1 YIELD MODELS FOR DEFECT-TOLERANT VLSI CIRCUITS: A REVIEW

From: DEFECT AND FAULT TOLERANCE IN VLSI SYSTEMS
Edited by Israel Koren
(Plenum Publishing Corporation, 1989)

YIELD MODELS FOR DEFECT-TOLERANT VLSI CIRCUITS: A REVIEW

Israel Koren and **Charles H. Stapper***

University of Massachusetts at Amherst
Department of Electrical and Computer Engineering
Amherst, MA 01003

* **IBM General Technology Division**
Essex Junction, VT 05452

ABSTRACT

The statistical models for estimating and predicting the manufacturing yields of VLSI circuits are reviewed. It is shown how defect clustering is taken into account, and how yield formulae for defect and fault tolerant VLSI circuits are developed. Different types of formulae for the yield of defect tolerant VLSI circuits have appeared in the literature. It is proven here for the first time that most of these approaches are equivalent.

INTRODUCTION

The designation *defect-tolerant* (or *fault-tolerant*) is often used in connection with integrated circuits that have some degree of tolerance to manufacturing flaws or defects. Such circuits are capable of functioning correctly when they contain certain types of faults. However, not all faults can be dealt with by the fault-tolerance schemes. As a result these circuits do not have 100% fabrication yields and the prediction of those yields is difficult. This difficulty is illustrated in at least three doctoral dissertations dealing with this subject, namely those of Mangir[1], Hedlund[2], and Harden[3].

The difficulty in modeling the yield of fault-tolerant integrated circuit chips is caused mainly by the clustering of manufacturing defects during chip fabrication. Several schemes for deriving expressions for the yield of fault-tolerant circuits in the presence of clustering have been proposed in different publications. We review these schemes and show that most of them are equivalent and result in the exact same value for the yield of such chips. We also present a simple way of handling the complexity of deriving yield expressions when defects are clustered.

In the next section we briefly review the most commonly used distributions for faults, namely, the Poisson distribution and the negative binomial distribution and show the resulting expressions for yield. We then present a method for determining the parameters of the negative binomial distribution and discuss the effect of the size of clusters relative to the size of chips. Next, the previously proposed models for the yield of fault tolerant chips are reviewed and the equivalence between them is proved. We also present a general method for deriving a yield expression for clustered faults from the simpler expression obtained when faults are assumed to be evenly distributed, i.e., follow the Poisson distribution. These results for chips with redundancy are then extended to partially good chips and multiple module-type chips. Final conclusions are presented in the last section.

YIELD MODELS

The earliest yield model published in the archival literature was the one used by Wallmark[4] in 1960. He expressed the probability of transistor failure as a ratio $S/100$, where S represents the number of failing transistors occurring in a batch of 100. For M transistors, the yield, denoted by Y_M, consequently becomes

$$Y_M = (1 - S/100)^M. \tag{1}$$

Wallmark used this result in a binomial distribution for estimating the yield of integrated circuits with redundant transistors. For obtaining exactly M good transistors out of a total of N this took the form

$$Y = \binom{N}{M} (S/100)^{N-M}(1 - S/100)^M. \tag{2}$$

This was the first step towards a yield model for integrated circuits with fault-tolerance. Several years later Hofstein and Heiman[5] also tackled the problem of manufacturing more circuits than were actually needed. Their chips contained field effect transistors, and they claimed that the transistor failures were predominately caused by pinholes in the oxide gates. Assuming a Poisson distribution for the number of defects per gate, they obtained a yield model for M transistors with the formula

$$Y_M = e^{-MA_GD} \tag{3}$$

where A_G is the gate area in each transistor and D an average defect density of pinholes per unit area. This expression was then used in a binomial formula for calculating the probability of having at least M good transistors on a chip with N transistors on it,

$$Y = \sum_{i=M}^{N} \binom{N}{i} e^{-iA_GD}(1 - e^{-A_GD})^{N-i}. \tag{4}$$

This was therefore, the first model for the yield of integrated circuits with fault-tolerance. Subsequently, Poisson statistics were commonly used for modeling the distribution of the number of faults per chip. According to this distribution the probability of having exactly x faults in a chip is given by,

$$Prob\ \{X = x\} = \frac{e^{-\lambda}\ \lambda^x}{x!} \tag{5}$$

where X is a random variable denoting the number of faults and λ is the average number of faults expected per chip. A fault is defined as a specific chip failure that can be caused by one or more manufacturing defects.

For chips with no redundancy the yield is therefore,

$$Y = Prob\ \{X = 0\} = e^{-\lambda}. \tag{6}$$

The average number of faults per chip is often expressed as $\lambda = AD$, the chip area A times a fault density D. This, however, is a simplification. The relationship between the average number of faults per chip and the chip area is more complicated; it depends on the circuit complexity, the density of photolithographic patterns, the number of photolithographic masks used in the process, etc. These however, are beyond the scope of this review.

It has been known since the beginning of integrated-circuit manufacture that the above yield formula is too pessimistic and leads to predicted chip yields that are too low when extrapolated from the yield of smaller chips or single circuits. It later became clear that the very low predicted yield was caused by the clustering of faults. This phenomenon has been observed in practice, but was not taken into account in equations (5) and (6).

Several modifications to the above yield formula have been proposed to account for fault clustering. The most commonly used modification is obtained by assuming the number of faults to be Poisson distributed as in (5) but to consider the parameter λ to be a random variable rather than a constant. The mere fact that λ is a random variable results in clustering of faults, no matter what type of distribution is assumed for λ.

The modified yield formula is then obtained by averaging yield formula (6) with respect to λ. Let $F(\lambda)$ be a cumulative distribution function of the average number of faults per chip. Associated with the cumulative distribution function $F(\lambda)$ is a probability density function $f(\lambda)$ given by

$$f(\lambda) = \frac{dF(\lambda)}{d\lambda} \tag{7}$$

where $f(\lambda)d\lambda$ is the probability of having an average number of faults per chip between λ and $\lambda + d\lambda$. Averaging (6) with respect to this probability density function results in a yield expression first used by Murphy[6]:

$$Y = \int_0^\infty e^{-\lambda} f(\lambda)\ d\lambda. \tag{8}$$

The function $f(\lambda)$ in this expression is known as a *compounder* or *mixing function.* A commonly used mixing function is the Gamma distribution [7–12] with two parameters α and β,

$$f(\lambda) = \frac{1}{\beta^\alpha\ \Gamma(\alpha)}\ \lambda^{\alpha-1}\ e^{-\lambda/\beta}. \tag{9}$$

Evaluating the integral in equation (8) results in the well known integrated-circuit yield formula

$$Y = (1 + \bar{\lambda}/\alpha)^{-\alpha} \tag{10}$$

where α is a clustering parameter and $\overline{\lambda} = \beta\alpha$ is the average number of faults per chip. It can be shown that $\overline{\lambda}$ is in effect the expected value of λ when the probability density function $f(\lambda)$ in (9) is used. This is therefore the grand average (average of averages) of the number of faults per chip.

The clustering parameter α also has physical significance. In the limit when $\alpha \to \infty$, the yield in expression (10) becomes equal to yield formula (6). This represents the case of random faults and complete absence of clustering. Smaller values of α indicate increased clustering. Actual values for α typically range between 0.3 and 5. Methods for determining this parameter are described in the next section.

