Bernd Meister

On Murphy's Yield Formula

Some properties of yield are presented, and one lower and three upper bounds for yield are derived. Some of these bounds represent yield formulas already known as useful approximations. Pure Poisson statistics for defect density provides the lower bound. The upper bounds are obtained with mixtures of Poisson distributions and a formula of Price and Stapper.

Introduction

A very general formula for the yield of a single processing step in the fabrication of integrated circuits was given by Murphy [1]. However, the complete probability distribution function of the defect density must be known in order to evaluate this formula, which can be used both for discrete and for continuous defect density distributions. The yield then becomes a function of the active area and the final yield for the whole fabrication process is the product of the yields of the individual steps.

It is shown in what follows that useful approximations for the yield can be obtained even if less than the complete defect density information is available, e.g., if only the first moment or the first two moments of the defect density are given. These approximations have the additional property that they are lower or upper bounds; i.e., it is known *a priori* whether they underestimate or overestimate the yield. The upper bounds are of special interest because if an optimistic approximation for the yield shows that it is insufficient, the corresponding process step has to be improved so that the defect density is suitably reduced.

The lower bound is obtained by pure Poisson statistics for the defect density. Several authors have observed that this use of statistics leads to a pessimistic approximation for the yield [1–14]. One of the upper bounds is a yield formula proposed by Price [15] and Stapper [11]. The other two upper bounds are obtained with mixtures of Poisson distributions [6, 12].

The main goals of this paper are 1) to show that the approximations for yield (some of which are already well known) are bounds in the strictest mathematical sense, and 2) to present two simple properties of yield.

Properties of yield

Murphy's yield formula [1] states that

$$Y_F(a) = \int_0^\infty e^{-ta} dF(t), \qquad (1)$$

which gives the yield as a function of the active area a. The defect density D is a random variable with a probability distribution function (or cumulative distribution function) $F(t) = \Pr\,[D < t]$, where Pr indicates probability. We are mainly interested in the interdependence between the yield and the distribution of defect density. If the derivative $f(t) = F'(t)$ exists, Eq. (1) can also be written as

$$Y_F(a) = \int_0^\infty e^{-at} f(t)\,dt.$$

In the following discussion, we need a certain ordering in the set of all defect-density distributions. This ordering has a simple physical interpretation, since we call a defect density F_1 "larger than" a defect density F_2 if the yield corresponding to F_1 is smaller than the yield corresponding to F_2.

First, the definition of a semi-ordering relation is given. A relation for which we shall use the symbol $(R)_\leq$ is called a semi-ordering relation if the following conditions are fulfilled:

1. $F\ (R)_\leq\ F$.
2. From $F_1\ (R)_\leq\ F_2$, $F_2\ (R)_\leq\ F_3$, it follows that $F_1\ (R)_\leq\ F_3$.
3. From $F_1\ (R)_\leq\ F_2$, $F_2\ (R)_\leq\ F_1$, it follows that $F_1 = F_2$.

These are the same conditions which hold for the $\geq$ relation in the set of real numbers. But in the field of real numbers, two numbers can always be compared with respect to $\geq$; i.e., for any two arbitrary numbers x and y, either $x \geq y$ or $y \geq x$. A semi-ordering relation does not require the property that all pairs of functions can be compared; i.e., two functions F* and F** may exist such that none of the following statements is true: F* $(R)_{\leq}$ F**, F** $(R)_{\leq}$ F*, F* = F**. An example of a semi-ordering relation is the $\geq$ relation for distribution functions, because two functions $X(t)$ and $Y(t)$ exist such that for some values of t, $X(t) \geq Y(t)$, and for some other values of t, $Y(t) \geq X(t)$.

It can be seen from Eq. (1) that the yield is a Laplace-Stieltjes transform of the distribution F of defect density D, wherein the active area a can only assume real values. Therefore we can use a known semi-ordering relation $(L)_{\leq}$, as defined by the Laplace-Stieltjes transform [16]. We define $F_1 \ (L)_{\leq} \ F_2$ if

$$\int_0^{\infty} e^{-at} dF_1(t) \geq \int_0^{\infty} e^{-at} dF_2(t) \tag{2}$$

or $Y_{F_1}(a) \geq Y_{F_2}(a)$ for all a, $\quad 0 \leq a < \infty$.

