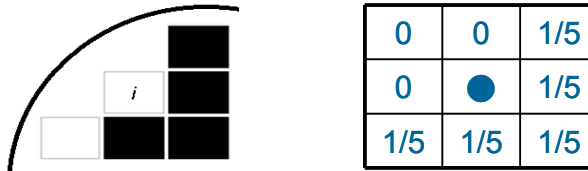
$$T_{\mathcal{N}_0} = N^{-1} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} w_i(j) I_{\mathcal{N}_0}(i) I_{\mathcal{N}_0}(j), \quad (3.2.4)$$

$$T_{\mathcal{N}_1} = N^{-1} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} w_i(j) I_{\mathcal{N}_1}(i) I_{\mathcal{N}_1}(j), \quad (3.2.5)$$

and, for each $i \in \mathcal{N}$, $\{w_i(j), j \in \mathcal{N}\}$ denotes a set of nonnegative weights normalized so that $\sum_{j \in \mathcal{N}} w_i(j) = 1$.



(a) an interior die



(b) a boundary die

Figure 20: An example to depict the calculation of the pair \mathcal{T} statistics.

Figure 20 depicts two illustrative examples with the right panel indicating the weight

for each neighboring site. In the examples, weights are equally distributed among all the neighbors for each die. Let

$$T_{\mathcal{N}_0}(i) = N^{-1} \sum_{j \in \mathcal{N}} w_i(j) I_{\mathcal{N}_0}(i) I_{\mathcal{N}_0}(j)$$

and

$$T_{\mathcal{N}_1}(i) = N^{-1} \sum_{j \in \mathcal{N}} w_i(j) I_{\mathcal{N}_1}(i) I_{\mathcal{N}_1}(j).$$

Then $T_{\mathcal{N}_0} = \sum_{i \in \mathcal{N}} T_{\mathcal{N}_0}(i)$ and $T_{\mathcal{N}_1} = \sum_{i \in \mathcal{N}} T_{\mathcal{N}_1}(i)$.

For Figure 20(a), we have $T_{\mathcal{N}_1}(i) = 0$ and $T_{\mathcal{N}_0}(i) = N^{-1}(\frac{1}{8} \cdot 1 \cdot 1 + \frac{1}{8} \cdot 1 \cdot 1)$; and for Figure 20(b), $T_{\mathcal{N}_1}(i) = 0$ and $T_{\mathcal{N}_0}(i) = N^{-1} \cdot \frac{1}{5} \cdot 1 \cdot 1$, corresponding to the king-move neighborhood and the associated weights of neighbors.

The following are the exact first and second moments of the pair \mathcal{T} statistic conditional on $\hat{p} = N_1/N$. Note that N is fixed, so being conditional on \hat{p} is equivalent to being conditional on N_0 or on N_1 . Thus, the exact first moments are

$$E(T_{\mathcal{N}_0} | N_0) = \frac{N_0(N_0 - 1)}{N(N - 1)} \quad (3.2.6)$$

and

$$E(T_{\mathcal{N}_1} | N_1) = \frac{N_1(N_1 - 1)}{N(N - 1)}, \quad (3.2.7)$$

assuming that $w_j(j) = 0$ for all $j \in \mathcal{N}$ (Cliff and Ord [26]). Next, the second moments of the pair \mathcal{T} are

$$E(T_{\mathcal{N}_0}^2 | N_0) = \alpha_1 \frac{N_0}{N} + \alpha_2 \frac{N_0(N_0 - 1)}{N(N - 1)} + \alpha_3 \frac{N_0(N_0 - 1)(N_0 - 2)}{N(N - 1)(N - 2)} + \alpha_4 \frac{N_0(N_0 - 1)(N_0 - 2)(N_0 - 3)}{N(N - 1)(N - 2)(N - 3)} \quad (3.2.8)$$

and

$$E(T_{\mathcal{N}_1}^2 | N_1) = \alpha_1 \frac{N_1}{N} + \alpha_2 \frac{N_1(N_1 - 1)}{N(N - 1)} + \alpha_3 \frac{N_1(N_1 - 1)(N_1 - 2)}{N(N - 1)(N - 2)} + \alpha_4 \frac{N_1(N_1 - 1)(N_1 - 2)(N_1 - 3)}{N(N - 1)(N - 2)(N - 3)}, \quad (3.2.9)$$

where $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$ and $\alpha_1 = 0$ if $w_j(j) = 0$ for all $j \in \mathcal{N}$. Note the symmetry between (3.2.6) and (3.2.7), also between (3.2.8) and (3.2.9). The four α_k 's are defined as

$$\alpha_k = N^{-2} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} \sum_{i' \in \mathcal{N}} \sum_{j' \in \mathcal{N}} w_i(j) w_{i'}(j') I_k(i, j, i', j'), \quad (3.2.10)$$

where I_k returns the value of 1 if there are only k different elements among its four arguments and 0 otherwise. Roughly speaking, quantities α_k 's are obtained by computing the proportions of the terms in the expansion of the second moment that involve two, three, and four unique indicators, respectively, and hence must be computed separately for every different die layout \mathcal{N} . Finally, if $w_j(j) = 0$ for all $j \in \mathcal{N}$ as one would usually set, then the covariance between $T_{\mathcal{N}_0}$ and $T_{\mathcal{N}_1}$ can be computed by

$$E(T_{\mathcal{N}_0} T_{\mathcal{N}_1} | N_0, N_1) = \alpha_4 \frac{N_1(N_1 - 1)N_0(N_0 - 1)}{N(N - 1)(N - 2)(N - 3)}. \quad (3.2.11)$$

Assume that N_1/N does not go to 0 as $N \rightarrow \infty$. Then, with the theory of U statistics (Lee [27]), it can be shown that, the statistic \mathcal{T} under null hypothesis has a bivariate normal limiting distribution as N (and N_1) tends to infinity (Hansen *et al.* [3]). This implies

$$\mathcal{T} = \begin{pmatrix} T_{\mathcal{N}_0} \\ T_{\mathcal{N}_1} \end{pmatrix} \bigg|_{N_0, N_1} \sim BN(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) \quad \text{as } N \rightarrow \infty, \quad (3.2.12)$$

and the exact conditional expectation ($\boldsymbol{\mu}_1$) and conditional variance and covariance matrix ($\boldsymbol{\Sigma}_1$) can be obtained by (3.2.6) – (3.2.11). Thus, conditional on N_0, N_1 , the test statistic

$$D = (\mathcal{T} - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_1^{-1} (\mathcal{T} - \boldsymbol{\mu}_1)$$

follows the chi-square distribution with degrees of freedom 2. Accordingly, the critical value of the test can be determined.

3.2.3 HNF Method Under Cluster

As mentioned in Section 1, the bad dies being distributed randomly across the wafer is the ideal situation for a wafer. Unfortunately, almost all wafer sort maps exhibit some regions of inherent failures in practice. Hansen *et al.* [3] also considered a *clustered* alternative hypothesis and derived the exact distribution of the \mathcal{T} statistic conditional on a cluster \mathcal{C} and $\hat{p} = N_1/N$.

To describe more precisely, let $\mathcal{C}_0, \mathcal{C}_1 \subset \mathcal{N}$ denote the set of good and bad die sites, respectively, and set $\mathcal{C} = \mathcal{C}_0 \cup \mathcal{C}_1$. Generically, we refer to \mathcal{C} as a cluster. Let C_0 and C_1

be the number of the good dies and bad dies in \mathcal{C} , respectively. For any sets $\mathcal{I}, \mathcal{J} \subset \mathcal{N}$, we define

$$T(\mathcal{I}, \mathcal{J}) = N^{-1} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} w_i(j) I_{\mathcal{I}}(i) I_{\mathcal{J}}(j), \quad (3.2.13)$$

where $I_{\mathcal{I}}(i) = 1$ if $i \in \mathcal{I}$; 0 otherwise.

Then the exact mean of \mathcal{T} conditional on the event that \mathcal{C}_0 contains all the good dies in \mathcal{C} can be computed by:

$$E(T_{\mathcal{N}_0} | N_0, N_1, \mathcal{C}_0, \mathcal{C}_1) = T(\mathcal{C}_0, \mathcal{C}_0) + [T(\mathcal{C}_0, \mathcal{C}^c) + T(\mathcal{C}^c, \mathcal{C}_0)] \frac{N_0 - C_0}{N - C} + T(\mathcal{C}^c, \mathcal{C}^c) \frac{(N_0 - C_0)(N_0 - C_0 - 1)}{(N - C)(N - C - 1)} \quad (3.2.14)$$

and

$$E(T_{\mathcal{N}_1} | \hat{p} = N_0, N_1, \mathcal{C}_0, \mathcal{C}_1) = T(\mathcal{C}_1, \mathcal{C}_1) + [T(\mathcal{C}_1, \mathcal{C}^c) + T(\mathcal{C}^c, \mathcal{C}_1)] \frac{N_1 - C_1}{N - C} + T(\mathcal{C}^c, \mathcal{C}^c) \frac{(N_1 - C_1)(N_1 - C_1 - 1)}{(N - C)(N - C - 1)}, \quad (3.2.15)$$

where $\mathcal{C}^c = \mathcal{N} \setminus \mathcal{C}$, the complement of \mathcal{C} .

The expressions for second moments are complicated and not presented here. These expressions also depend on the α_k 's defined in (3.2.10). Recall that α_k 's vary with the wafer layout. In addition, when there exists a baseline \mathcal{C} , these α_k 's depend not only on the set \mathcal{C} but also on \mathcal{C}_0 and \mathcal{C}_1 .

The joint distribution of the pair $\mathcal{T} = (T_{\mathcal{N}_0}, T_{\mathcal{N}_1})$ conditional on $(\mathcal{N}_0, \mathcal{N}_1, \mathcal{C}_0, \mathcal{C}_1)$ is approximately bivariate normal when $N - C$ is large, that is

$$\mathcal{T} = \begin{pmatrix} T_{\mathcal{N}_0} \\ T_{\mathcal{N}_1} \end{pmatrix} \bigg|_{N_0, N_1, \mathcal{C}_0, \mathcal{C}_1} \sim BN(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) \quad \text{as } N - C \rightarrow \infty, \quad (3.2.16)$$

where $\boldsymbol{\mu}_2$ given in (3.2.14) and (3.2.15) and $\boldsymbol{\Sigma}_2$ can be found in Ooi *et al.* [28]. With (3.2.16), the critical value of the randomness test can be determined.

However, the computation of the exact second moments are too extensive to carry out in practice, we will propose an alternative method to compute $\boldsymbol{\mu}_2$ and $\boldsymbol{\Sigma}_2$ by constructing some pseudo wafers.

3.3 Mode Estimation Via Kernel Density Estimation

In this subsection, we give a brief description on kernel density estimation, a tool we use to estimate the mode of a distribution in proposed schemes.

Kernel density estimation has a long tradition of estimating the probability density function of a random variable in nonparametric way. With the kernel density estimate, the mode of the distribution can be estimated by the maximizer of the kernel density estimate. Parzen [29] gave consistency, asymptotic normality, and mean squared error evaluation of the kernel sample mode, and his results have been extended in several directions by Chernoff [30], Eddy [31], and Romano [32].

Given the random sample x_1, x_2, \dots, x_n , which follow a continuous, univariate probability density function f , the kernel density estimator is

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right), \quad (3.3.1)$$

where K is some kernel function and h is a smoothing parameter called the bandwidth. The bandwidth controls the smoothness of the estimated curve (Silverman [33], Chen [34], Shi and Zhang [35], and Gasser and Muller [36]). Small values of h force the expected value of the estimate $\hat{f}_h(x)$ to be close to the true value $f(x)$, but the price to pay is the high variability of the estimate, since it is based on comparatively few observations. On the other hand, variability can be decreased by increasing h . Quite often K is taken to be a standard Gaussian function with mean 0 and variance 1. That is,

$$K\left(\frac{x - x_i}{h}\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x - x_i)^2}{2h^2}}. \quad (3.3.2)$$

Given the parameter h and kernel K , we take the maximizer of the estimated kernel density \hat{f}_h to be the estimate of the mode of f .

4 Proposed Schemes

4.1 Data Encoding

The WS data are three-dimensional data with the (x, y) position and the binary test result value for each die. Figure 21 illustrates how to transform WS data (left panel) to a one-dimensional binary sequence according to the ordering shown in the site map (right panel).

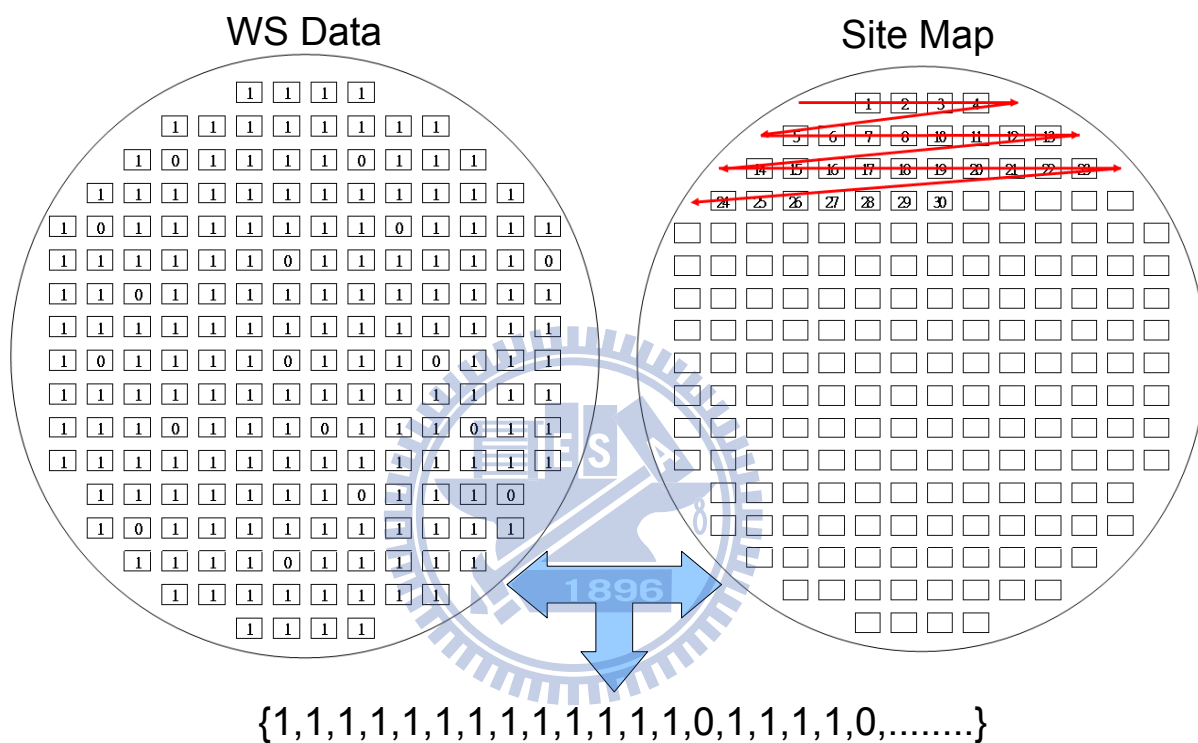


Figure 21: Data encoding.