Applying the same averaging (or compounding) procedure to the Poisson probability function for the number of faults in (5), results in the negative binomial distribution,

$$Prob\ \{X = x\} = \frac{\Gamma(\alpha + x)}{x!\ \Gamma(\alpha)} \frac{(\overline{\lambda}/\alpha)^x}{(1 + \overline{\lambda}/\alpha)^{\alpha+x}}. \tag{11}$$

The yield formula in (10) is a special case of (11) with $x = 0$.

To illustrate the effect of fault clustering consider a chip containing a varying number of identical circuits. Let us start with a single circuit that has a hypothetical yield of 0.999 and an average of 0.001 faults per circuit. If we use model (6), the yield of a chip with 600 of these circuits is equal to $e^{-600\times0.001}$ which is approximately equal to 55%. For a chip with 40,000 logic circuits, we expect a yield of $e^{-40{,}000\times0.001} = 4.248 \cdot 10^{-18}$, or for all practical purposes 0%.

If clustering is taken into account, the yield formula for a chip with M identical circuits is given by

$$Y_M = (1 + M \cdot \overline{\lambda}_1/\alpha)^{-\alpha} \tag{12}$$

where the average number of faults in a single circuit is denoted by $\overline{\lambda}_1$. Assuming again that this number is equal to 0.001, it is possible to estimate the yield for chips with any number of circuits. Calculated yields for chips with a single circuit, chips with 600 circuits, and chips with 40,000 circuits are tabulated in Table 1 for values of $\alpha = 0.5, 1, 2$ and ∞. These results show that even if $M\overline{\lambda}_1$ is high, the presence of a high degree of clustering leads to surprisingly high yields. This effect has been observed in many manufacturing lines.

In most cases a gross yield factor Y_0 must be included in the yield model. Gross yield losses usually are the result of systematic processing problems that affect whole

Table 1: Yield as a function of the number of circuits per chip and the clustering parameter α, when $\overline{\lambda}_1 = .001$.

α	Yield in % M=1	Yield in % M=600	Yield in % M=40,000
0.5	99.9	67.4	11.1
1	99.9	62.5	2.4
2	99.9	59.2	0.2
∞	99.9	55.0	0

wafers or parts of wafers. Such losses may, for example, be caused by misalignment, over- or under-etching or out-of-spec semiconductor parameters such as beta transconductance or threshold voltage. Paz and Lawson[10] have shown that fault clusters with very high fault densities can also be modeled by Y_0.

Introduction of the gross yield into the yield formula leads to

$$Y = Y_0(1 + \bar{\lambda}/\alpha)^{-\alpha} \tag{13}$$

This three parameter model has been used successfully for yield modeling since 1975. Its parameters have physical significance and can be determined by a straightforward technique described in the next section. It must be pointed out, however, that the simplicity of this model can be deceptive. Some of the hidden complexities are discussed in subsequent sections.

SIZE OF CLUSTERS AND DETERMINATION OF PARAMETERS

Defect clusters in integrated circuits can be roughly categorized into three classes. The first includes clusters much larger than the chip size. Most papers on integrated-circuit yield which take fault clustering into account have assumed, sometimes unknowingly, that clusters are larger than the chip size. This is implied by the assumption that the value of the clustering parameter α is the same when the whole chip is considered or when only part of the chip is considered. The success of these yield models can be attributed to the fact that this is not a bad assumption. According to Stapper[13] most of the clustering is expected to be caused by wafer-to-wafer variations of fault densities. In that case, the cluster area is equal to the wafer size, which is indeed larger than the area of individual chips. Another source of clustering is the radial variation in the average number of faults per chip. This effect was originally described by Yanagawa[14-15], confirmed by others[10,16] and studied more recently by Ferris-Prabhu *et al*[17], Walker[18] and Gandemer[19]. It leads to a lower chip yield along the periphery of integrated-circuit wafers. This peripheral region can therefore in effect be considered a large-area cluster.

The radial variation of chip yield has led to the use of concentric wafer zones for yield analysis [10,13,16,20]. In such analyses, it is usually assumed that the faults per chip within each zone are distributed according to Poisson's distribution. Each zone has its own average number of faults per chip λ. The yield inside a zone can therefore be estimated by using formula (6). The yield of chips in all zones from many wafers can be combined and results in a *compound* or *mixed Poisson* yield model.

Another class of clusters deals with fault clusters that are smaller than the chip area. It is sometimes believed that the faults in such small clusters should distribute themselves according to Poisson's distribution. This distribution, however, is too constrained, because it has a variance that is equal to the mean. Clusters, by their very nature, tend to increase the variability in the number of faults per chip. As a result, clustering leads to distributions with variances that are larger than the mean. Statistics applicable to this type of clustering are described by Stapper[21].

The third class of fault clusters deals with clusters that vary in dimension. This area has been investigated by Warner[22,23], Hu[24], Stapper[25], and in an approximate point defect model for wafer-scale-integration by Ketchen[26]. A simulation technique

for modeling this situation has, furthermore, been described by Foard Flack[27] and Stapper[28]. These efforts, however, have not been definitive.

To understand how in practice one determines whether the large-area clustering assumption is valid (for a given size of chip) we need first to review the method used for determining the parameters of the yield model.

The values of Y_0, $\bar{\lambda}$ and α in formula (13) are usually determined with the "window method." This method was first described by Seeds[29-30] and subsequently by Okabe *et al*[31], Warner[22,23], Paz *et al*[10] and Hemmert[11]. The objective is to determine the yield as a function of chip multiples. This is done with wafer maps that show the location of functioning and failing chips at final test. The maps are analyzed using overlays with grids, or windows. These windows contain blocks of chips. Each block usually contains two, four, six, or nine chips. For each chip multiple, the number of windows containing only fault free chips can be counted. Dividing this number by the total number of windows in the sample gives us the yield for that multiple.

The results of the window analysis must next be matched to a yield formula. For the negative binomial model this has the form

$$Y_M = Y_0(1 + M\bar{\lambda}/\alpha)^{-\alpha} \tag{14}$$

where M is the chip multiple. Values for the parameters Y_0, $\bar{\lambda}$ and α are usually determined by means of a nonlinear regression analysis. Here it is implicitly assumed that the value of α is the same for all different sizes of windows, i.e., large-area clusters are assumed.

Note that high values of α obtained by this method do not necessarily mean that there is less clustering. It only implies that there is less large-area clustering. Small area clusters can still exist, but this method is impervious to them. The smaller clusters are essentially counted as single faults.

It is not difficult to use the window method. An example of a window method analysis is tabulated in Table 2. The data in that table came from 24 wafers, each one containing 89 memory chips. For each wafer a map was obtained to show the location of fault free and faulty chips. One of these maps is shown in Figure 1. Also shown on the map are locations taken up by test sites used to measure processing parameters.

The first step in evaluating the wafer map data is the determination of the chip yield. In this case 701 out of a total of 2136 chips were fault free. The yield was therefore 32.82%. Next, a transparent overlay was made with a grid containing pairs of chips. It was found that only 42 pairs could be placed on the grid of each wafer. This resulted in a sample of 1008 pairs. Only 140 of these were found to be free of faulty chips. The yield for these windows with blocks of two chips was therefore 13.89%.