(As mentioned earlier, a smaller defect density means a higher yield for all possible active areas.)

The relation $(L)_{\leq}$ has Properties 1 to 3 of a semi-ordering [16]. From known theorems [16] about the relation $(L)_{\leq}$, if the following two properties of yield, which are stated without proof, can be derived.

- *Property 1*

From $F_1(t) \leq F_2(t)$, $0 \leq t < \infty$, it follows that $F_2(t) \ (L)_{\leq} \ F_1(t)$ or equivalently $Y_{F_2}(a) \leq Y_{F_1}(a)$. A defect-density distribution which is smaller for all values of t than another distribution gives the lower yield. This property may be used to compare different distributions.

- *Property 2*

If we have $Y_{F_1}(a) \geq Y_{F_2}(a)$ for all a, $0 \leq a < \infty$, then it follows for the expected values of the defect densities that

$$\mathrm{E}\,[D_1] = \int_0^{\infty} t dF_1(t) \geq \mathrm{E}\,[D_2] = \int_0^{\infty} t dF_2(t),$$

provided that these expected values are finite. (A higher yield also means a smaller mean number of defects.)

Finally, it should be mentioned that a Laplace-Stieltjes transform can be written as a Laplace transform, which leads to the following expression for yield as being equivalent to Eq. (1):

$$Y_F(a) = a \int_0^{\infty} e^{-at} F(t) dt - F(0)$$

[with $F(0) = 0$ in this case].

Bounds for yield

Here we present some bounds for yield, i.e., the minimum and the maximum yield which can be achieved if the defect-density distributions belong to certain classes of distribution functions. Such bounds can be obtained in the following way. A lower bound $Y_{MIN}(a)$ requires minimization of the yield over all distributions F of a certain class K,

$$Y_{MIN}(a) = \min_{F \in K} \int_0^{\infty} e^{-at} dF(t). \tag{3}$$

Then we have, for all $F \in K$, $Y_{MIN}(a) \leq Y_F(a)$. The maximization of the same expression yields upper bounds $Y_{MAX}(a)$ for the yield. A few of these extremal distributions are known [16] and are presented in the following propositions. The proofs for Propositions 1–3 are based on the fact that under certain conditions the integral of Eq. (1) is maximized or minimized by distribution functions which are step functions with two or three jumps [17–19]. The proof is outlined in [20]. The proof for Proposition 4 makes use of a bound for distribution functions of the type NBUE (new better than used in expectation), which is given in [21, p. 188]. We use the following notation for the degenerate distribution:

$$\theta_m(t) = \begin{cases} 0 & \text{for } t \leq m, \\ 1 & \text{for } t > m. \end{cases} \tag{4}$$

- *Proposition 1*

Let K_1 be the set of all distribution functions F with expected value m,

$$m = \int_0^{\infty} t dF(t) = \mathrm{E}\,[D].$$

Then, θ_m minimizes the yield in class K_1. This gives a lower bound for yield with a given expected value of defect density

$Y_{MIN}(a) = e^{-am}$, or

$$Y_F(a) \geq e^{-am} \text{ for all } F \in K_1. \tag{5}$$

Several authors have observed that the yield given by Eq. (5) is too pessimistic [1–14]. It even represents a minimum yield which is achieved with constant defect density. That means that clustering of defects increases the yield if the mean remains fixed. This result is hardly surprising to a specialist in the yield field; but this author has never seen a rigorous proof of it in the yield literature. The upper bound equals 1 for functions in K_1. Therefore subclasses of K_1 have to be investigated to get nontrivial upper bounds. The minimum yield for the classes to be considered in what follows is again given by (5); i.e., we obtain nontrivial maxima, but no other minima.

• *Proposition 2*
Let $K_2 \subseteq K_1$ (given E $[D] = m$) be the set of all distribution functions which are 0 or 1 outside a finite interval (α, β), $F(\alpha) = 0, F(\beta + 0) = 1, 0 \leq \alpha < \beta < \infty$.