4.2 Proposed Method

Loosely speaking, the baseline is a region treated by engineers as an accepted flaw in practice. When they judge a wafer to be random or patterned, they simply ignore whether the situation is inside the baseline. Therefore, to them, an automatic classification method should also ignore whether within the baseline. Also, baselines have various shapes and they will definitely change as the manufacturing process improves/changes over time. So desirable automatic pattern detection method should be able to adapt to any kind of baselines set by engineers.

Unfortunately, the odds ratio test and HNF test described in the previous section cannot distinguish between the baseline and the “genuine” non-random patterns. The HNF method under cluster dose takes the baseline into account, but it is too computationally intensive to be feasible for practice use. To meet engineers’ need, we modify these three randomness tests with a simple approach that takes the baseline into account. The idea is fairly simple: masking the unwanted area, e.g., the baseline, with a randomly generated Bernoulli sequence to remove the unwanted patterns so that the existing methods are applicable.

Two types of pseudo wafers, called pseudo-inside and pseudo-outside WSMs, are considered. For the WSM under test, a pseudo-inside WSM is generated by replacing the die data in the baseline with a set of generated variates that follows an SHBP with the probability of $\{Y = 0\}$ to be the yield outside the baseline. If the dies outside the baseline indeed follow an SHBP, then the whole pseudo WSM also follows the same SHBP approximately. Then the first two randomness tests can be applied to the pseudo WSM.

Similarly, a pseudo-outside WSM is generated by replacing the binary die data outside the baseline with a sequence of SHBP data. The procedure for generating a sequence of pseudo-inside (pseudo-outside) is as follows.

1. Given the baseline region, calculate $y_{outside}$, the yield outside the baseline region.
2. Generate a Bernoulli sequence with the parameter $1 - y_{outside}$ and the length is the total number of dies inside (outside) the baseline region.
3. Replace the inside-baseline (outside-baseline) data with the Bernoulli sequence according to the site map.
4. Repeat steps 2–3 n times.

Figure 22 and 23 illustrate the generation of pseudo-inside and pseudo-outside WSMs, respectively. The left-hand-side of the arrow is the original WSM with a central-disc baseline and the right-hand-side are n pseudo WSMs obtained by the above procedure.

The purpose of generating n pseudo-inside WSMs is to find a “good” substitute for the original WSM. We could have simply generated just one pseudo WSM and apply the odds ratio or HNF test to it. But we are concerned about getting not-so-good substitute due to sampling. With n replicates, one could use the sample mean of n test statistics for

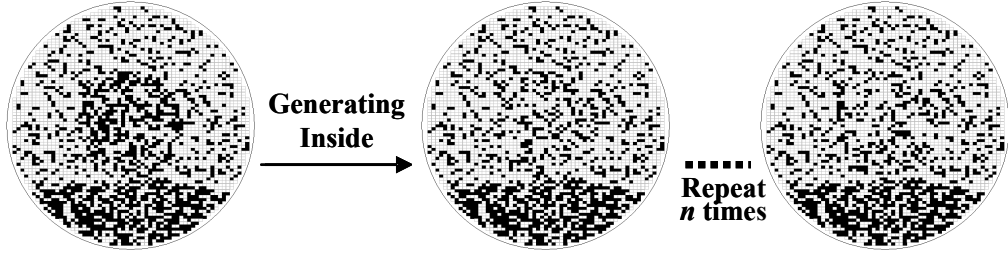


Figure 22: Generating a sequence of pseudo-inside WSMs from a WSM.

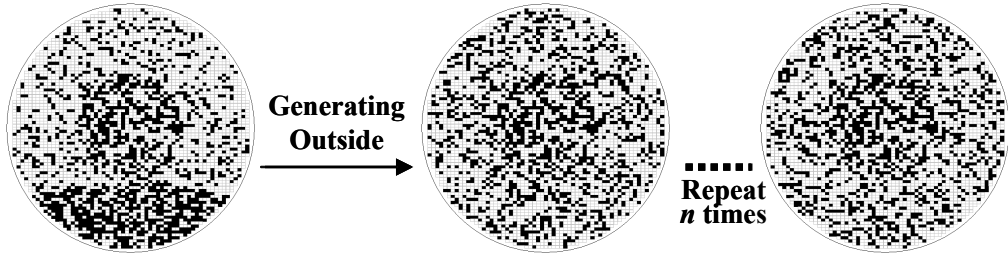


Figure 23: Generating a sequence of pseudo-outside WSMs from a WSM.

reducing the sampling error, or the sample median for robustness, or the sample mode for the high likelihood. Here we describe the sample mode approach because it is slightly complicated than using mean and median.

With this generating procedure, we now are ready to describe the proposed schemes.

4.2.1 Mode Odds Ratio Method

Note that, for the same original WSM, the approximate variance (3.2.3) of the test statistic $\hat{\theta}$ of a pseudo-inside WSM depends on the statistics obtained from the pseudo WSM, hence each pseudo WSM has its own value of σ^2 . By normalizing $\hat{\theta}$, we have

$$Z = \frac{\hat{\theta}}{\sigma} \stackrel{\text{approx.}}{\sim} N(0, 1). \quad (4.2.1)$$

To describe the proposed mode odds ratio method, we use the WSM shown in Figure 22 as an illustrative example. Generate 100 pseudo WSMs. For the j th pseudo WSM, $j = 1, 2, \dots, 100$, obtain $\hat{\theta}_j$ then Z_j . Applying the kernel density estimation method to $\{Z_j, j = 1, \dots, 100\}$ to obtain the kernel density estimate and then its mode. The mode represents the most probable Z value for the WSM under study when pretending that the baseline region is random. Figure 24 presents the kernel density estimate of the 100

Z_j 's in this case. At the significance level $\alpha = 0.01$, the mode estimate is $Z^* = 11.7558 > z_{.005} = 2.576$, hence we reject the null hypothesis of randomness, i.e., special patterns other than the baseline exist for this WSM, which indeed is the case.

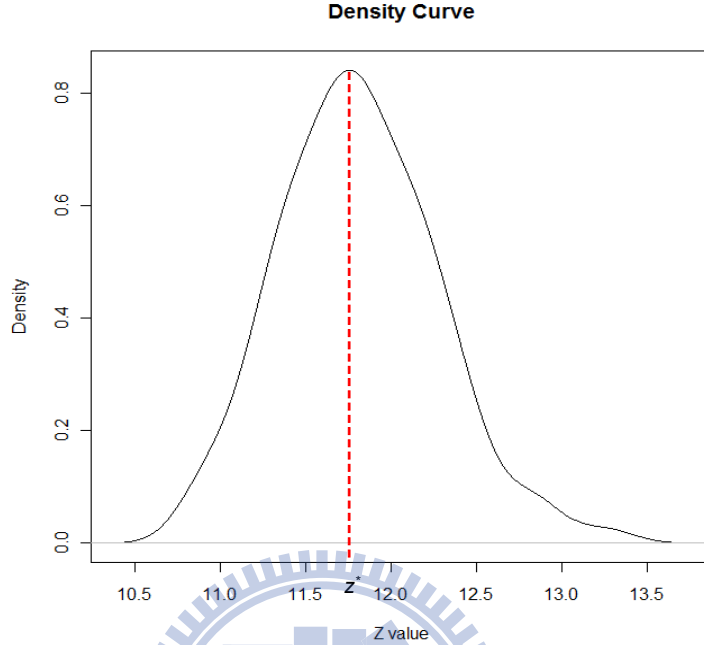


Figure 24: Mode estimation of Z_j values

4.2.2 Mode HNF Method

Generate n pseudo-inside WSMs from the WSM to be tested as described earlier in subsection 2. For the j th pseudo-inside WSM data, $j = 1, \dots, n$, carry out the following statistic

$$D_j = (\mathcal{T}_j - \boldsymbol{\mu}_j)' \boldsymbol{\Sigma}_j^{-1} (\mathcal{T}_j - \boldsymbol{\mu}_j), \quad (4.2.2)$$

where $\boldsymbol{\mu}_j$ and $\boldsymbol{\Sigma}_j$ can be obtained from (3.2.6)–(3.2.9). By the result given in Hansen *et al.* [3], D_j follows $\chi^2(2)$ asymptotically, the chi-square distribution with 2 degrees of freedom. Apply the kernel density estimation method to these D_j values and obtain the mode of the kernel density estimate, denoted by D^* . Then we can judge whether there exist special patterns by comparing the mode with the critical value $\chi_\alpha^2(2)$, the $100(1 - \alpha)\%$ percentile of $\chi^2(2)$. Take the same example as described in the last subsection as an example. The same 100 pseudo-inside WSMs given a kernel density estimate and its mode $D^* = 147.02$ as shown in Figure 25. Take $\alpha = 0.01$ as an example. Since D^* is far greater than the

critical value $\chi_{0.01}^2(2) = 9.21034$, we reject the null hypothesis of randomness, and claim that the original WSM exhibits special patterns other than the baseline.

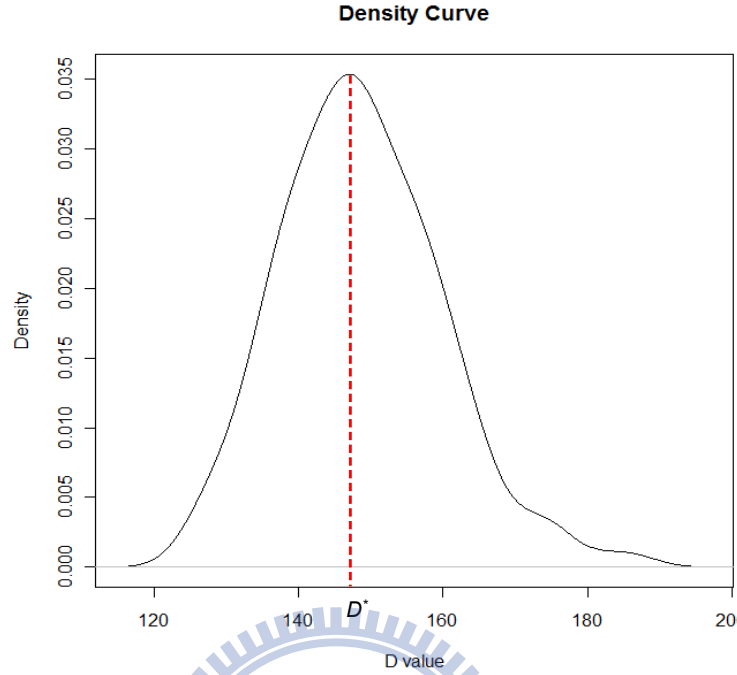


Figure 25: Mode estimation of D_j values

4.2.3 Empirical HNF Method

The method described in Subsection 3.2 does take the baseline into account, but the computation is too intensive to be feasible for online testing. For example, it took almost 30 minutes on a regular PC to complete the randomness test for one single wafer. In this subsection, we propose an alternative method by estimating the mean and covariance matrix empirically with pseudo-outside WSMs.

For a given WSM, we first generate n pseudo-outside WSMs as described earlier. As in Subsection 3.2, compute the statistic \mathcal{T}_j for the j th pseudo WSM, $j = 1, \dots, n$, and obtain a pseudo sample of size n . Compute the sample mean $\bar{\mathcal{T}}$ and sample covariance matrix \mathbf{S} of this sample. Consider the statistic

$$F = \frac{n(n-p)}{(n^2-1)p} (\mathcal{T} - \bar{\mathcal{T}})' \mathbf{S}^{-1} (\mathcal{T} - \bar{\mathcal{T}}), \quad (4.2.3)$$

where \mathcal{T} is the statistic of the original WSM and $p = 2$ is the dimension of \mathcal{T} . Recall that conditional on N_0, N_1, C_0 , and C_1 , the distribution of each \mathcal{T}_j follows a bivariate normal

asymptotically. Therefore, if \mathcal{T}_j 's are independent and identically distribution (i.i.d.), then the test statistic F in (4.2.3) has an \mathcal{F} distribution with degrees of freedom 2 and $n - 2$ asymptotically and the critical value would be its $100(1 - \alpha)$ th percentile accordingly.

However, \mathcal{T}_j 's are not independent because all the pseudo WSMs have the same baseline. Fortunately, this loop hole can be mended with the following argument. Let \mathcal{C} be the set of dies in the baseline and $\mathcal{C}^c = \mathcal{N} \setminus \mathcal{C}$, the complement of \mathcal{C} . We say a die is in the boarder of \mathcal{C} and \mathcal{C}^c if it has a neighbor from the other set. Denote the boarder by \mathcal{B} . First note that, the contribution of all the dies in the boarder to $T_{\mathcal{N}_0}$ and $T_{\mathcal{N}_1}$, $\sum_{i \in \mathcal{B}} T_{\mathcal{N}_0}(i)$ and $\sum_{i \in \mathcal{B}} T_{\mathcal{N}_1}(i)$, can be neglected because $|\mathcal{B}|/|\mathcal{N}| \rightarrow 0$ as $N \rightarrow \infty$. Next, without involving neighbors in \mathcal{C}^c , we have $\sum_{i \in \mathcal{C} \setminus \mathcal{B}} T_{\mathcal{N}_0}(i) = c_0$ and $\sum_{i \in \mathcal{C} \setminus \mathcal{B}} T_{\mathcal{N}_1}(i) = c_1$ for all pseudo-outside WSMs, where c_0 and c_1 are two constants. On the other hand, the statistic $\sum_{i \in \mathcal{C}^c \setminus \mathcal{B}} T_{\mathcal{N}_0}(i)$ for each pseudo WSM are i.i.d. by their construction. The same argument

of Hansen *et al.* [3] would imply that $\left(\sum_{i \in \mathcal{C}^c \setminus \mathcal{B}} T_{\mathcal{N}_0}(i), \sum_{i \in \mathcal{C}^c \setminus \mathcal{B}} T_{\mathcal{N}_1}(i) \right)'$ follows a bivariate normal asymptotically. Combining all these terms, $(T_{\mathcal{N}_0}, T_{\mathcal{N}_1})'$ also follows a bivariate normal asymptotically. Thus we conclude that conditional on $(N_0, N_1, \mathcal{C}_0, \mathcal{C}_1)$, statistic F in (4.2.3) follows \mathcal{F} distribution with degrees of freedom 2 and $n - 2$ asymptotically.

For the same example as before, we generate 100 pseudo-outside WSMs to obtain 100 \mathcal{T}_j 's. For these \mathcal{T}_j 's, $\mathcal{T} = \begin{pmatrix} 0.4723 \\ 0.1350 \end{pmatrix}$, $\bar{\mathcal{T}} = \begin{pmatrix} 0.4519 \\ 0.1148 \end{pmatrix}$, and $\mathbf{S} = \begin{pmatrix} 2.16e - 05 & -5.84e - 06 \\ -5.84e - 06 & 5.54e - 06 \end{pmatrix}$. Let $\alpha = 0.01$. Then by (4.2.3), $F = 380.1747 > \mathcal{F}_{2,98}(0.01) = 4.8285$. Thus, the example shows that the Empirical HNF test rejects the null hypothesis of being random and claim the WSM has some patterns other than the baseline.

Finally, we summarize these three proposed schemes in the following flow chart (Figure 26).