The third step consisted of making an overlay grid that contained four chips in a 2 x 2 arrangement. Seventeen such windows could be fitted unambiguously on a wafer. To increase the sample size, and to include as much of the circumferential area as possible, three additional odd-shaped windows containing four chips were formed along the wafer edge. The total sample therefore contained 480 windows. For 18 of

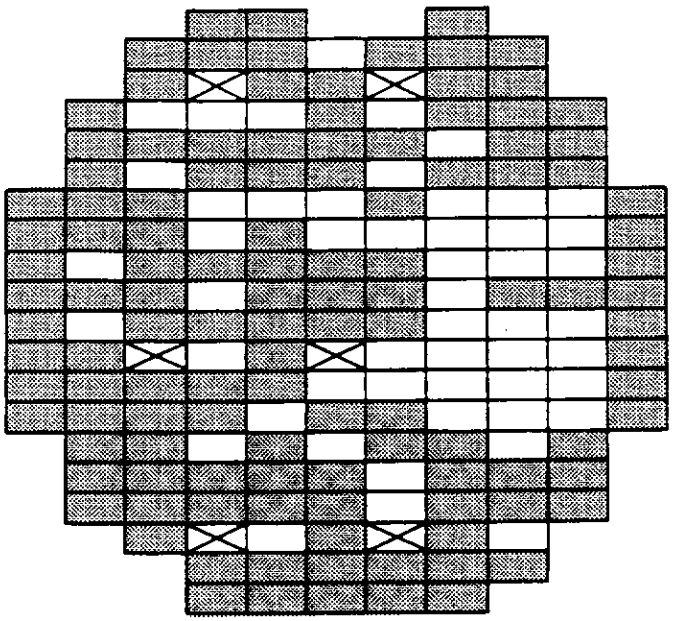

Fig. 1: Wafer map showing the locations of fault-free (light) and faulty (dark) chips. Test sites are marked with crosses.

Table 2: Illustrative use of the window method to determine model parameters. Here, $Y_0 = 1$, $\overline{\lambda} = 1.2934$, and $\alpha = 3.8274$.

Chip Multiples	Sample Size	Number Perfect	Yield in % Data	Yield in % Model
1	2136	701	32.82	32.82
2	1008	140	13.89	13.86
4	480	18	3.75	3.79

these windows it was found that all four chips were free of faults, thus resulting in a yield of 3.75%.

It is possible to obtain an additional data point by analyzing blocks of three chips. Such blocks, however, have odd-shaped windows, which makes them awkward to use. The three data points in Table 2 supply sufficient data for determining the parameters of the yield model. The values for λ, α and Y_0 were obtained by fitting equation (14) to these data points with a computer program that minimized the sum of the squares of the differences between model and data. With three data points and three parameters in equation (14), this was equivalent to solving three nonlinear equations with three unknowns. For these data, furthermore, it was possible to set $Y_0 = 1$. This led to the values $\overline{\lambda} = 1.2934$, and $\alpha = 3.8274$ for the other two parameters. Putting these values into equation (14) led to the numbers shown in the column labeled "Model Yield" in Table 2. The experimental yields are also tabulated

and are in good agreement. Because of the non-linearity, even with three data points, such agreement is not always guaranteed for this three parameter model.

The window method analysis is used regularly in the industry. A variation of such an analysis was described by Hemmert[11]. His data were obtained from wafer maps of logic chips and read only memories (ROMs). He used a least square fitting technique to determine $\bar{\lambda}$ and α in equation (14) while keeping Y_0 at 10C% yield. His results on seven manufacturing lots of wafers had an average cluster parameter of 2.2 with a standard deviation of 0.22. The values of α were therefore tightly grouped, indicating that they were stable during the fabrication of those lots.

An alternative use of equation (14) has been described in references [12,32]. The yield of different read only memory chips was analyzed as a function of the number of bits in those chips. This number was represented by M in equation (14). The values of Y_0, $\bar{\lambda}$, and α in that case were also determined with a nonlinear least square minimization technique. This analysis was performed on data from three different manufacturing lines and resulted in values for α of 1.27, 0.86, and 0.75. The lowest value, and therefore the highest degree of clustering, occurred on wafers fabricated in the manufacturing line with the highest chip yields. The highest value of α, suggesting less clustering, resulted from the wafers that were made in the line with the lowest chip yields.

The yield analysis of these read-only memory chips also showed that the gross yield Y_0 varied between 70.8% and 90.4%. Although these numbers include the yield of the support circuits on these chips, this range of gross yields is typical for most integrated circuits. The lowest value of Y_0 occurred in the low yield line and the highest value of Y_0 in the high yield line.

It must be noted here that the values of $\bar{\lambda}$ obtained by this method tend to be lower than the actual average number of faults observed on chips. This difference can be caused by the effect of clusters that are smaller than the chip. As mentioned before, such clusters are counted as single faults by this technique.

To find the chip sizes for which the large-area clustering assumption is valid particle distributions on actual wafers can be studied. This was done, for example, in reference [33], where wafer surfaces were subdivided into squares called *quadrats*. Negative binomial distribution were found in this study to be in good agreement with the frequency distribution of the number of particles in each quadrat for a wide range of quadrat sizes. The values of the cluster parameter α, however, differed for quadrats with different areas.

The data obtained with quadrat analysis can be analyzed using a maximum likelihood estimation technique described by Foard Flack[34]. This approach makes it possible to determine the variability in the estimated values of α. The results of such an analysis are shown in Figure 2. The bars around the data point indicate the range of $\pm\sigma_\alpha$, where σ_α is the standard deviation of each estimate. Note that the horizontal scale is logarithmic and represents a range of two orders of magnitude in area.

Of interest in Figure 2 are the results for the three smallest quadrat areas. The ranges of standard deviations overlap, thus suggesting that these points represent the condition for large-area clustering. The increase in values of α for the other points

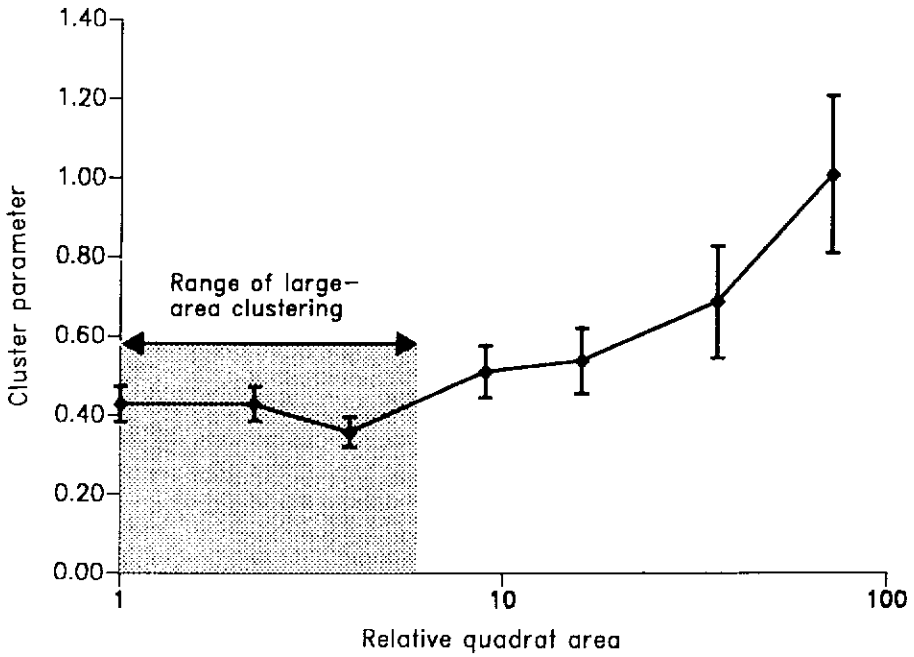

Fig. 2: Experimental dependence of the clustering parameter α on quadrat area.

on the curve indicate that the associated quadrat areas are exceeding the range for which the large-area clustering assumption is valid.