Then the yield is maximized by the following defect-density distribution F_{MAX_1}:

$$F_{MAX_1} = \frac{\beta - m}{\beta - \alpha}\theta_\alpha + \frac{m - \alpha}{\beta - \alpha}\theta_\beta . \tag{6}$$

The corresponding maximum yield is

$$Y_{MAX_1}(a) = \frac{\beta - m}{\beta - \alpha}e^{-a\alpha} + \frac{m - \alpha}{\beta - \alpha}e^{-a\beta} \tag{7}$$

or $Y_F(a) \leq Y_{MAX_1}(a)$ for all a and all $F \in K_2$.

For $m = \beta$, the lower and the upper bounds coincide, since θ_m is then the only element of K_2.

This bound is very good if the number of defects is small, since β is then not much bigger than m. The bound permits the following interpretation: The defect density which maximizes yield only assumes its minimum value α and its maximum value β with the proper probabilities. This represents a form of maximum clustering.

Other authors get approximate yield formulas of this type by mixing Poisson distributions [6, 12]. In [6], for example, it was even suggested that approximately 10^5 Poisson distributions be combined to obtain a better yield formula, i.e., to use in Eq. (1) a defect density distribution of the form

$$F(\lambda) = \sum_{i=1}^{N} \alpha_i \theta_{\lambda - \lambda_i}$$

with $N \approx 10^5$ and λ_i, α_i properly chosen.

• *Proposition 3*
Let K_3 be the subset of K_1 containing all distribution functions having a given variance σ^2. Then the following defect-density distribution maximizes yield:

$$F_{MAX_2} = p\theta_0 + (1 - p)\theta_{(m+\sigma^2/m)}, \tag{8}$$

with $p = \dfrac{\sigma^2}{\sigma^2 + m^2}$.

The corresponding maximum yield is

$$Y_{MAX_2}(a) = p + (1 - p)e^{-a(m+\sigma^2/m)} \tag{9}$$

or $Y_F(a) \leq Y_{MAX_2}(a)$ for all $F \in K_3$.

The interpretation is similar to that for F_{MAX_1}. The defect density again assumes only its extreme values 0 and $m + \sigma^2/m$ to achieve maximum clustering.

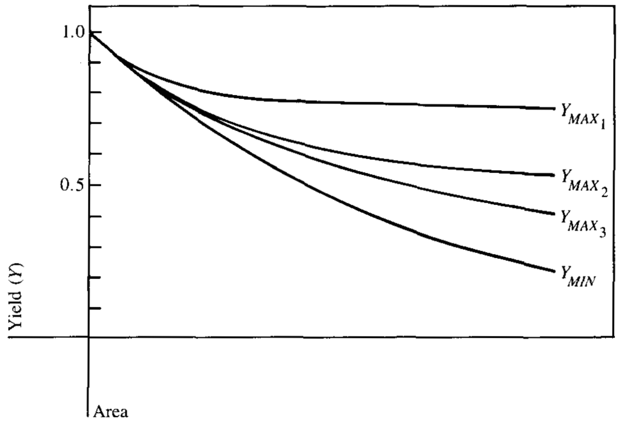

Figure 1 Bounds for yield as a function of area with normal variance. Here Y_{MAX_3} comes closest to Y_{MIN}.

The last upper bound requires some knowledge of reliability theory [16]. Let $F(t)$ be a distribution function, and define

$$F_\tau(t) = \begin{cases} \dfrac{F(t+\tau) - F(\tau)}{1 - F(\tau)} & \text{for } F(\tau) \neq 1, \\ \theta_0(t) & \text{for } F(\tau) = 1. \end{cases}$$

Then F is of type NBUE if

$$\int_0^\infty t dF_\tau(t) \leq \int_0^\infty t dF(t) = m. \tag{10}$$

In terms of reliability, F is the distribution of the duration of life until failure of a new unit, and F_τ is the distribution of the residual life of a unit of age τ. Equation (10) just means that the expected time to failure is greater for a new unit than for a unit already in use for a time τ.