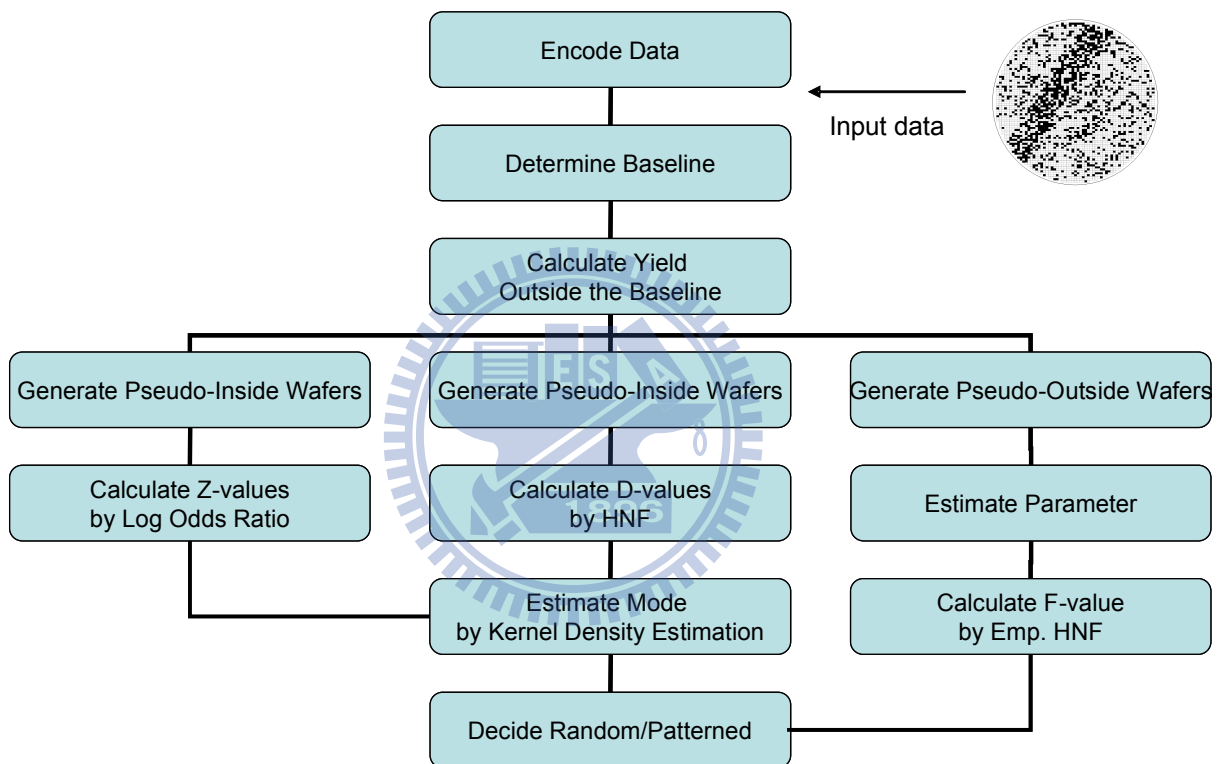


Figure 26: The flow chart of the proposed schemes.

5 Simulation Studies

We first discuss how to create the platform of the WSM. In this part, the total number of dies on a wafer is determined by the simulation parameters. After the simulation parameters are set, we then focus on how to generate various patterns on a wafer. Finally, we evaluate the performances of the three proposed schemes in terms of the false alarm rate and detection power under various yield levels and patterns.

5.1 Simulation Settings

In our simulation, the number of rows and columns and total number of dies are determined by the 5 spatial parameters. The 5 parameters are as follows and illustrated in Figure 27:

- r denotes the radius of the wafer
- h denotes the interval between two dies
- k denotes the length of the retained wafer edge
- l denotes the length of each die
- w denotes the width of each die

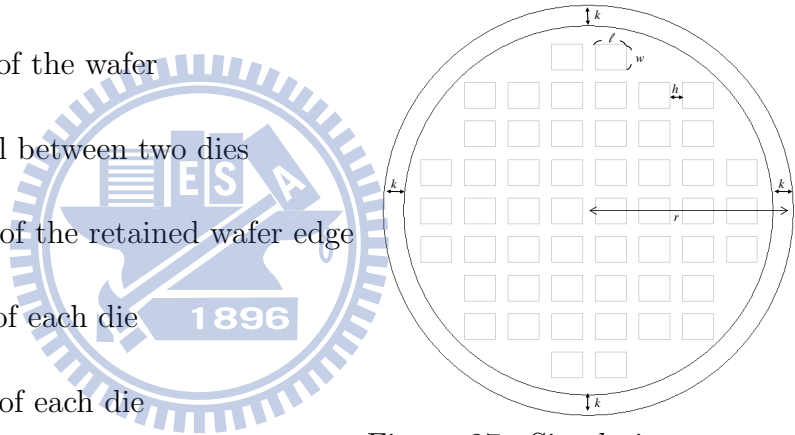


Figure 27: Simulation parameters for wafer.

For example, the simulation parameters $(r, h, k, l, w) = (8, 0.5, 0.8, 1.2, 1)$ would correspond to a wafer with 9 rows and 8 columns, and the total number of dies is 52. Figure 27 and Table 1 illustrate the details of the wafer layout under this setting.

Table 1: An example for wafer creation.

row	# of dies	x -coordinate		y -coordinate
		start	end	
1	2	-0.85	0.85	6.0
2	6	-4.25	4.25	4.5
3	6	-4.25	4.25	3.0
4	8	-5.95	5.95	1.5
5	8	-5.95	5.95	0.0
6	8	-5.95	5.95	-1.5
7	6	-4.25	4.25	-3.0
8	6	-4.25	4.25	-4.5
9	2	-0.85	0.85	-6.0

5.2 Generating Patterned Wafer Sort Maps

After the wafer layout is set, we focus on how to generate patterned wafer sort maps. In this study, we pick four wafer patterns, the linear scratch, ring, bottom, and crescent moon. Here, we explain how to generate the linear scratch pattern as an illustrative example.

For the linear scratch pattern, the wafer is separated into two disjoint areas, the linear scratch (B) area and its complement (A), as illustrated in Figure 28. To generate a linear scratch pattern, simply generate a Bernoulli variate for each die with a “success” probability to be a bad die higher for dies in patterned area B than in A . That is, the yield in A is higher than in B .

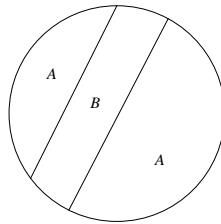


Figure 28: Illustration for generating a linear scratch pattern.

5.3 Comparisons

In this simulation, the wafer layout parameters are $(r, h, k, l, w) = (8, 0.01, 0.15, 0.24, 0.18)$ and the total number of dies per wafer is 3913. Assume the baseline is the centered disc of radius 3.47 as shown in Figure 29. The area of the baseline is almost 20% of a wafer.

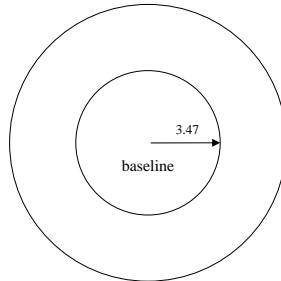


Figure 29: Baseline setting.

For patterned wafers with the baseline, Figure 30 illustrates four different patterns in our simulation studies. Divide a wafer into three areas, the baseline, the patterned area, and the random area. Let the yield in the random area be y_r , the yield in the baseline be $0.8y_r$, and the yield in the patterned area be $0.6y_r$. Denote the proportion of the patterned area is a wafer by b . Then the wafer yield is $y = (0.96 - 0.4b)y_r$.

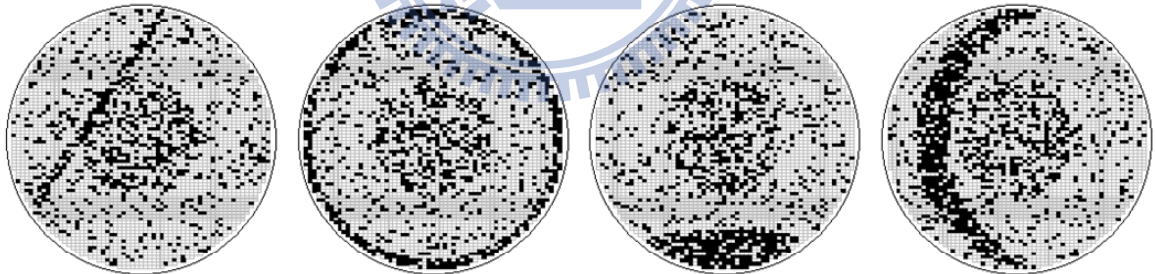


Figure 30: 4 different patterns with a central-disc baseline region.

In order to study and compare the performances of the proposed schemes under various conditions, we conduct the following three simulation studies.

The first simulation study compares the false-alarm rate of the three proposed methods and two existing methods with wafers of no special patterns. The second study compares the detecting power of the three methods under various yield level. The third study compares the detection power of the three methods under various sizes of the patterned area when the yield level is fixed. Two yield levels, including middle and high wafer

yield, are considered. For each case under study, we generate 150 WSMs and the level of significance is $\alpha = 0.01$.

5.3.1 False-alarm Rate

We generate 150 random WSMs with the baseline for each yield level y_r , for $y_r = 0.01(0.05)0.95$. We apply the existing odds ratio test and HNF tests as well as our three methods to the generated WSMs. The false-alarm counts are listed in Table 2 and the corresponding false-alarm rates are plotted in Figure 31.

From Figure 31, the two existing methods that do not take the baseline into account lead to a higher false-alarm rate, especially when the yield y is high. A high false-alarm rate will make engineers lose confidence in the auto detection method. On the other hand, our proposed schemes take the baseline into account and the resulting false-alarm rates are controlled at an about right level.

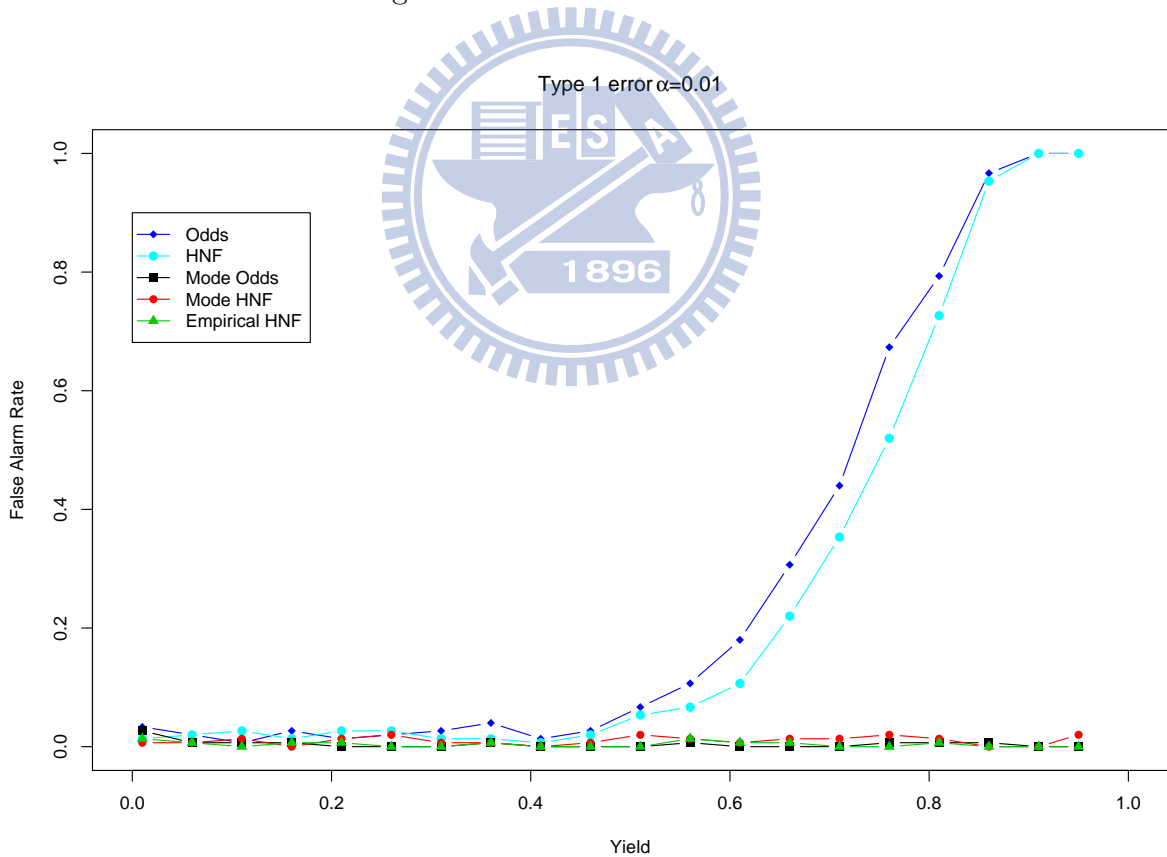


Figure 31: False-alarm rate of the two existing method and three proposed methods for various yield level.

5.3.2 Detecting Power vs. Yield

It is interesting to evaluate the performance in detecting various patterns. Because the size of the patterned area affects the performance, we fix the size at 20% of the wafer in this study.

Figure 32 and Tables 3–6, show the simulation results. It is observed that these methods have the same trend—the higher yield, the larger power. The detecting power of the mode odds ratio test is greater than the other two tests; and the mode HNF test seems performing slightly better than empirical HNF test. When the yield (y) is higher than 0.6, all three methods have 100% detecting power for all the patterns.

For the comparison of various patterns, it is observed that the linear scratch pattern is the hardest to detect. The bottom pattern is the easiest, followed by the crescent moon and edge ring pattern.

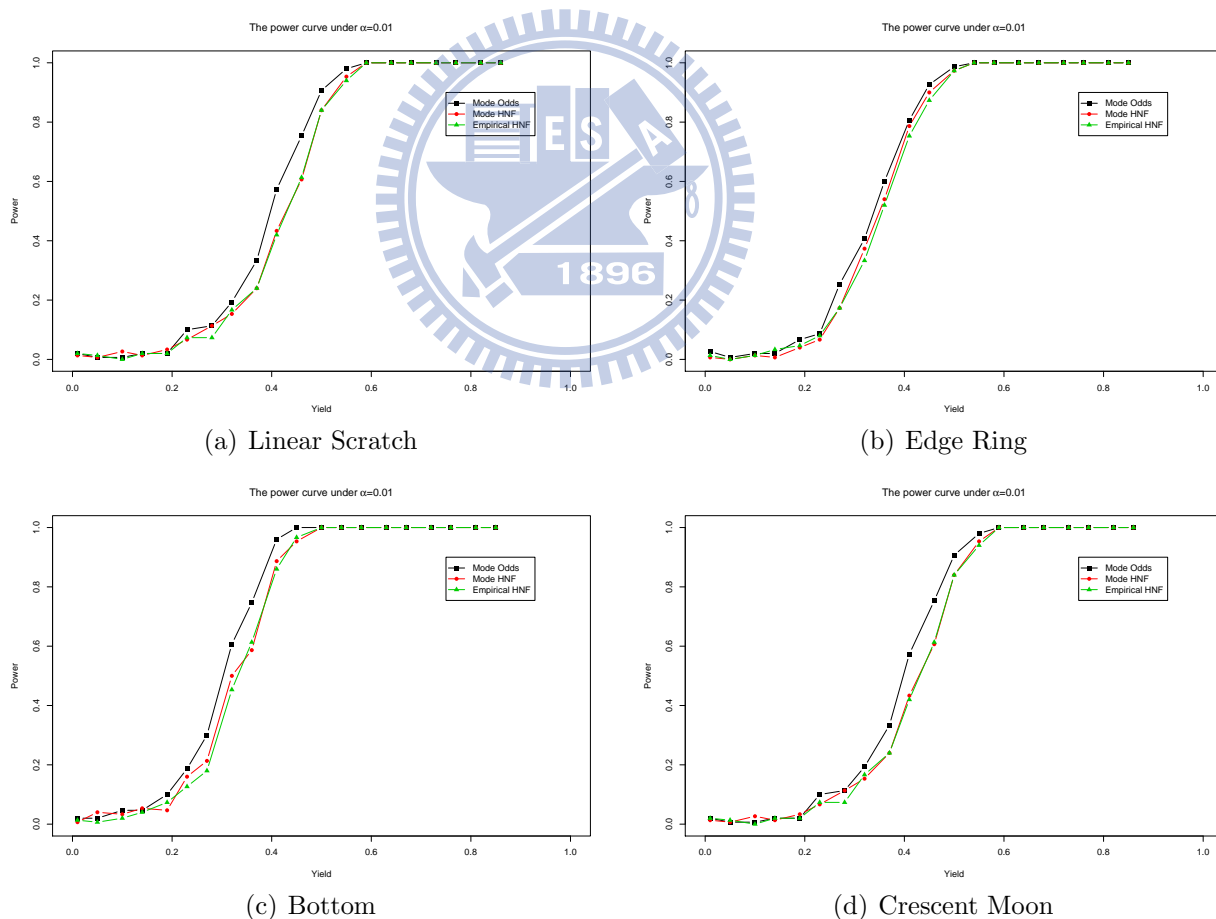


Figure 32: Detecting power for various patterns under a fixed patterned area size.