YIELD MODELS FOR CHIPS WITH REDUNDANCY

In many integrated circuit chips, identical blocks of circuits are often replicated. In memory chips, these are blocks of memory cells which are also known as *sub-arrays*. In processor arrays these basic circuit blocks are referred to as *processing elements*, or *PEs*. In other digital chips they are referred to as *macros*. The terminology depends not only on the type of circuitry that is used, but also by whom it is used. The designation *modules* is used in this paper. It is meant to be general and include all these designations.

Chips containing a number of identical modules (of one type or more) can often be used even if some of the modules do not function correctly. We obtain this way *partially good chips*. Alternatively, we can add a few redundant modules to our design and accept only those chips which have the necessary number of fault-free modules. We will first consider chips with a single type of identical modules and then extend our results to multiple module-type chips.

Let N be the number of identical circuit modules. Define the following probability

$$a_{M,N} = Prob\ \{\text{Exactly M out of the N modules are fault} - \text{free}\} \tag{15}$$

This probability can be used to calculate the yield of chips with redundancy and that of partially good chips. For example, if R out of the N modules are spares meaning that a chip with at least $(N - R)$ fault-free modules is acceptable, then the yield of the chip is given by

$$Y = \sum_{M=N-R}^{N} a_{M,N} \tag{16}$$

Two methods have been used to calculate the probability $a_{M,N}$. In the first method it is assumed that for any given subset of n modules we can compute the probability that this subset is fault-free. The latter is the yield of this subset and is denoted by

$$y_n = Prob\ \{X_n = 0\}; \qquad n = 1, 2, ..., N \tag{17}$$

where X_n is a random variable denoting the number of faults in n modules. For example, if a Poisson distribution is assumed then,

$$y_n = e^{-n\lambda} \tag{18}$$

while if the negative binomial distribution is assumed then,

$$y_n = (1 + n\bar{\lambda}/\alpha)^{-\alpha} \tag{19}$$

where λ (and similarly, $\bar{\lambda}$) is the average number of faults per module. Note that equation (19) is based on the large-area clustering assumption.

If the faults occurring in different modules are independent (as in the case where faults follow the Poisson distribution) then $y_n = y^n$ where $y = y_1$ is the yield of a single module, i.e., the probability that the module is fault-free. We can use in this case the binomial distribution to obtain the following expression for $a_{M,N}$

$$a_{M,N} = \binom{N}{M} y^M (1 - y)^{N-M} \tag{20}$$

If however, the faults in different modules are dependent (as in the case where faults follow the negative binomial distribution), then the Inclusion and Exclusion principle must be used to calculate the probability $a_{M,N}$. Defining the event - the i-th module is fault-free, then $a_{M,N}$ is the probability of exactly M such events occurring simultaneously, i.e.,

$$a_{M,N} = \binom{N}{M} \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} y_{M+k} \tag{21}$$

Notice that for the Poisson distribution equations (20) and (21) are equivalent.

In the second method for calculating $a_{M,N}$ we compute the probability that a given number of faults occur in the complete chip (containing N modules) and then

distribute these faults uniformly among the N modules. (Variants of this scheme were used in references[1-3,20,35-40].) Again, this is justified if large-area clustering is assumed. Thus, the probability that exactly $(N - M)$ modules will contain faults is[39,40]

$$a_{M,N} = \sum_{x=N-M}^{\infty} Q_{x,(N-M)}^{(N)} \cdot Prob\{X_N = x\} \tag{22}$$

where $Prob\{X_N = x\}$ is the probability that the chip has x faults and $Q_{x,j}^{(N)}$ is the probability that the x faults are distributed into exactly j out of N modules given that there are x faults. Assuming that faults are distinguishable, the latter equals[40],

$$Q_{x,j}^{(N)} = \sum_{k=0}^{j} (-1)^k \binom{N}{k, j-k, N-j} \left[\frac{j-k}{N}\right]^x \quad \text{for } x \geq j \text{ and } 0 < j \leq N \tag{23}$$

where $\binom{N}{k,j-k,N-j} = \binom{N}{k}\binom{N-k}{N-j}$ is the multinomial coefficient.

For $x < j$ we have $Q_{x,j}^{(N)} = 0$ and for $x = j = 0$ $Q_{0,0}^{(N)} = 1$ and consequently, we can rewrite (22) as follows,

$$a_{M,N} = \sum_{x=0}^{\infty} Q_{x,(N-M)}^{(N)} \cdot Prob\{X_N = x\} \tag{24}$$

We show next that the two methods in (21) and (24) yield the same expression for $a_{M,N}$ when the probability of having x faults follows the Poisson distribution, i.e.,

$$Prob\{X_N = x\} = \frac{e^{-N\lambda}(N\lambda)^x}{x!} \tag{25}$$

Substituting (23) and (25) in (24) results in,

$$\begin{aligned}
a_{M,N} &= \sum_{x=0}^{\infty} Prob\{X_N = x\} \cdot Q_{x,(N-M)}^{(N)} \\
&= \sum_{x=0}^{\infty} \frac{e^{-N\lambda}(N\lambda)^x}{x!} \cdot \sum_{k=0}^{N-M} (-1)^k \binom{N}{k, M, N-M-k} \left[\frac{N-M-k}{N}\right]^x \\
&= \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} \binom{N}{M} e^{-N\lambda} \sum_{x=0}^{\infty} \frac{[\lambda(N-M-k)]^x}{x!} \\
&= \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} \binom{N}{M} e^{-N\lambda} e^{(N-M-k)\lambda} \\
&= \binom{N}{M} e^{-N\lambda} e^{(N-M)\lambda} \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} e^{-k\lambda} \\
&= \binom{N}{M} e^{-M\lambda}(1 - e^{-\lambda})^{N-M} = \binom{N}{M} y^M (1-y)^{N-M}
\end{aligned}$$

The equivalence of the above two methods is not restricted to the simple case of the Poisson distribution for faults but holds for other distributions as well, in particular for the negative binomial distribution as shown in what follows. Substituting (23) in (24) we obtain,

$$a_{M,N} = \sum_{x=0}^{\infty} Prob\{X_N = x\} \cdot Q_{x,(N-M)}^{(N)}$$

$$= \sum_{x=0}^{\infty} Prob\{X_N = x\} \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} \binom{N}{M} \left[\frac{N-M-k}{N}\right]^x$$

$$= \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} \binom{N}{M} \sum_{x=0}^{\infty} Prob\{X_N = x\} \left[\frac{N-M-k}{N}\right]^x$$

$$= \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} \binom{N}{M} G\left(\frac{N-M-k}{N}\right)$$

where $G(s)$ is the generating function of the probability distribution $Prob\{X_N = x\}$. For the negative binomial distribution we have

$$G(s) = \left[1 + \frac{(1-s)N\bar{\lambda}}{\alpha}\right]^{-\alpha} \tag{26}$$

Substituting (26) into the above equation yields,

$$a_{M,N} = \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} \binom{N}{M} \left[1 + \frac{\left(1 - \frac{N-M-k}{N}\right) N\bar{\lambda}}{\alpha}\right]^{-\alpha}$$

$$= \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} \binom{N}{M} \left[1 + \frac{(M+k)\bar{\lambda}}{\alpha}\right]^{-\alpha} \tag{27}$$

which is identical to (21) after substituting y_{M+k} by its proper expression from (19). Note that the equivalence of the two schemes for the Poisson distribution can be proved similarly using the generating function for the Poisson distribution which is given by

$$G(s) = e^{N\lambda(s-1)} \tag{28}$$

The negative binomial distribution is obtained from the Poisson distribution by averaging over all values of λ, using the Gamma distribution function. This compounding procedure can be applied to any statistical measure. We can derive an expression for the desired measure assuming the very convenient Poisson distribution (whose most useful property is the statistical independence between faults in different modules), and then apply the compounding procedure to obtain the required expression for the negative binomial model.