5.3.3 Detecting Power vs. Patterned Area

In Subsection 5.3.2, we have seen that the yield level has a great impact on the detecting power for all three proposed methods, and the higher the yield is, the better is the detecting power. This is very intuitive because it is easier to spot unusual patterns when the yield is high. Conversely, for a wafer with a very low yield, the bad dies in the random area will mask the pattern so that it would be hard to detect patterns, if they exist. We remark that detecting patterns has no practical meaning when the yield level is low, because it would be a time of emergency and engineers must stop the process to take necessary actions to find the root cause of the low yield.

Therefore, in studying how the proposed methods would react to the size of the patterned area, we consider two yield levels, $y_r = 0.75$ (middle) and 0.95 (high). In our simulation study, for each pattern, we start with almost no pattern and then gradually increase the size of the patterned area to about 20% of the wafer. For 20 area size considered in the simulation, the wafer yield is in the range of 0.7 ± 0.03 and 0.9 ± 0.03 for the middle and high yield levels, respectively.

For the middle and high yield level, Figure 33 and Figure 34 display the detecting power of the three methods over the size of the patterned area with panel (a) for linear scratch, (b) for edge ring, (c) for bottom, and (d) for crescent moon pattern, respectively. Table 7–14 show the simulation results.

From these figures, we observe

- Except for the edge ring pattern, the detecting power of all three methods increases as the area size increases. This is quite intuitive since the non-random pattern gets more apparent as the size gets larger.
- Except for the edge ring pattern, the mode odds ratio method performs the best and the two HNF methods perform very similar with the mode HNF method slightly better.
- In general, the bottom pattern is the easiest and the linear scratch pattern is the hardest to detect.
- The performances of the three methods behave quite differently from the other patterns. For this pattern, the mode odds ratio method performs the worst, but not

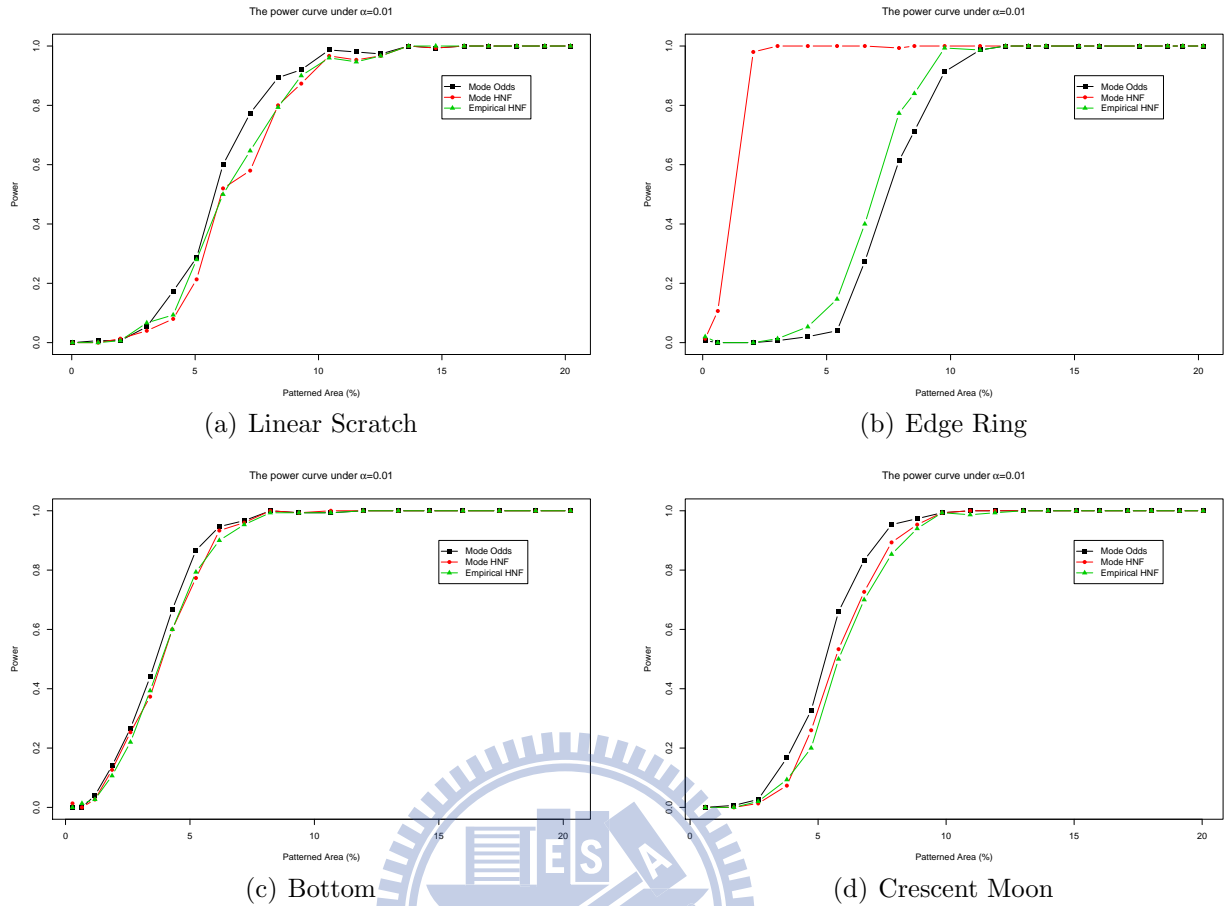


Figure 33: Detecting power for various patterns under middle yield level.

too far from the empirical HNF method. The behavior of the mode HNF method is worth mentioning. The detecting power increases very rapidly at low area, which is somewhat against the general impression that a small patterned area would be harder to detect than a lower patterned area. This could be due to some kind of edge effect. The edge ring pattern generates more bad dies at the edge than other patterns. A bad die at the edge has fewer neighbors so that each of its neighbors has a larger weight $w_i(j)$ than that of the interior dies. This may lead to a larger statistic T_{N_1} and hence higher power. This finding makes the mode HNF method a good choice for detecting a pattern with many dies at the edge.

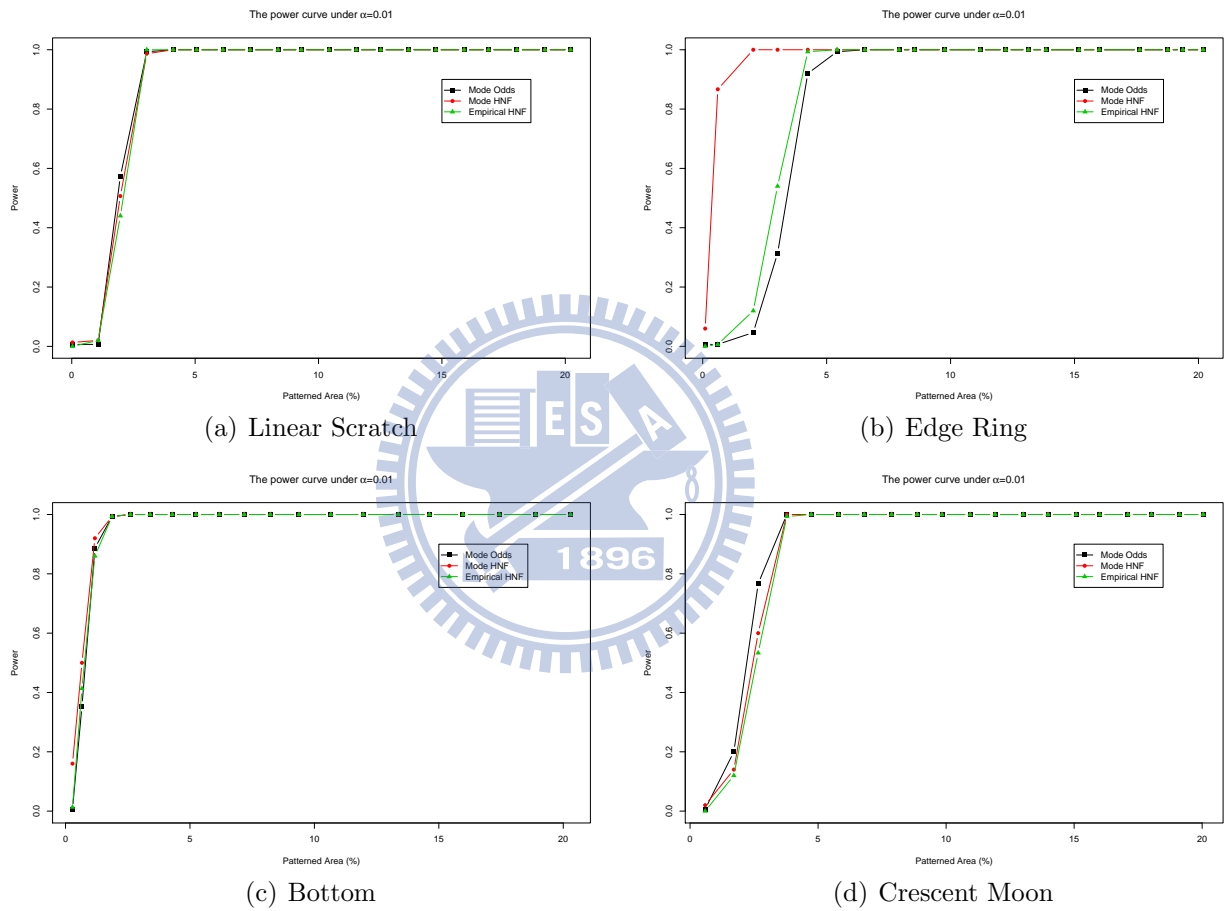


Figure 34: Detecting power for various patterns under high yield level.

6 Conclusion

The semiconductor industry is a very competitive industry. Continual improvement of the process stability and yield has become essential for each semiconductor company to stay competitive. Among quality improvement tools, the wafer sort map is commonly used for detecting abnormal processes since non-random patterns on a WSM often can provide clues about process problems and perhaps lead to possible causes. In addition, dies in a wafer with a random WSM are graded higher than those in a wafer that shows some non-random patterns in the die market. Thus, it is common practice to separate wafers into two categories—“random” and “non-random” (or “patterned”). However, most existing automatic methods cannot distinguish the “genuine” patterns and the “baseline’ patterns. The former are caused by the real process problems that engineers would like to detect and the later are inherent from the process limitations induced by present technology that engineers accepted as “normal” for now.

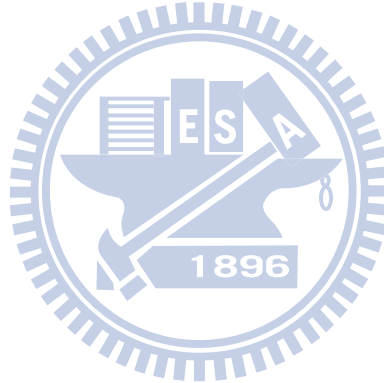
In this study, we take the baseline into consideration and propose three automatic schemes for detecting “genuine” patterned wafers with an approach of generating pseudo WSMs on which the existing randomness tests can be applicable. Simulation studies are conducted to evaluate and compare these new schemes. It is found that

- the yield level affects the detecting power greatly—the higher the yield, the larger the power;
- the size of the patterned area also affects the effectiveness of the detection—the larger the patterned area, the larger the power;
- not counting the the edge ring pattern, the mode odds ratio test performs the best but not too far from the rest;
- there exists an edge effect and the mode HNF method is very sensitive to it.

Therefore, considering the overall performance, the mode HNF method may be a good choice. However, in practice, a very small area edge pattern in human eyes most likely will be ignored and the wafer will be considered as a good wafer. So being very sensitive may not be a good feature under practical (economical) concerns. However, giving up opportunities for process improvement may cost a lot more in the long run.

The methods we propose may have another application for KGD vendors in pricing KGDs. KGD vendors could grade their KGDs with our methods by setting different levels of the baseline area. A wafer that passes the test with smaller baseline area could be graded higher. For example, set three baseline levels. Start with the most stringent level (smallest baseline area), if the wafer passes the test with no apparent patterns, grade it as A^+ ; otherwise, test it with the second level. If it passes, then grade it is as A ; otherwise test it with the third level. If it passes, grade it A^- ; otherwise, classify it as “patterned”.

In this study, we assume the baseline is pre-determined. In practice, how to determine the baseline from historical data or engineers’ knowledge/experiences is an important issue. It would be useful if we can automatically generate probable baselines from historical data for engineers. This could be a potential future research topic. Another topic could be developing techniques for automatic pattern classifications under the consideration of the baseline.



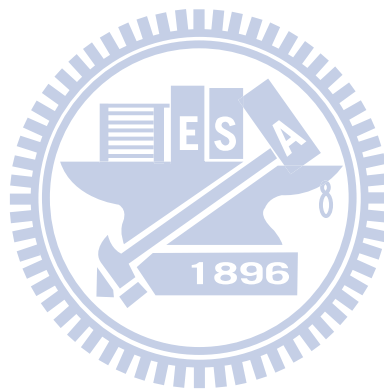
References

- [1] S. P. Cunningham and S. McKinnon, "Statistical Methods for Visual Defect Metrology," *IEEE Transactions on Semiconductor Manufacturing*, vol. 11, no. 1, pp. 48-53, 1998.
- [2] C. K. Hansen and P. Thyregod, "Use of Wafer Maps in Integrated Circuit Manufacturing," *Microelectronics Reliability*, vol. 38, pp. 1155-1164, 1998.
- [3] M. H. Hansen, V. N. Nair, and D. J. Friedman, "Monitoring Wafer Map Data from Integrated Circuit Fabrication Processes for Spatially Clustered Defects," *Technometrics*, vol. 39, no. 3, pp. 241-253, 1997.
- [4] W. Taam and M. Hamada, "Detecting Spatial Effects from Factorial Experiment: An Application from IC Manufacturing," *Technometrics*, vol. 35, no. 2, pp. 149-160, 1993.
- [5] R. Uzsoy, C. Y. Lee, and L. A. Martin-Vega, "A Review of Production Planning and Scheduling Models in the Semiconductor Industry, Part I," *IIE Transactions*, vol. 24, no. 4, pp. 47-60, 1992.
- [6] K. Knutson, K. Kempf, J. Fowler, and M. Carlyle, "Lot-to-Order Matching for a Semiconductor Assembly and Test Facility," *IIE Transactions*, vol. 31, no. 11, pp. 1103-1111, 1999.
- [7] G. S. May and C. J. Spanos, *Fundamentals of Semiconductor Manufacturing and Process Control*. New York: Wiley, 2006.
- [8] C. M. Scanlan, and N. Karim, "System-In-Package Technology, Application and Trends," *Amkor Technology, Inc.*
- [9] M. S. Lin, "SiP - System in Package," *MEGIC, Inc.*, 2001.
- [10] R. Buck, "RF System-in-Package Competes with SoCs," *Analog Devices, Inc.*, 2002.
- [11] M. Aguirre, "Super High Density Packaging Technologies - Known Good Die Workshop," *Fujitsu Microelectronics America, Inc.*, 2002.

- [12] U. Kaempf, "The Binomial Test: A Simple Tool to Identify Process Problems," *IEEE Transactions on Semiconductor Manufacturing*, vol. 8, no. 2, pp. 160-166, 1995.
- [13] J. Y. C. Pan and J. M. Tenenbaum, "P.I.E.S: An Engineer's 'Do-It-Yourself' Knowledge System for Interpretation of Parametric Test Data," *Proceedings of Artificial Intelligence*, 1986, pp. 836-843.
- [14] C. Mary, "Artificial Intelligence in Semiconductor Manufacturing for Process Development, Functional Diagnostics, and Yield Crash Prevention," *Proceedings of International Test Conference*, 1986, pp. 939-946.
- [15] W. Maly, B. Trifilo, R. A. Hughes, and A. Miller, "Yield Diagnosis through Interpretation of Tester Data," *Proceedings of International Test Conference*, 1987, pp. 10-20.
- [16] L. Breiman, J. Friedman, A. R. Olshen, and C. Stone. *Classification and Regression Trees*. New York: Chapman and Hall, 1993.
- [17] V. Raghavan, "Application of Decision Trees for Integrated Circuit Yield Improvement," *Proceedings of Advanced Semiconductor Manufacturing Conference*, 2002, pp. 262-265.
- [18] J. A. Cunningham, "The Use and Evaluation of Yield Models in Integrated Circuit Manufacturing," *IEEE Transactions on Semiconductor Manufacturing*, vol. 3, no. 2, pp. 60-71, 1990.
- [19] S. S. Gleason, K. W. Tobin, T. P. Karnowski, and F. Lakhani, "Rapid Yield Learning through Optical Defect and Electrical Test Analysis," *Proceedings of SPIE-The International Society for Optical Engineering*, 1998, pp. 232-242.
- [20] F. L. Chen and S. F. Liu, "A Neural-Network Approach to Recognize Defect Spatial Pattern in Semiconductor Fabrication." *IEEE Transactions on Semiconductor Manufacturing*, vol. 13, no. 3, pp. 366-373, 2000.
- [21] S. F. Liu, F. L. Chen, and W. B. Lu, "Wafer Bin Map Recognition Using a Neural Network Approach," *International Journal of Production Research*, vol. 40, pp. 2207-2223, 2002.

- [22] J. H. Lee, S. J. Yu, and S. C. Park, "Design of Intelligent Data Sampling Methodology Based on Data Mining," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 5, pp. 637-649, 2001.
- [23] L. C. Chao, and L. I. Tong, "Wafer Defect Pattern Recognition by Multi-Class Support Vector Machines by Using a Novel Defect Cluster Index," *Expert Systems with Applications*, vol. 36, no. 6, pp. 10158-10167, 2009.
- [24] J. S. Weszka, C. R. Dyer, and A. Rosenfeld, "A Comparative Study of Texture Measures for Terrain Classification," *IEEE Transactions on Systems Man Cybernet*, vol. 6, no. 4, pp. 269-285, 1976.
- [25] A. Agresti, *Categorical Data Analysis*. New York: Wiley, 1990.
- [26] A. D. Cliff and J. K. Ord, *Spatial Processes: Models and Applications*. New York: Pion, 1981.
- [27] A. J. Lee, *U-Statistics, Theory and Practice*. New York: Marcel Dekker, 1990.
- [28] M. P.-L. Ooi, Y. C. Kuang, W. J. Tee, A. A. Mohanan, and C. Chan, "Accurate Defect Cluster Detection and Localisation on Fabricated Semiconductor Wafers using Joint Count Statistics," *ASQED*, 2009, pp. 225-232.
- [29] E. Parzen, "On Estimation of a Probability Density Function and Mode," *The Annals of Mathematical Statistics*, vol. 33, pp. 1065-1076, 1962.
- [30] H. Chernoff, "Estimation of the Mode", *Annals of the Institute of Statistical Mathematics*, vol. 16, pp. 31-41, 1964.
- [31] W. F. Eddy, "Optimum Kernel Estimators of the Mode", *The Annals of Statistics*, vol. 8, pp. 870-882, 1980.
- [32] J. P. Romano, "On Weak Convergence and Optimality of Kernel Density Estimates of the Mode", *The Annals of Statistics*, vol. 16, pp. 629-647, 1988.
- [33] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall, 1986.

- [34] D. Cheng, "Controllability of Switched Bilinear Systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 4, 511-515, 2005.
- [35] S. J. Shi and W. Zhang, "Asymptotic Theory of the Error Distribution of Nonparameter Regression Models," *Journal of Sichuan University*, vol. 32, no. 1, pp. 16-22, 1995.
- [36] T. Gasser and H. G. Muller, "Nonparametric Estimation of Regression Functions and Their Derivatives by the Kernel Method," *Scandinavian Journal of Statistics*, vol. 25, no. 11, 171-175, 2002.



Appendices

A Table and WSMs Related to Simulation Studies

Table 2: Counts of false alarm's of the two existing methods and three proposed methods among 150 random WSMs under various yield levels.

Case	Yield	Odds Ratio Test	HNF	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.01	5	2	4	1	2
2	0.06	3	3	1	1	1
3	0.11	1	4	1	2	0
4	0.16	4	2	1	0	1
5	0.21	2	4	0	2	1
6	0.26	3	4	0	3	0
7	0.31	4	2	0	1	0
8	0.36	6	2	1	1	1
9	0.41	2	1	0	0	0
10	0.46	4	3	0	1	0
11	0.51	10	8	0	3	0
12	0.56	16	10	1	2	2
13	0.61	27	16	0	1	1
14	0.66	46	33	0	2	1
15	0.71	66	53	0	2	0
16	0.76	101	78	1	3	0
17	0.81	119	109	1	2	1
18	0.86	145	143	1	0	0
19	0.91	150	150	0	0	0
20	0.95	150	150	0	3	0

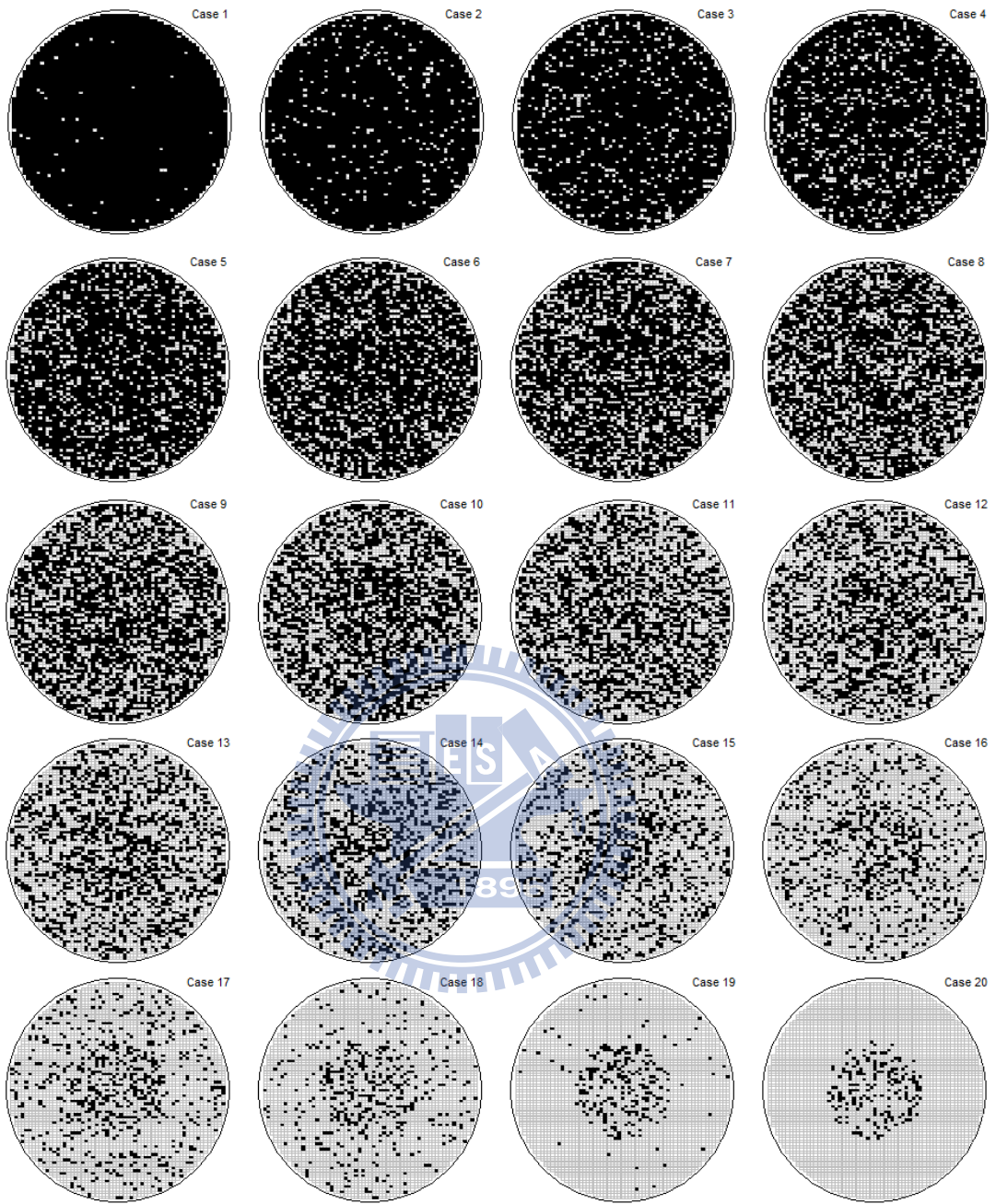


Table 3: Counts of detection among 150 WSMs with the linear scratch pattern under various yield.

Case	Yield	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.01	3	2	3
2	0.05	1	1	2
3	0.1	1	4	0
4	0.14	3	2	3
5	0.19	3	5	3
6	0.23	15	10	11
7	0.28	17	17	11
8	0.32	29	23	25
9	0.37	50	36	36
10	0.41	86	65	63
11	0.46	113	91	92
12	0.5	136	126	126
13	0.55	147	143	141
14	0.59	150	150	150
15	0.64	150	150	150
16	0.68	150	150	150
17	0.73	150	150	150
18	0.77	150	150	150
19	0.82	150	150	150
20	0.86	150	150	150

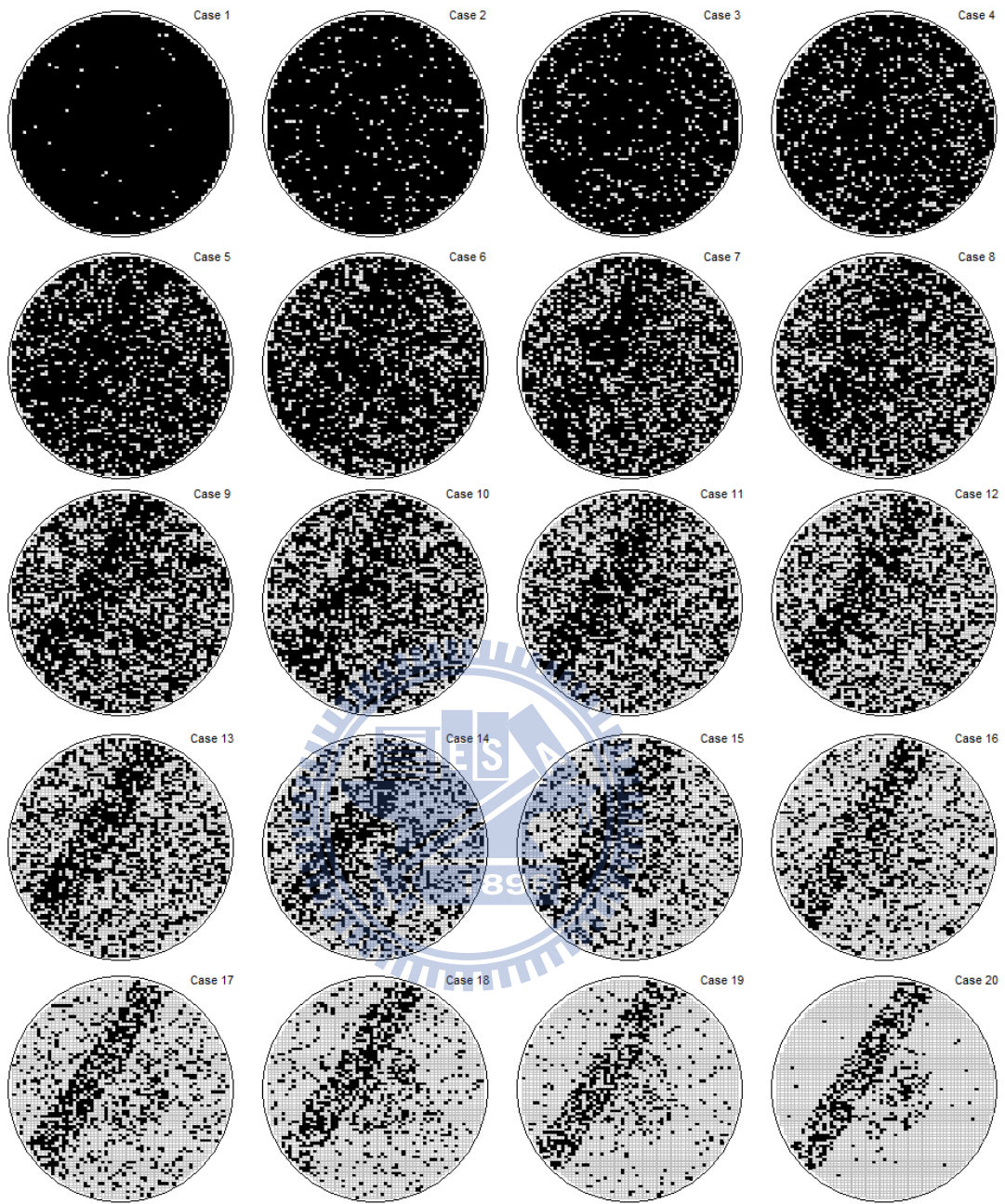


Table 4: Counts of detection among 150 WSMs with the ring pattern under various yield.