To illustrate this procedure we show next that $a_{M,N}$ in (27) can be obtained by compounding (20) when $y = e^{-\lambda}$. Equation (20) can be rewritten in the form of (21) by expanding $(1-y)^{N-M}$ into the following binomial series

$$(1-y)^{N-M} = \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} y^k \tag{29}$$

Substituting this series expansion into (20) results in,

$$a_{M,N} = \binom{N}{M} \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} y^{M+k} \tag{30}$$

calculated using

$$Y_i = \sum_{M_i=N_i-R_i}^{N_i} a_{M_i,N_i} \tag{40}$$

The expression for Y will consist of $R_1 \cdot R_2$ terms. However, in many practical chip architectures there is no such *architectural independence* and a fault in a module of type 1 may affect the usefulness of type 2 modules, i.e., a type 2 module may become useless when a type 1 module is defective. In such a case not all $R_1 \cdot R_2$ possible terms should be included in the expression for Y. In well-structured architectures we can easily identify those terms that should be included and we can therefore define a "coverage factor" as follows, $c_{M_1,M_2} = 1$ if the chip is acceptable with M_1 and M_2 fault-free modules of type 1 and 2, respectively. Otherwise, $c_{M_1,M_2} = 0$. Consequently,

$$Y = \sum_{M_1=N_1-R_1}^{N_1} \sum_{M_2=N_2-R_2}^{N_2} a_{M_1,N_1} \cdot a_{M_2,N_2} \cdot c_{M_1,M_2} \tag{41}$$

c_{M_1,M_2} serves to select all the *fixable* combinations out of all combinations of fault-free modules of type 1 and type 2.

In less structured architectures, the number of fault-free modules of either type may be insufficient to determine whether the chip is fixable or not; we may also need to know the exact position of the fault-free modules. In such a case, c_{M_1,M_2} will not be a factor assuming only the values 0 and 1, but the fraction of fixable patterns out of all patterns consisting of M_1 and M_2 fault-free modules of type 1 and 2, respectively.

The final expression for the yield of a chip with two types of identical modules and support circuitry when the Poisson distribution is assumed, is as follows,

$$Y = Y_0 \sum_{\substack{M_1= \\ N_1-R_1}}^{N_1} \sum_{\substack{M_2= \\ N_2-R_2}}^{N_2} \sum_{k_1=0}^{N_1-M_1} \sum_{k_2=0}^{N_2-M_2} (-1)^{k_1}(-1)^{k_2} \binom{N_1}{M_1} \binom{N_1-M_1}{k_1} \binom{N_2}{M_2} \binom{N_2-M_2}{k_2}$$

$$\cdot\, e^{-(M_1+k_1)\lambda_1} e^{-(M_2+k_2)\lambda_2} e^{-\lambda_{CK}} \cdot c_{M_1,M_2} \tag{42}$$

Next we have to apply the compounding procedure to calculate the yield when clustering of faults is allowed. We should not however, perform three separate compounding steps (for the two types of modules and the support circuits) since the clustering of faults in one type of circuits is not independent of the clustering in the other two. We must therefore, perform a single compounding step using the average number of faults in the complete chip, i.e.,

$$\lambda = \lambda_1 N_1 + \lambda_2 N_2 + \lambda_{CK} \tag{43}$$

To simplify the integration of the various summands in (42) which contain different multiples of λ_1 and λ_2 we define,

$$\delta_1 = \frac{\lambda_1 N_1}{\lambda}, \quad \delta_2 = \frac{\lambda_2 N_2}{\lambda} \quad \text{and} \quad \delta_3 = \frac{\lambda_{CK}}{\lambda}$$

Note that δ_1 (and similarly, δ_2) is a constant which depends mainly on the area ratio of a type 1 module to the whole chip. The exponential terms in (42) now take the form,

$$e^{-(M_1+k_1)\lambda_1-(M_2+k_2)\lambda_2-\lambda_{CK}} = e^{-[(M_1+k_1)\delta_1/N_1+(M_2+k_2)\delta_2/N_2+\delta_3]\lambda}$$

Some practical modifications

The simple architecture analyzed in the preceding section is an idealization because actual chips rarely consist entirely of identical circuit modules. In all chips there are support circuits in addition to such modules. These support circuits are shared by the replicated modules. The chips, however, become unusable if such support circuits are damaged beyond use. In principle, this effect can be included in the expression for $a_{M,N}$ by multiplication with the yield of the support circuits. Doing so, however, would assume that the clustering of the support circuit faults is completely independent of the clustering of the module faults. In most practical cases there is a dependence between the average number of faults in different circuits. This effect can be taken into account by including in formula (19) the average number of faults that cause these support circuits to be defective. This results in

$$y_n = [1 + (\bar{\lambda}_{CK} + n\bar{\lambda})/\alpha]^{-\alpha} \tag{37}$$

where $\bar{\lambda}_{CK}$ is the average number of fatal faults or *chip-kill* faults in the support circuits. Chips with these faults cannot be used. Substituting expression (37) into formula (31) makes it possible to take these types of faults into account when calculating the yields of partially-good chips (or chips with redundancy) with support circuitry.

Another effect that must be included in yield estimates is the gross yield. Unless the chips are very large, this yield is independent of chip area. It is used as a yield multiplier, which has been denoted by Y_0 in the preceding sections. Introducing it into yield formula (37) results in

$$y_n = Y_0[1 + (\bar{\lambda}_{CK} + n\bar{\lambda})/\alpha]^{-\alpha} \tag{38}$$

Introduction of this expression into equation (31) results in a formula that can be used to estimate yields of partially-good chips and chips with redundancy, with support circuits and gross yield losses.

Multiple Module-type Chips

The discussion above was restricted to the case where redundancy is provided to tolerate faults in a single type of circuit modules. In this section we extend the previous results to fault tolerant chips with multiple types of modules. We derive yield expressions for chips with two different types of modules, say, Type 1 and Type 2. The extension to a larger number of module types is straightforward and is therefore, not presented here.

Suppose that there are redundant modules of both types and that the modules of each type can be reconfigured separately when necessary. Then, we can calculate the yield of each module type separately (assuming that faults follow the Poisson distribution), and multiply the two results to obtain the overall yield,

$$Y = Y_1 \cdot Y_2 \tag{39}$$

where Y_i $(i = 1, 2)$ is the yield of the set of N_i modules of type i. This yield can be

By compounding (30) with the Gamma distribution in (9) we obtain,

$$a_{M,N} = \binom{N}{M} \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} \int_0^\infty e^{-(M+k)\lambda} \cdot f(\lambda) \cdot d\lambda$$

$$= \binom{N}{M} \sum_{k=0}^{N-M} (-1)^k \binom{N-M}{k} \left[1 + \frac{(M+k)\overline{\lambda}}{\alpha}\right]^{-\alpha} \tag{31}$$

This very powerful compounding procedure was employed to derive yield expressions for interconnection buses in VLSI chips[41], for wafer scale cube-connected-cycles[42], and for partially good memory chips[28,37].