Case	Yield	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.01	4	1	2
2	0.05	1	0	0
3	0.1	3	2	2
4	0.14	3	1	5
5	0.19	10	6	7
6	0.23	13	10	12
7	0.27	38	26	26
8	0.32	61	56	50
9	0.36	90	81	78
10	0.41	121	118	113
11	0.45	139	135	131
12	0.5	148	146	146
13	0.54	150	150	150
14	0.58	150	150	150
15	0.63	150	150	150
16	0.67	150	150	150
17	0.72	150	150	150
18	0.76	150	150	150
19	0.81	150	150	150
20	0.85	150	150	150

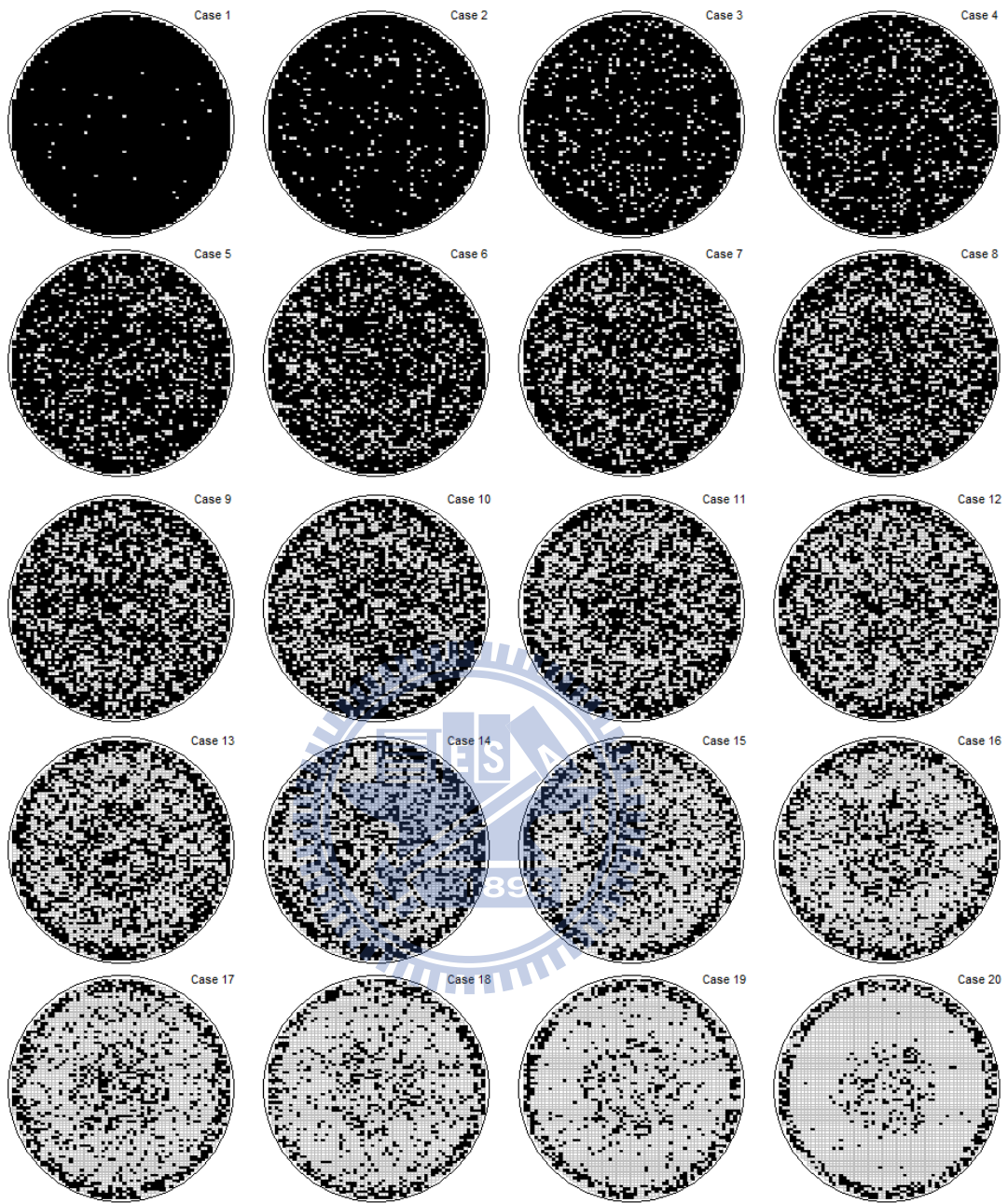


Table 5: Counts of detection among 150 WSMS with the bottom pattern under various yield.

Case	Yield	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.01	3	1	2
2	0.05	3	6	1
3	0.1	7	5	3
4	0.14	7	8	6
5	0.19	15	7	11
6	0.23	28	24	19
7	0.27	45	32	27
8	0.32	91	75	68
9	0.36	112	88	92
10	0.41	144	133	129
11	0.45	150	143	145
12	0.5	150	150	150
13	0.54	150	150	150
14	0.58	150	150	150
15	0.63	150	150	150
16	0.67	150	150	150
17	0.72	150	150	150
18	0.76	150	150	150
19	0.81	150	150	150
20	0.85	150	150	150

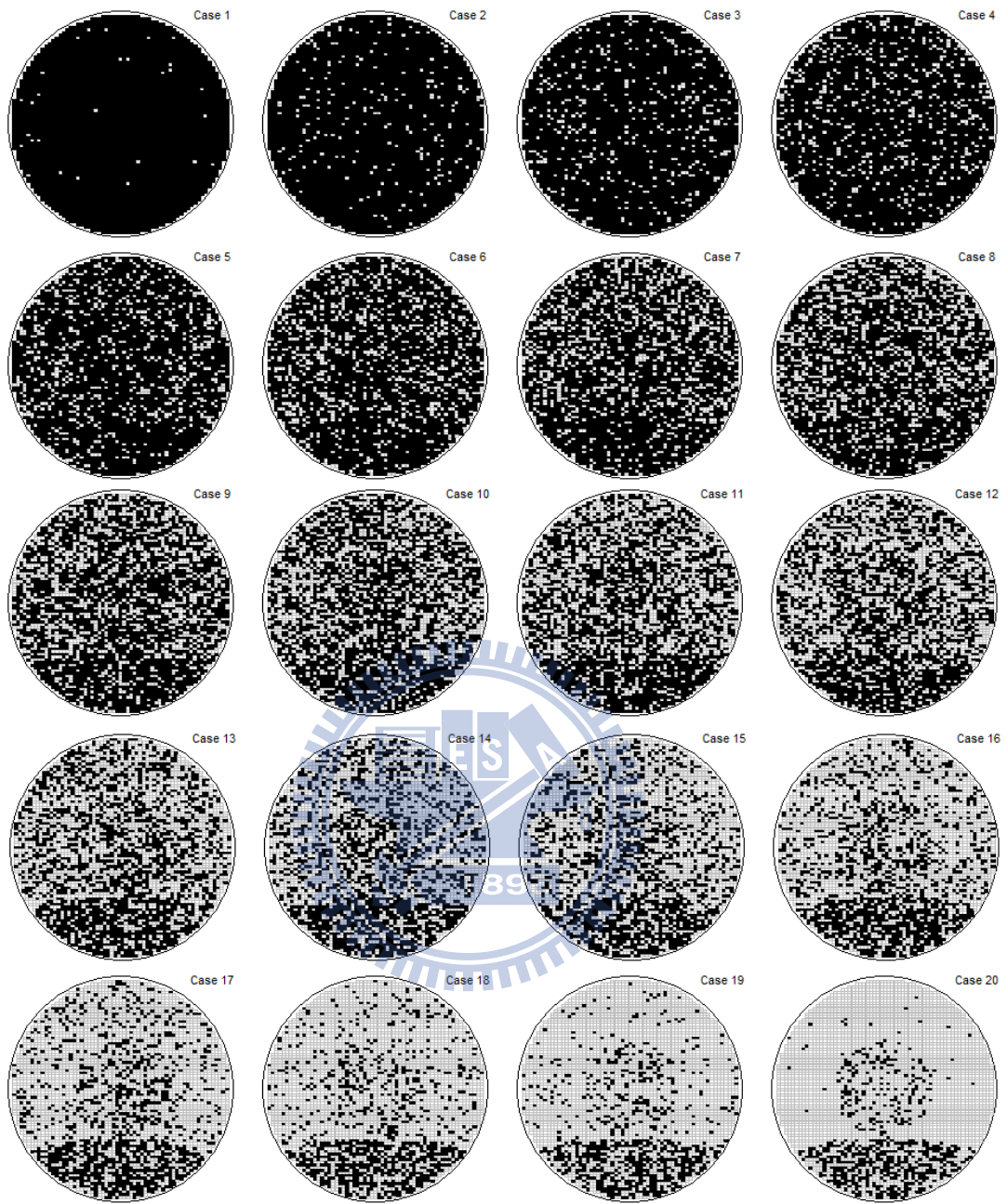
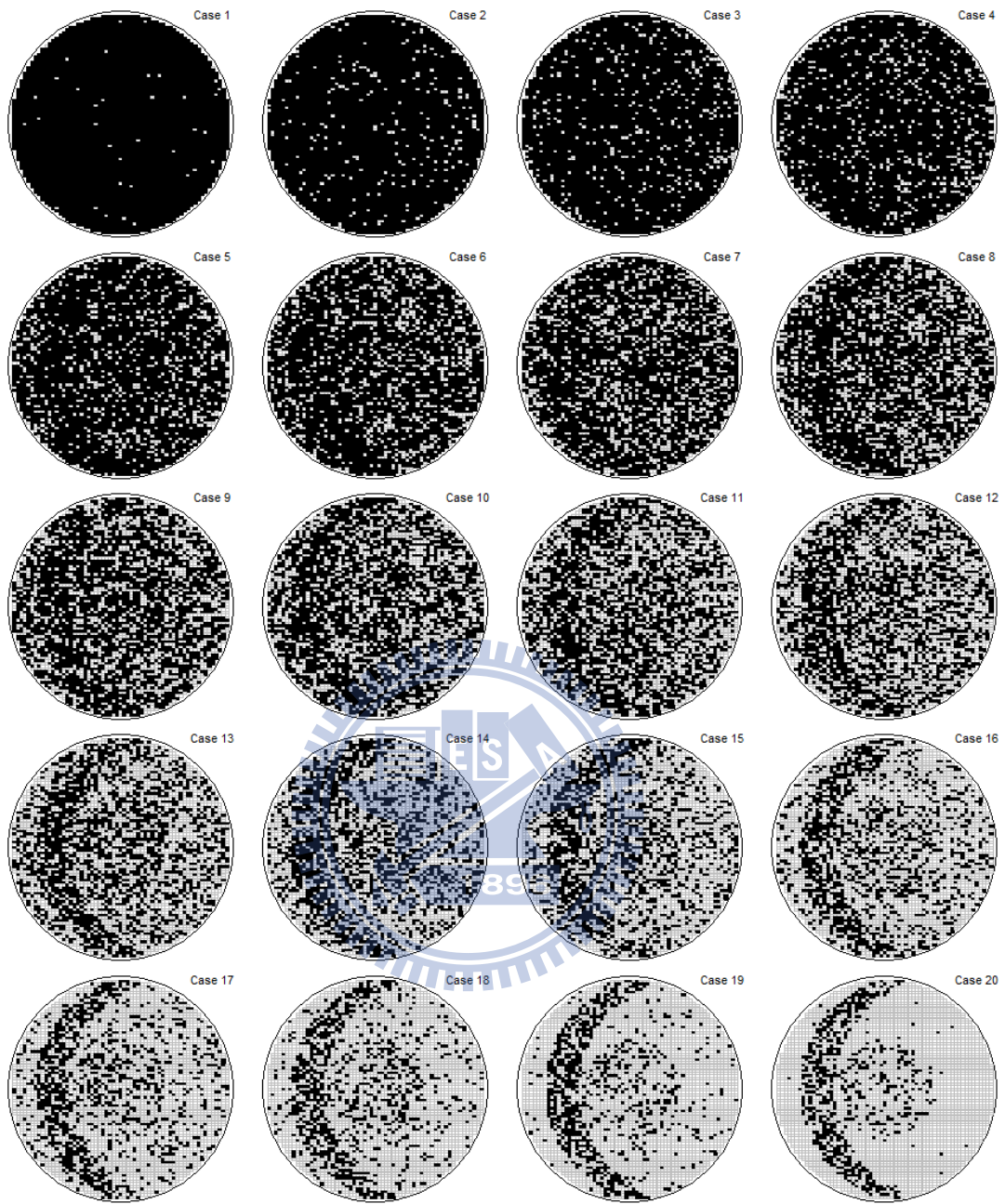


Table 6: Counts of detection among 150 WSMs with the moon pattern under various yield.

Case	Yield	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.01	4	1	1
2	0.05	4	5	4
3	0.1	2	1	2
4	0.14	2	2	3
5	0.19	9	6	7
6	0.23	28	21	18
7	0.27	40	32	32
8	0.32	69	54	50
9	0.36	107	89	86
10	0.41	132	112	112
11	0.45	148	139	138
12	0.5	149	146	147
13	0.54	150	150	150
14	0.58	150	150	150
15	0.63	150	150	150
16	0.67	150	150	150
17	0.72	150	150	150
18	0.76	150	150	150
19	0.81	150	150	150
20	0.85	150	150	150



- Middle Yield

Table 7: Counts of detection among 150 WSMs with the linear scratch pattern under various patterned areas.