The expression for $a_{M,N}$ in (31) can be used to calculate the yield of chips with redundancy, using equation (16), and partially good chips as will be shown subsequently. To calculate the yield of a chip with a single type of redundant modules we can derive a somewhat simpler expression involving one summation instead of two. Employing the previous notation, the required yield is

$$Y = Prob\ \{\text{There are at least (N - R) fault - free modules}\} \tag{32}$$

Define the event - the i-th module is fault-free, then Y is the probability that at least $(N - R)$ such events occur simultaneously. According to the Inclusion and Exclusion principle, Y can be written as,

$$Y = \sum_{n=N-R}^{N} (-1)^{n-N+R} \binom{n-1}{N-R-1} \binom{N}{n} y_n \tag{33}$$

It should be noted that we may use either equation (18) or (19) for y_n in (33) and obtain a yield formula for evenly distributed faults or clustered faults, respectively. Replacing the index in (33) by $i = N - n$ yields,

$$Y = \sum_{i=0}^{R} (-1)^{R-i} \binom{N}{N-i} \binom{N-1-i}{N-R-1} y_{N-i} \tag{34}$$

This is the yield formula (with a different notation, i.e., $M = N - R$) presented by Harden and Strader[43] and deduced from several special cases of N and R. The yield expressions in (34) and (16) are equivalent.

We conclude this section with an example illustrating the effect of clustering on the yield of chips with redundancy. Consider a chip on which ten identical circuit modules must be functioning correctly if the chips are to be usable. Let the yield of the ten circuit modules be equal to 10%. We can then investigate how the chip yield is affected when we have one to five redundant circuits. This is done in Table 3, where yields (in %) correspond to different values of the clustering parameter α.

The pure random fault model corresponds to $\alpha = \infty$. In this case, according to Table 3, the use of five redundant circuits increases the yield from 10% to 93.1%. If, however, the clustering parameter is $\alpha = 0.5$, the yield is expected to improve from 10% to 24.1%. This indicates that the yield prediction for purely random faults is four times higher than the prediction for clustered faults. Miscalculations by a factor of four in the productivity of semiconductor manufacturing plants can be very costly.

Table 3: Yield (in %) with different amounts of redundancy for varying degrees of large-area fault clustering.

α	R=0	R=1	R=2	R=3	R=4	R=5
0.5	10	14.6	17.8	20.3	22.4	24.1
1	10	18.3	25.2	31.1	36.1	40.6
2	10	22.4	34.3	44.8	53.7	61.1
∞	10	30.6	53.8	73.0	85.8	93.1

Inclusion of clustering in redundancy yield calculation is therefore of considerable importance.

Partially Good Chips

Partially good chips are chips which are usable even if only some of their identical modules are fault-free. Consider for example, chips consisting of four identical modules. These chips are known as *perfect* if all four modules are fault-free. The fraction of chips falling in this category represents the perfect chip yield. The chips with three operating modules and one defective module are referred to as being *three-quarter-good.* The yield of these chips is known as the three-quarter-good yield and is equal to $a_{3,4}$. Similar designations apply to chips that are *half-good* and *quarter-good* and their yields are $a_{2,4}$ and $a_{1,4}$, respectively. In general, the yield of a partially-good chip with exactly M fault-free modules out of N, is given by $a_{M,N}$. The difference between a partially-good chip and a chip with redundancy is that *all* fault-free modules in a partially-good chip are considered to be usable while only $(N-R)$ fault-free modules in a chip with redundancy are expected to be used. Thus, a three-quarter-good chip is more valuable than a half-good chip. Consequently, instead of summing up all $a_{M,N}$'s to obtain the yield as is done for chips with redundancy, we multiply each $a_{M,N}$ by a weight equal to M/N to obtain what is called the *equivalent yield,*

$$Y_{EQ} = \sum_{M=J}^{N} \frac{M}{N} a_{M,N} \tag{35}$$

where J is the minimum number of modules which have to be fault-free.

We show in what follows that if $J = 1$ then $Y_{EQ} = y$, i.e., the equivalent yield in this case is equal to the yield of the individual circuit module. This implies that the use of partially good chips results in utilization of all the fault-free modules; none have been wasted. We first prove it for the Poisson distribution using equation (20),

$$Y_{EQ} = \sum_{M=1}^{N} \frac{M}{N} a_{M,N} = \sum_{M=1}^{N} \frac{M}{N} \binom{N}{M} y^M (1-y)^{N-M} = y \sum_{M=1}^{N} \binom{N-1}{M-1} y^{M-1} (1-y)^{N-M}$$

Substituting $m = M - 1$ and $n = N - 1$ yields,

$$Y_{EQ} = y \sum_{m=0}^{n} \binom{n}{m} y^m (1-y)^{n-m} = y \tag{36}$$

Applying the compounding procedure to the above equation proves our claim for the negative binomial distribution as well.

Substituting the above in (42) and compounding with respect to λ results in,

$$Y = Y_0 \sum_{\substack{M_1=\\ N_1-R_1}}^{N_1} \sum_{\substack{M_2=\\ N_2-R_2}}^{N_2} \sum_{k_1=0}^{N_1-M_1} \sum_{k_2=0}^{N_2-M_2} (-1)^{k_1}(-1)^{k_2} \binom{N_1}{M_1} \binom{N_1-M_1}{k_1} \binom{N_2}{M_2} \binom{N_2-M_2}{k_2}$$

$$\cdot \left[1 + \frac{((M_1+k_1)\delta_1/N_1 + (M_2+k_2)\delta_2/N_2 + \delta_3)\,\overline{\lambda}}{\alpha}\right]^{-\alpha} \cdot c_{M_1,M_2} \qquad (44)$$

Finally, we define $\overline{\lambda}_1 = \overline{\lambda}\delta_1/N_1$, $\overline{\lambda}_2 = \overline{\lambda}\delta_2/N_2$, $\overline{\lambda}_{CK} = \overline{\lambda}\delta_3$ and obtain,

$$Y = Y_0 \sum_{\substack{M_1=\\ N_1-R_1}}^{N_1} \sum_{\substack{M_2=\\ N_2-R_2}}^{N_2} \sum_{k_1=0}^{N_1-M_1} \sum_{k_2=0}^{N_2-M_2} (-1)^{k_1}(-1)^{k_2} \binom{N_1}{M_1} \binom{N_1-M_1}{k_1} \binom{N_2}{M_2} \binom{N_2-M_2}{k_2}$$

$$\cdot \left[1 + \frac{\left((M_1+k_1)\overline{\lambda}_1 + (M_2+k_2)\overline{\lambda}_2 + \overline{\lambda}_{CK}\right)}{\alpha}\right]^{-\alpha} \cdot c_{M_1,M_2} \qquad (45)$$

A simpler expression for the yield of a multiple module-type chip can be derived if it is possible to determine for any single fault whether it can be tolerated or not. In this case, instead of considering modules which may have any number of faults, we examine individual faults. An example is a memory chip which can have the following types of faults: single cell faults, adjacent cell faults, single word line faults, adjacent word line faults, single bit line faults and adjacent bit line faults[37]. A fault-tolerant memory chip has three types of identical circuits for which some form of redundancy is provided, namely, memory cells, word lines and bit lines. However, only two types of redundant circuits are added to the memory chip: spare word lines and bit lines. These two types of redundant circuits are used to replace defective word lines, bit lines and memory cells.

We first derive a yield expression for a chip with two types of faults, type a and type b, and then extend our result to memory chips with a larger number of fault types. Let λ_a and λ_b denote the average number of faults of type a and b, respectively. Then,

$$Y = Y_0 \sum_{k_a,k_b} Prob\ \{\text{There are } k_a \text{ faults of type } a \text{ and } k_b \text{ faults of type } b\} \cdot h_{k_a,k_b}$$

where h_{k_a,k_b} is the probability that the combination of k_a and k_b faults can be tolerated. The above yield expression can be derived by first assuming independency between the two types of faults (i.e., faults follow the Poisson distribution) and then applying the compounding procedure to allow for fault clustering. We also take into account the support circuitry and denote by λ_{CK} the average number of faults in this part of the chip. Consequently,

$$Y = Y_0 \sum_{k_a,k_b} \frac{\lambda_a^{k_a} e^{-\lambda_a}}{k_a!} \cdot \frac{\lambda_b^{k_b} e^{-\lambda_b}}{k_b!} \cdot e^{-\lambda_{CK}} \cdot h_{k_a,k_b}$$

$$= Y_0 \sum_{k_a,k_b} \frac{\lambda_a^{k_a}\lambda_b^{k_b}}{k_a!k_b!} e^{-(\lambda_a+\lambda_b+\lambda_{CK})} \cdot h_{k_a,k_b} \qquad (46)$$

The compounding procedure has to be applied only once due to the dependence between the clustering of the three types of faults. Therefore, we define as before

$$\lambda = \lambda_a + \lambda_b + \lambda_{CK}, \quad \delta_1 = \frac{\lambda_a}{\lambda}, \quad \delta_2 = \frac{\lambda_b}{\lambda}, \quad \text{and} \quad \delta_3 = \frac{\lambda_{CK}}{\lambda}$$

Substituting these in (46) yields

$$Y = Y_0 \sum_{k_a,k_b} \frac{\lambda^{k_a+k_b}\,\delta_1^{k_a}\,\delta_2^{k_b}}{k_a!k_b!}\, e^{-\lambda} \cdot h_{k_a,k_b} \tag{47}$$

Compounding now with respect to λ results in,

$$Y = Y_0 \sum_{k_a,k_b} \frac{\Gamma(k_a+k_b+\alpha)}{k_a!k_b!\,\Gamma(\alpha)} \cdot \frac{(\frac{1}{\alpha})^{k_a+k_b}\,\delta_1^{k_a}\,\delta_2^{k_b}\,\bar{\lambda}^{k_a+k_b}}{(1+\frac{\bar{\lambda}}{\alpha})^{k_a+k_b+\alpha}} \cdot h_{k_a,k_b} \tag{48}$$

Defining $\bar{\lambda}_a = \bar{\lambda}\delta_1$ and $\bar{\lambda}_b = \bar{\lambda}\delta_2$ and substituting in (48) we obtain

$$Y = Y_0 \sum_{k_a,k_b} \frac{\Gamma(k_a+k_b+\alpha)}{k_a!k_b!\,\Gamma(\alpha)} \cdot \frac{(\frac{1}{\alpha})^{k_a+k_b}\,\bar{\lambda}_a^{k_a}\,\bar{\lambda}_b^{k_b}}{(1+\frac{\bar{\lambda}}{\alpha})^{k_a+k_b+\alpha}} \cdot h_{k_a,k_b} \tag{49}$$

We now extend the above result to memory chips with seven types of faults: single cell faults, adjacent cell faults along word lines, adjacent cell faults along bit lines, single word line faults, adjacent word line faults, single bit line faults and adjacent bit line faults. The corresponding fault averages are denoted by λ_{sc}, λ_{acwl}, λ_{acbl}, λ_{swl}, λ_{awl}, λ_{sbl} and λ_{abl}. Following the same procedure as outlined in equations (46) through (49) we arrive at the following yield expression,

$$Y = Y_0 \sum_{i,j,k,l,m,n,q} \frac{\Gamma(i+j+k+l+m+n+q+\alpha)}{i!\,j!\,k!\,l!\,m!\,n!\,q!\,\Gamma(\alpha)}$$

$$\cdot \frac{(\frac{1}{\alpha})^{i+j+k+l+m+n+q}\,\bar{\lambda}_{sc}^{i}\bar{\lambda}_{acwl}^{j}\bar{\lambda}_{acbl}^{k}\bar{\lambda}_{swl}^{l}\bar{\lambda}_{awl}^{m}\bar{\lambda}_{sbl}^{n}\bar{\lambda}_{abl}^{q}}{(1+\frac{\bar{\lambda}}{\alpha})^{i+j+k+l+m+n+q+\alpha}} \cdot h_{i,j,k,l,m,n,q} \tag{50}$$

where the definitions of $\bar{\lambda}_{sc}$, $\bar{\lambda}_{acwl}$, etc are similar to those above.

Finally, note that we must assume that the probability of multiple faults (of any of the seven types) occurring in a single word line (or bit line) is negligible. Otherwise, the summation in (50) will be infinite. In VLSI memory chips with thousands of word lines and bit lines and a small number of expected faults, this assumption is well justified.

Also, it must be noted here that the above approach is valid only if there is a perfect correlation between the distribution of faults of different types. Such a perfect correlation was assumed in the derivation of (45), (49) and (50). If there is no correlation, independent negative binomial distributions, each with their own clustering parameter α, have to be used for each fault type. Separate compounding steps for each type of faults have to be performed, as was done in reference[41]. Small correlation between failure types requires the use of a multivariate model like the one described by Stapper *et al*[20].

CONCLUSIONS

The statistical models used to estimate the manufacturing yield of defect tolerant VLSI chips have been reviewed in this paper. We have shown how to take fault clustering into account when deriving yield expressions for VLSI chips with redundancy or partially-good chips, with one or more types of circuit modules. Also, the method used to determine the parameters of the yield model was reviewed and the assumed size of fault clusters was discussed.

ACKNOWLEDGMENTS

The work of the first co-author was supported in part by NSF under contract MIP-8805586.

The figures and tables in this manuscript also appeared in the IBM Journal of Research and Development, Vol. 33, No. 2., March 1989 in a paper by the second co-author entitled "Large Area Fault Clusters and Fault Tolerance in VLSI Circuits: A Review." These figures and tables are copyrighted in 1989 by the International Business Machines Corporation and are reprinted here with permission.

REFERENCES

1. T.E. Mangir, "Use of On-Chip Redundancy for Fault-Tolerant Very Large Scale Integrated Circuit Design," Ph.D. Dissertation, Univ. of Calif., Los Angeles, 1981.
2. K.S. Hedlund, "Wafer Scale Integration of Configurable, Highly Parallel Processors," Ph.D. Dissertation, Purdue Univ., Indiana, 1982.
3. J.C. Harden, "A Wafer Scale Cellular Tree Architecture," Ph.D. Dissertation, Texas A & M University, College Station, TX, 1985.
4. T. J. Wallmark, "Design Considerations for Integrated Electron Devices," *Proc. IRE,* Vol. 48, pp. 293-300, March 1960.
5. S. R. Hofstein and F. P. Heiman, "The Silicon Insulated-Gate Field Effect Transistor," *Proc. IEEE,* Vol. 51, pp. 1190-1202, Sept. 1963.
6. B.T. Murphy, "Cost-size Optima of Monolithic Integrated Circuits," *Proc. IEEE,* Vol. 52, pp. 1537-1545, Dec. 1964.
7. C.H. Stapper, "Defect Density Distribution for LSI Yield Calculations," *IEEE Trans. Electron Devices,* Vol. ED-20, pp. 655-657, July 1973.
8. A.P. Turley and D.S. Herman, "LSI Yield Projections Based Upon Test Patterns Results: An Application to Multilevel Metal Structures," *IEEE Trans. Parts, Hybrids, Packag.*, Vol. PHP-10, pp. 230-234, Dec. 1974.
9. C.H. Stapper, "On a Composite Model of the IC Yield Problem," *IEEE J. Solid-State Circuits,* Vol. SC-10, pp. 537-539, Dec. 1975.
10. O. Paz and T.R. Lawson, Jr., "Modification of Poisson Statistics: Modeling Defects Induced by Diffusion," *IEEE J. Solid-State Circuits,* Vol. SC-12, pp. 540-546, Oct. 1977.
11. R.S. Hemmert, "Poisson Process and Integrated Circuit Yield Prediction," *Solid-State Electronics,* Vol. 24, pp. 511-515, June 1981.