Case	Patterned Area (%)	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.03	0	0	0
2	1.07	1	0	0
3	1.97	1	2	1
4	3.04	8	6	10
5	4.11	26	12	14
6	5.06	43	32	42
7	6.13	90	78	75
8	7.23	116	87	97
9	8.36	134	120	119
10	9.3	138	131	135
11	10.43	148	145	144
12	11.53	147	143	142
13	12.52	146	145	145
14	13.65	150	150	150
15	14.75	149	149	150
16	15.9	150	150	150
17	16.89	150	150	150
18	18.04	150	150	150
19	19.17	150	150	150
20	20.21	150	150	150

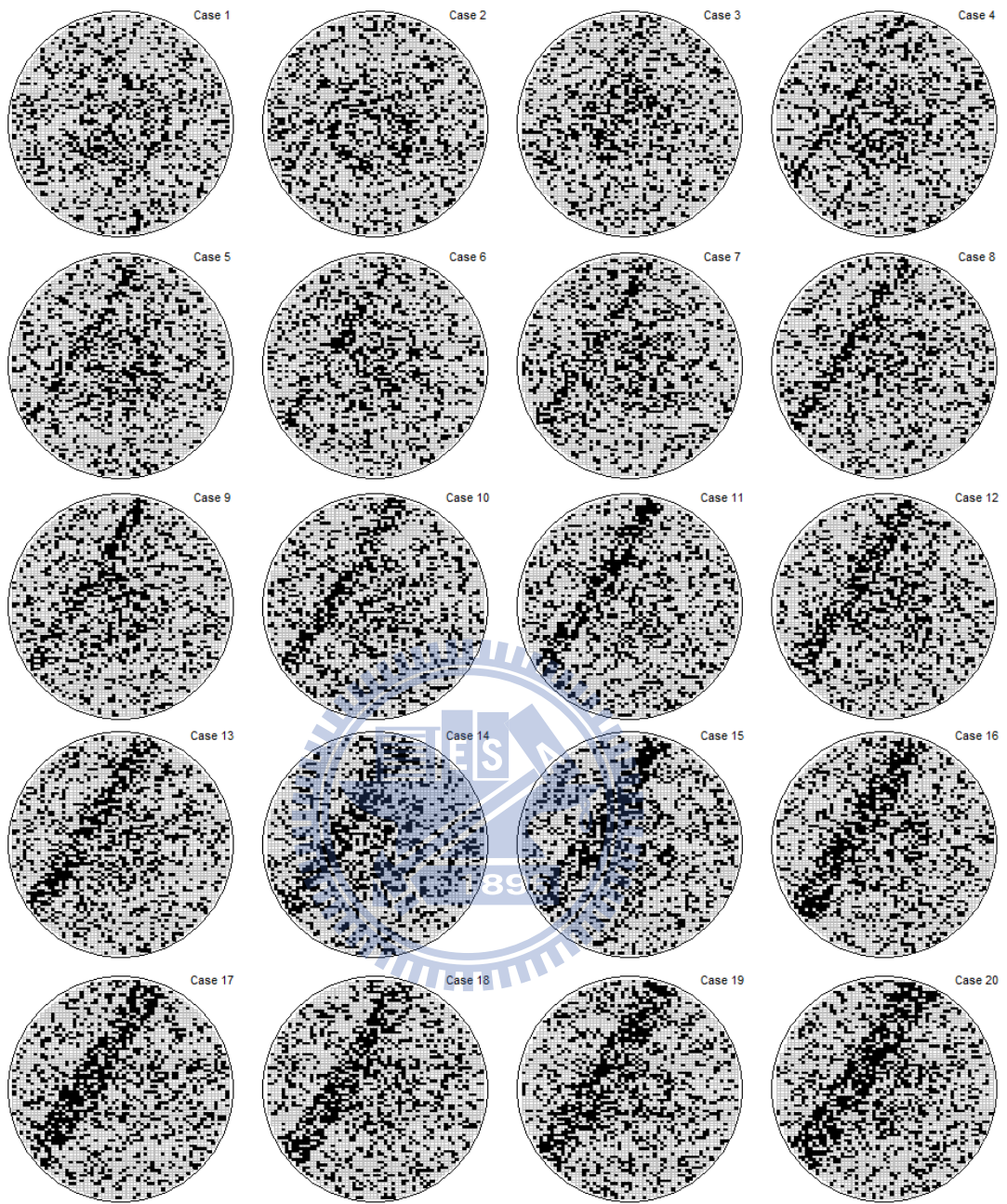


Table 8: Counts of detection among 150 WSMs with the ring pattern under various patterned areas.

Case	Patterned Area (%)	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.1	1	2	3
2	0.61	0	16	0
3	2.04	0	147	0
4	3.02	1	150	2
5	4.24	3	150	8
6	5.42	6	150	22
7	6.54	41	150	60
8	7.92	92	149	116
9	8.54	107	150	126
10	9.76	137	150	149
11	11.19	148	150	148
12	12.22	150	150	150
13	13.14	150	150	150
14	13.85	150	150	150
15	15.18	150	150	150
16	16	150	150	150
17	17.63	150	150	150
18	18.76	150	150	150
19	19.37	150	150	150
20	20.19	150	150	150

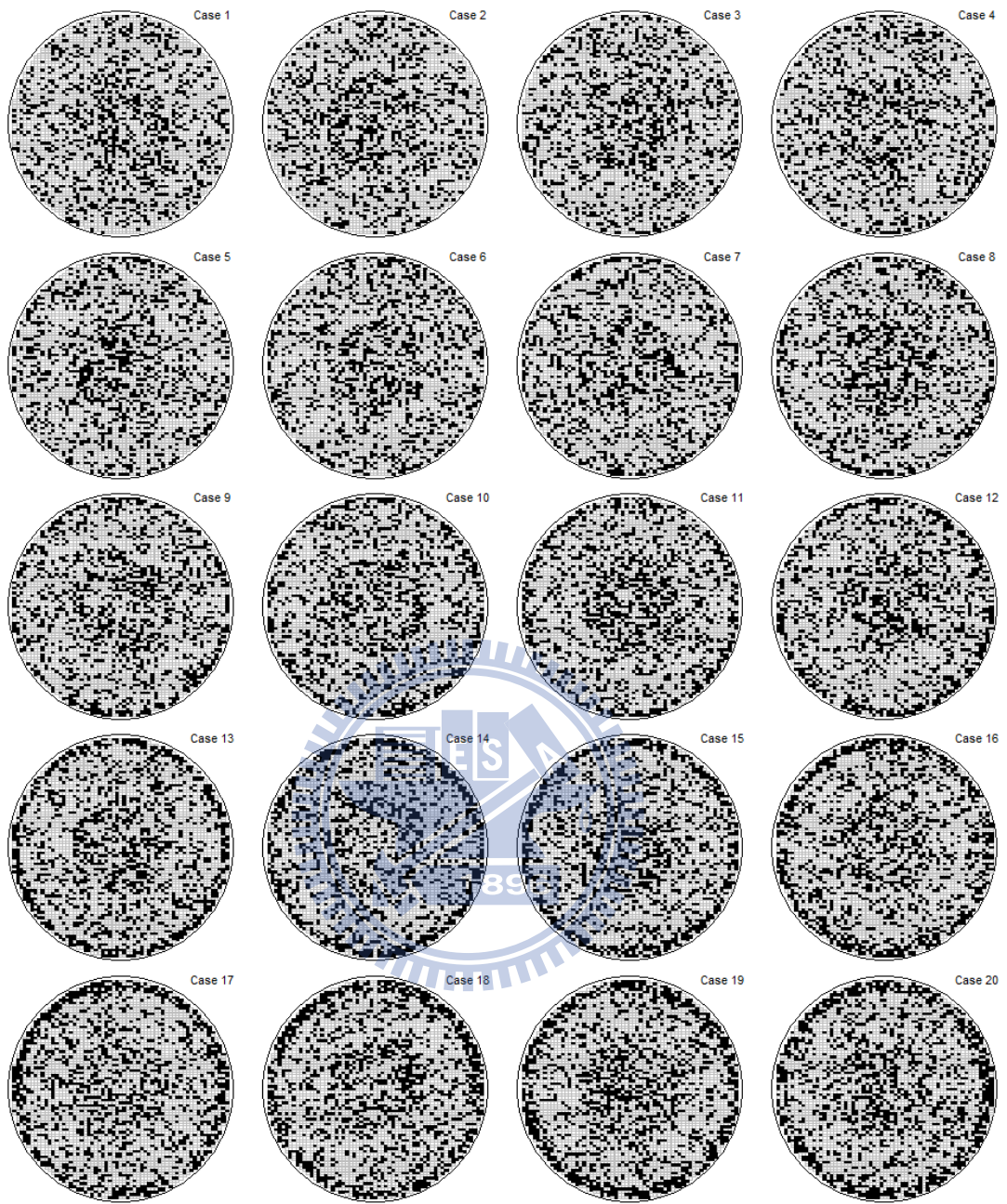


Table 9: Counts of detection among 150 WSMs with the bottom pattern under various defect areas.

Case	Patterned Area (%)	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.28	0	2	0
2	0.66	0	0	2
3	1.18	6	4	4
4	1.87	21	19	16
5	2.61	40	38	33
6	3.4	66	56	59
7	4.29	100	90	90
8	5.24	130	116	119
9	6.18	142	140	135
10	7.18	145	144	143
11	8.23	150	150	149
12	9.38	149	149	149
13	10.66	149	150	149
14	11.96	150	150	150
15	13.37	150	150	150
16	14.62	150	150	150
17	15.97	150	150	150
18	17.43	150	150	150
19	18.89	150	150	150
20	20.29	150	150	150

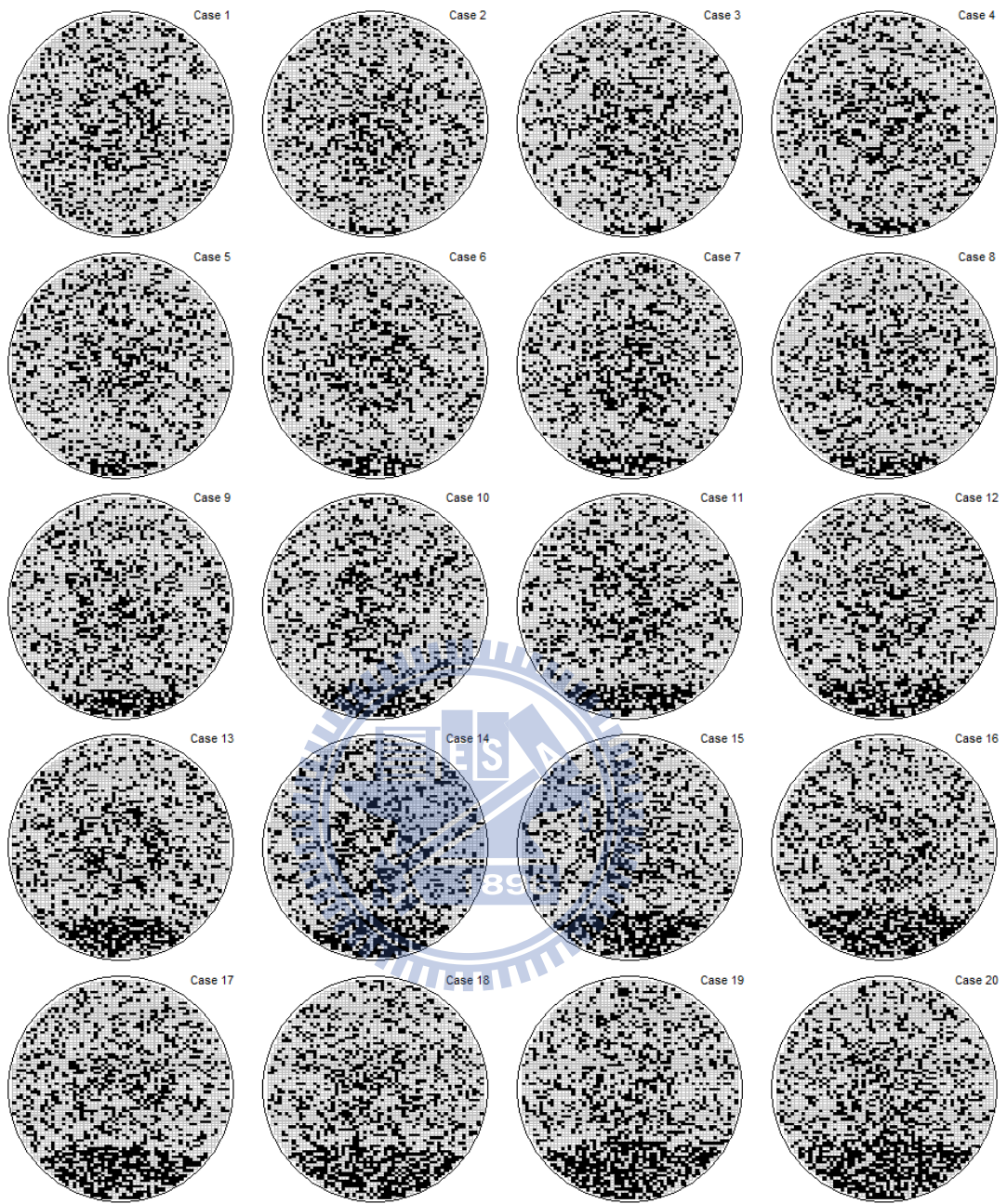
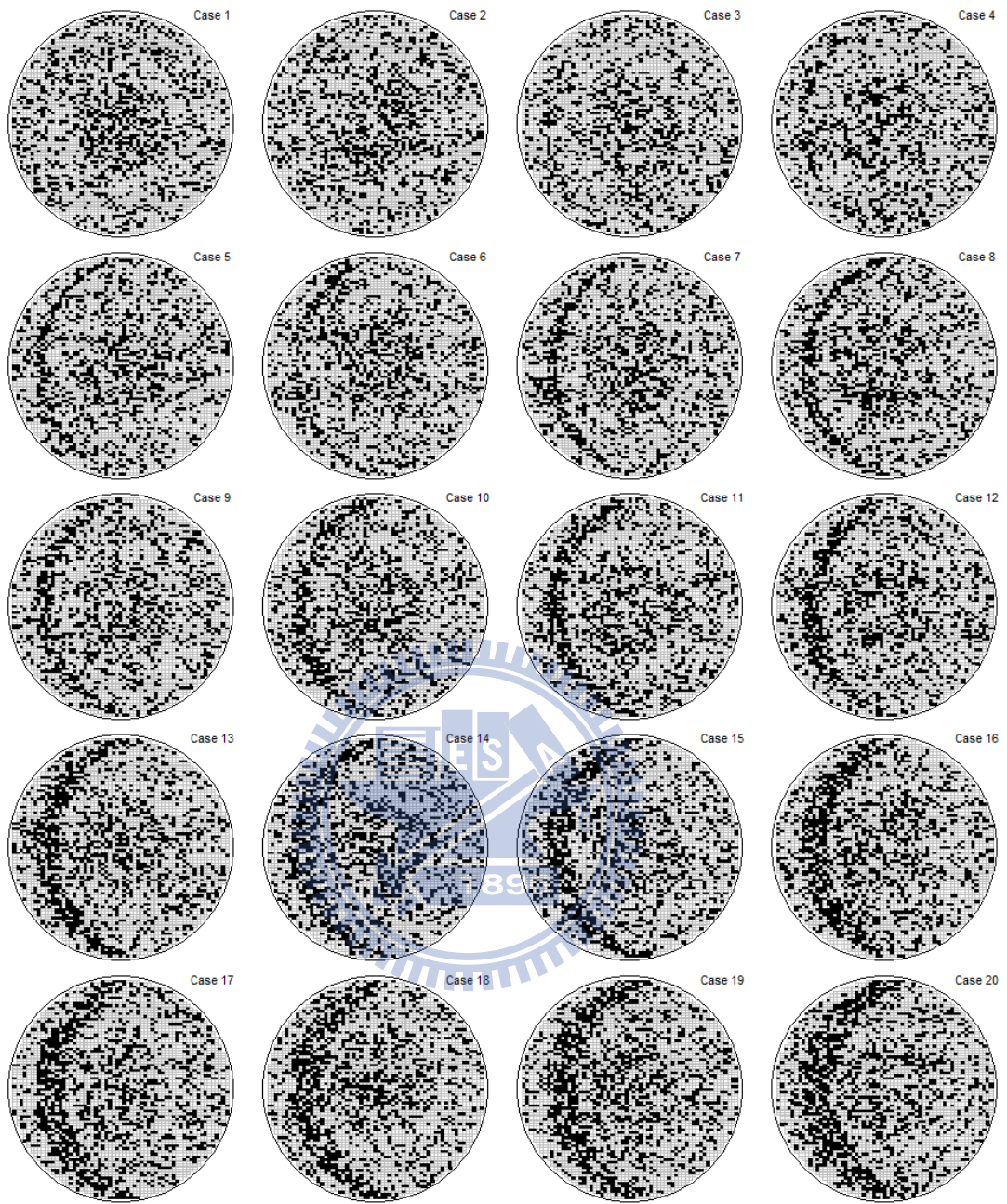


Table 10: Counts of detection among 150 WSMs with the moon pattern under various patterned areas.