12. C.H. Stapper and R.J. Rosner, "A Simple Method for Modeling VLSI Yields," *Solid-State Electronics*, Vol. 25, pp. 487-489, June 1982.
13. C.H. Stapper, "The Effects of Wafer to Wafer Density Variations on Integrated Circuit Defect and Fault Distributions," *IBM J. Res. Develop.*, Vol. 29, pp. 87-97, January 1985.
14. T. Yanagawa, "Influence of Epitaxial Mounds on the Yield of Integrated Circuits," *Proc. IEEE*, Vol. 57, pp. 1621-1696, Sept. 1969.
15. T. Yanagawa, "Yield Degradation of Integrated Circuits Due to Spot Defects," *IEEE Trans. Electron Devices*, Vol. ED-19, pp. 190-197, Feb. 1972.
16. C.H. Stapper, "LSI Yield Modeling and Process Monitoring," *IBM J. Res. Develop.*, Vol. 20, pp. 228-234, May 1976.
17. A.V. Ferris-Prabhu, L.D. Smith, H.A. Bonges and J.K. Paulsen, "Radial Yield Variations in Semiconductor Wafers," *IEEE Circuits and Devices Magazine*, Vol. 3, pp. 42-47, March 1987.
18. D.M.H. Walker, *Yield Simulation for Integrated Circuits*, Boston: Kluwer Academic Publishers, 1987, Ch. 4, pp. 45-49, Ch. 8, pp. 158-160.
19. S. Gandemer, "Modelisation de l'Impact des Defauts de Fabrication sur le Rendement des Microcircuits Integres Fabriques en Technologie Silicium," Doctoral Dissertation, Ecole Nationale Superieure des Telecommunications, Sept. 1987.
20. C.H. Stapper, F.M. Armstrong and K. Saji, "Integrated Circuit Yield Statistics," *Proc. IEEE*, Vol. 71, pp. 453-470, April 1983.
21. C.H. Stapper, "Small-Area Fault Clusters and Fault-Tolerance in VLSI Circuits," *IBM J. Res. Develop.*, Vol. 33, March 1989.
22. R.M. Warner, "Applying a Composite Model to the IC Yield Problem," *Solid-State Circuits*, Vol. SC-9, pp. 86-95, June 1974.
23. R.M. Warner, "A Note on IC Yield Statistics," *Solid-State Electronics*, Vol. 24, pp. 1045-1047, Dec. 1981.
24. S.M. Hu, "Some Considerations in the Formulation of IC Yield Statistics," *Solid-State Electronics*, Vol. 22, pp. 205-211, Feb. 1979.
25. C.H. Stapper, "Comments on 'Some Considerations in the Formulation of IC Yield Statistics'," *Solid-State Electronics*, Vol. 24, pp. 127-132, Feb. 1981.
26. M.B. Ketchen, "Point Defect Yield Model for Wafer Scale Integration," *IEEE Circuits and Devices Magazine*, Vol. 1, No. 4, pp. 24-34, July 1985.
27. V. Foard Flack, "Introducing Dependency into IC Yield Models," *Solid-State Electronics*, Vol. 28, No. 6, pp. 555-559, June 1985.
28. C.H. Stapper, "Block Alignment: A Method for Increasing the Yield of Memory Chips that are Partially Good," *Defect and Fault Tolerance in VLSI Systems*, I. Koren (ed.), pp. 243-255, New York: Plenum, 1989.
29. R.B. Seeds, "Yield, Economic, and Logistic Models for Complex Digital Arrays," in *1967 IEEE Int. Conv. Rec.*, 1967, pt. 6, pp. 61-66.
30. R.B. Seeds, "Yield and Cost Analysis of Bipolar LSI," presented at the 1967 Int. Electron Device Meeting Keynote Session, p. 12, Oct. 1967.
31. T. Okabe, M. Nagata, and S. Shimada, "Analysis on Yield of Integrated Circuits and a New Expression for the Yield," *Elec. Eng. Japan*, Vol. 92, pp. 135-141, Dec. 1972.
32. C.H. Stapper, "The Defect-Sensitivity Effect of Memory Chips," *IEEE J. Solid-State Circuits*, Vol. SC-21, pp. 193-198, Feb. 1986.

33. C.H. Stapper, "On Yield, Fault Distributions and Clustering of Particles," *IBM J. Res. Develop.*, Vol. 30, pp. 326-338, May 1986.
34. V. Foard Flack, "Estimating Variations in IC Yield Estimates," *IEEE J. of Solid-State Circuits*, Vol. SC-21, pp. 362-365, April 1986.
35. T.E. Mangir, "Sources of Failures and Yield Improvement for VLSI and Restructurable Interconnects for RVLSI and WSI: Part I - Sources of Failures and Yield Improvement for VLSI," *Proc. IEEE*, Vol. 72, pp. 690-708, June 1984.
36. J.C. Harden, "Comments on 'Sources of Failures and Yield Improvement for VLSI and Restructurable Interconnects for RVLSI and WSI'," *Proc. IEEE*, Vol. 74, pp. 515-516, March 1986.
37. C.H. Stapper, A.N. McLaren, and M. Dreckmann, "Yield Model for Productivity Optimization of VLSI Memory Chips with Redundancy and Partially Good Product," *IBM J. Res. Develop.*, Vol. 20, pp. 398-409, 1980.
38. I. Koren and D.K. Pradhan, "Introducing Redundancy into VLSI Designs for Yield and Performance Enhancement," *Proc. 15th Annual Int. Symp. on Fault-Tolerant Computing*, pp. 330-335, 1985.
39. I. Koren and D.K. Pradhan, "Yield and Performance Enhancement through Redundancy in VLSI and WSI Multiprocessor Systems," *Proc. IEEE*, Vol. 74, pp. 699-711, May 1986.
40. I. Koren and D.K. Pradhan, "Modeling the Effect of Redundancy on Yield and Performance of VLSI Systems," *IEEE Trans. on Computers*, Vol. C-36, pp. 344-355, March 1987.
41. I. Koren, Z. Koren and D.K. Pradhan, "Designing Interconnection Buses in VLSI and WSI for Maximum Yield and Minimum Delay," *IEEE J. of Solid-state Circuits*, pp. 859-866, June 1988.
42. J-J. Shen and I. Koren, "Yield Enhancement Designs for WSI Cube Connected Cycles," *Proc. of Int. Conf. on WSI*, pp. 289-298, Jan. 1989.
43. J.C. Harden and N.R. Strader, "Architectural Yield Optimization for WSI," *IEEE Trans. on Computers*, Vol. C-37, pp. 88-110, Jan. 1988.