Case	Patterned Area (%)	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.59	0	0	0
2	1.71	1	0	0
3	2.66	4	2	3
4	3.78	25	11	14
5	4.73	49	39	30
6	5.8	99	80	75
7	6.8	125	109	105
8	7.87	143	134	128
9	8.87	146	143	141
10	9.84	149	149	149
11	10.94	150	150	148
12	11.91	150	150	149
13	13.01	150	150	150
14	13.98	150	150	150
15	15.08	150	150	150
16	16	150	150	150
17	17.1	150	150	150
18	18.02	150	150	150
19	19.12	150	150	150
20	20.04	150	150	150



- High Yield

Table 11: Counts of detection among 150 WSMs with the linear scratch pattern under various patterned areas.

Case	Patterned Area (%)	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.03	1	2	0
2	1.07	1	3	3
3	1.97	86	76	66
4	3.04	149	148	150
5	4.11	150	150	150
6	5.06	150	150	150
7	6.13	150	150	150
8	7.23	150	150	150
9	8.36	150	150	150
10	9.3	150	150	150
11	10.43	150	150	150
12	11.53	150	150	150
13	12.52	150	150	150
14	13.65	150	150	150
15	14.75	150	150	150
16	15.9	150	150	150
17	16.89	150	150	150
18	18.04	150	150	150
19	19.17	150	150	150
20	20.21	150	150	150

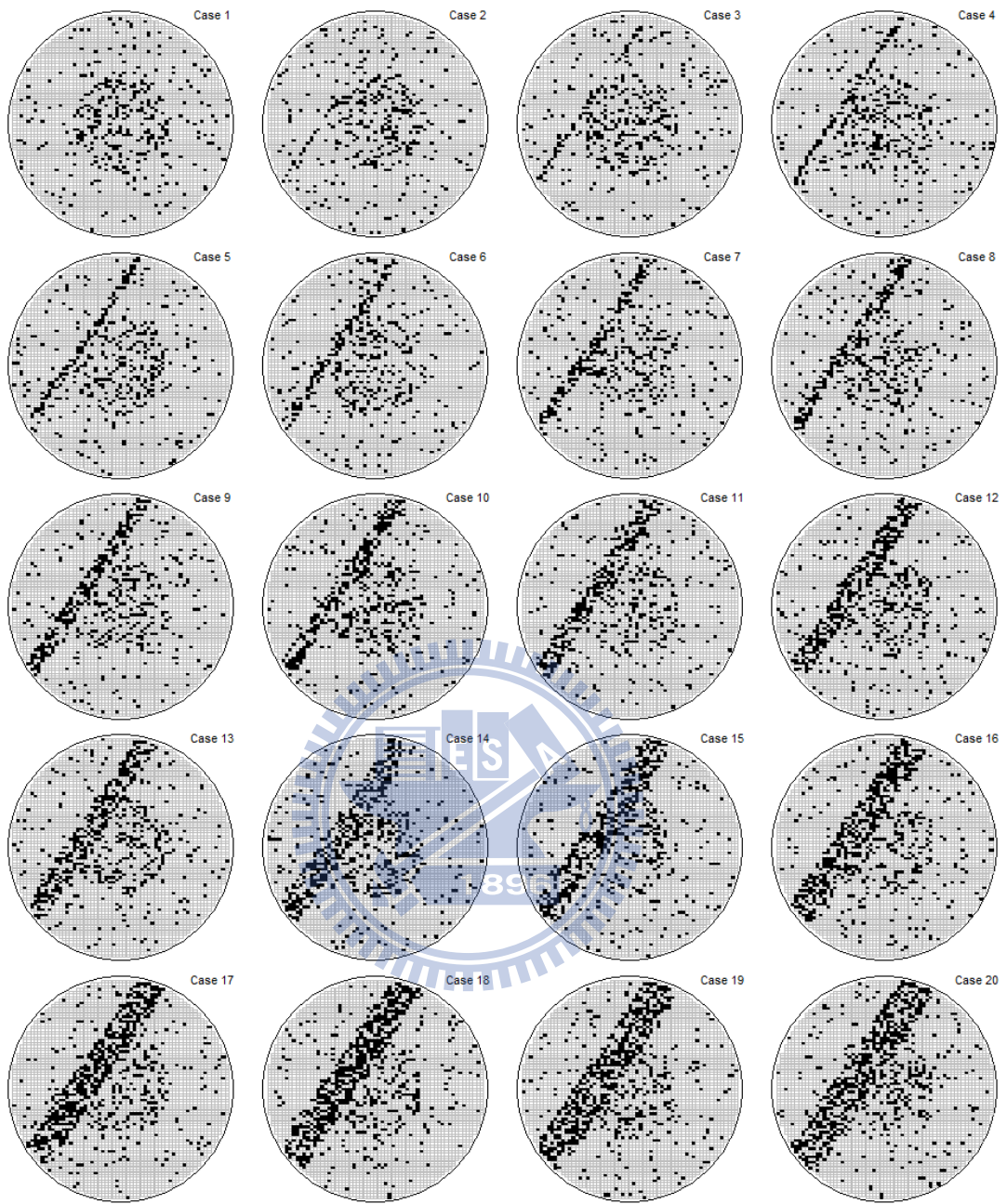


Table 12: Counts of detection among 150 WSMs with the ring pattern under various patterned areas.

Case	Patterned Area (%)	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.1	1	9	0
2	0.61	1	130	1
3	2.04	7	150	18
4	3.02	47	150	81
5	4.24	138	150	149
6	5.42	149	150	150
7	6.54	150	150	150
8	7.92	150	150	150
9	8.54	150	150	150
10	9.76	150	150	150
11	11.19	150	150	150
12	12.22	150	150	150
13	13.14	150	150	150
14	13.85	150	150	150
15	15.18	150	150	150
16	16	150	150	150
17	17.63	150	150	150
18	18.76	150	150	150
19	19.37	150	150	150
20	20.19	150	150	150

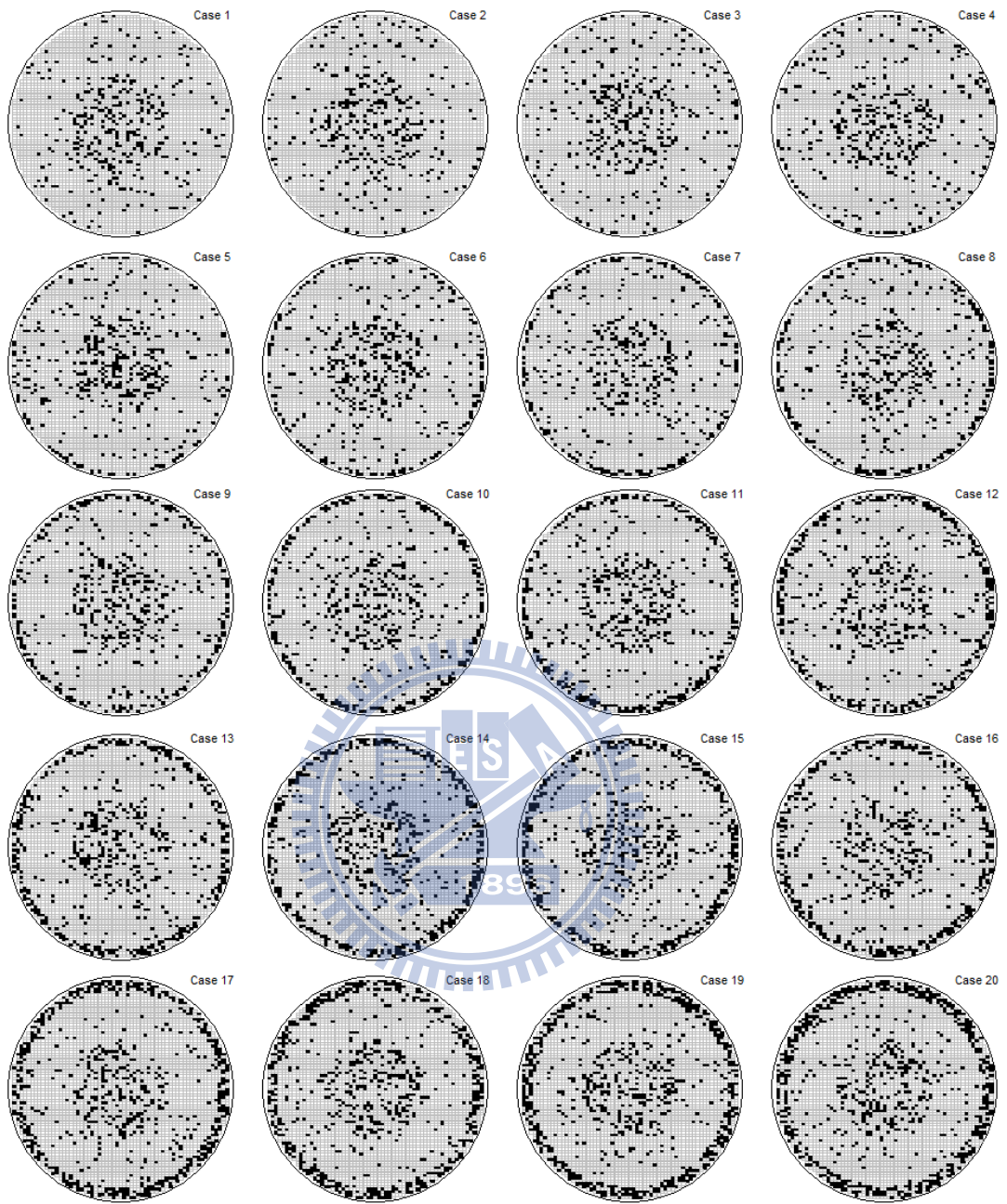


Table 13: Counts of detection among 150 WSMs with the bottom pattern under various patterned areas.

Case	Patterned Area (%)	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.28	1	24	2
2	0.66	53	75	62
3	1.18	133	138	129
4	1.87	149	149	149
5	2.61	150	150	150
6	3.4	150	150	150
7	4.29	150	150	150
8	5.24	150	150	150
9	6.18	150	150	150
10	7.18	150	150	150
11	8.23	150	150	150
12	9.38	150	150	150
13	10.66	150	150	150
14	11.96	150	150	150
15	13.37	150	150	150
16	14.62	150	150	150
17	15.97	150	150	150
18	17.43	150	150	150
19	18.89	150	150	150
20	20.29	150	150	150

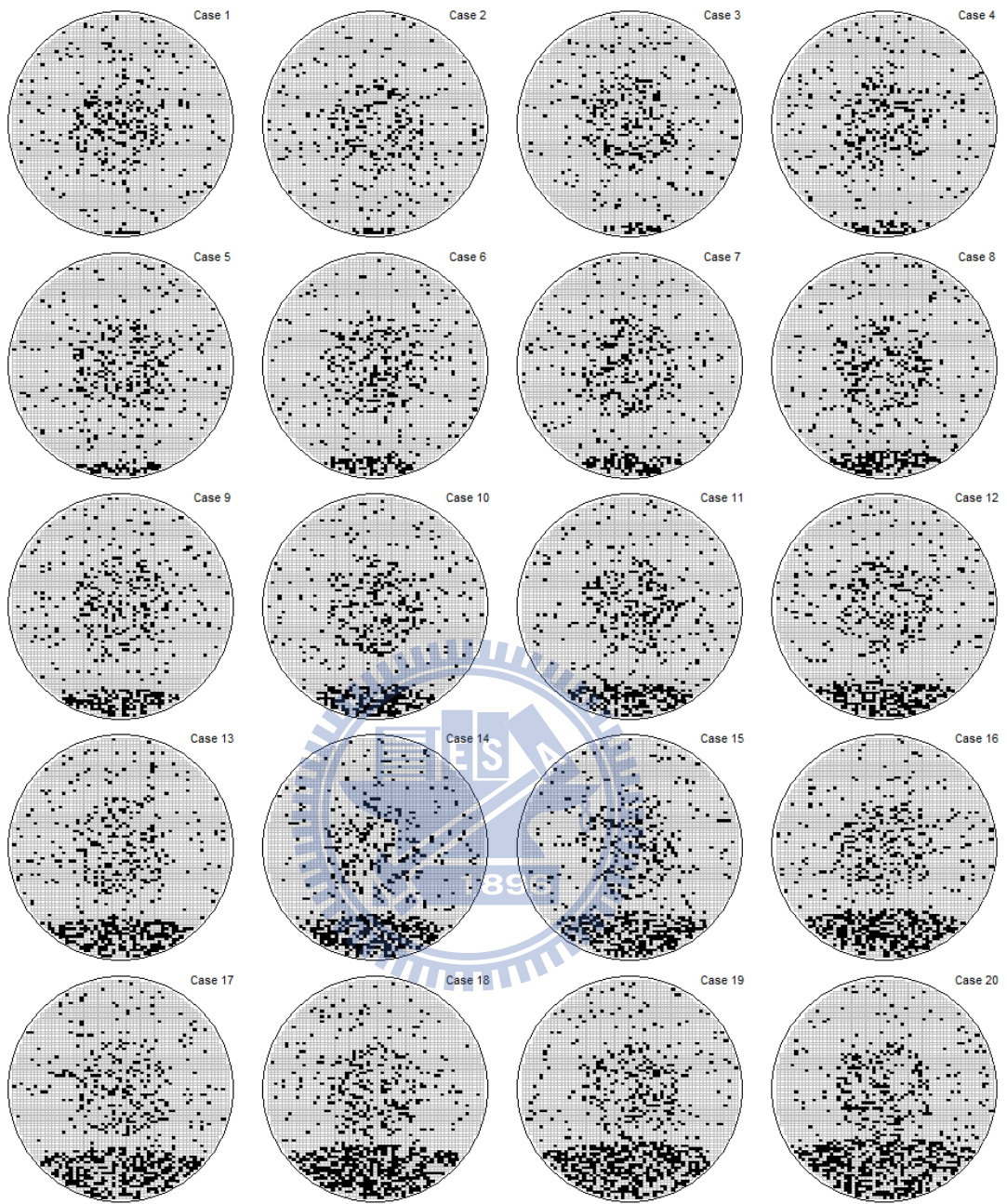


Table 14: Counts of detection among 150 WSMs with the moon pattern under various patterned areas.

Case	Patterned Area (%)	Mode Odds Ratio Test	Mode HNF	Empirical HNF
1	0.59	1	3	0
2	1.71	30	21	18
3	2.66	115	90	80
4	3.78	150	150	149
5	4.73	150	150	150
6	5.8	150	150	150
7	6.8	150	150	150
8	7.87	150	150	150
9	8.87	150	150	150
10	9.84	150	150	150
11	10.94	150	150	150
12	11.91	150	150	150
13	13.01	150	150	150
14	13.98	150	150	150
15	15.08	150	150	150
16	16	150	150	150
17	17.1	150	150	150
18	18.02	150	150	150
19	19.12	150	150	150
20	20.04	150	150	150