The condition (10) is fulfilled for distributions where $\log(1 - F(t))$ is concave. Examples are gamma and Weibull distributions with $\alpha \geq 1$. Now we can formulate the next proposition.

• *Proposition 4*
Let K_4 be the subset of all distributions in K_1 of type NBUE. Then the yield is maximized by the defect-density distribution

$$F_{MAX_3}(t) = 1 - e^{-t/m}.$$

The maximum yield is

$$F_{MAX_3}(t) = \frac{1}{1 + am}, \tag{11}$$

and

$$Y_F(a) \leq \frac{1}{1 + am},$$

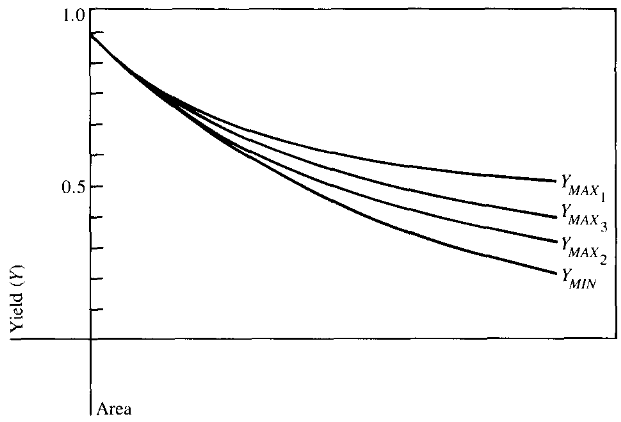

Figure 2 Bounds for yield as a function of area with small variance. Notice that now Y_{MAX_2} comes closest to Y_{MIN}.

for all $F \in K_4$. This formula was derived in [15] but not recognized as an upper bound. It also represents the special case $\alpha = 1$ for a formula proposed by Stapper [10–12]. For $\alpha > 1$, Eq. (11) always gives upper bounds for Stapper's yield because his defect-density distribution is a gamma distribution, and in this case is of type NBUE. This is no longer true for $\alpha < 1$, where Stapper's formula may give a higher yield than Y_{MAX_3}, as examples show.

Numerical examples

Two numerical examples follow. In Fig. 1, the lower bound and the three upper bounds are given as functions of the active area (linear scale) for the parameter values E $[D] = m = 2$, Var $[D] = \sigma^2 = 4$, $\alpha = 0$, and $\beta = 8$.

The variance is reduced in Fig. 2: Var $[D] = 1$, $\alpha = 0$, and $\beta = 4$. This leads to much narrower bounds. It is interesting to note that in Fig. 1, Y_{MAX_3} is nearest to Y_{MIN}, but in Fig. 2, Y_{MAX_2} comes closest to Y_{MIN}. This is due to the fact that Y_{MAX_2} approximated the yield better whenever the variance of the defect density is smaller. Therefore it is useful to have several upper bounds to choose from. Depending on the special case of interest, the "best" bound can be chosen.

References

1. B. T. Murphy, "Cost-Size Optima of Monolithic Integrated Circuits," *Proc. IEEE* **52,** 1537–1545 (1964).
2. W. G. Ansley, "Computation of Integrated-Circuit Yields from Distribution of Slice Yields for the Individual Devices," *IEEE Trans. Electron Devices* **ED-15,** 405–406 (1968).
3. A. G. F. Dingwall, "High-Yield-Processed Bipolar LSI Arrays," presented at the IEEE Int. Electron Devices Meeting, Washington, DC, Oct. 1968.
4. A. Gupta and J. W. Lanthrop, "Yield Analysis of Large Integrated-Circuit Chips," *IEEE J. Solid-State Circuits* **SC-7,** 389–395 (1972).
5. A. Gupta, W. A. Porter, and J. W. Lathrop, "Defect Analysis and Yield Degradation of Integrated Circuits," *IEEE J. Solid-State Circuits* **SC-9,** 96–103 (1974).
6. S. M. Hu, "Some Considerations in the Formulation of the IC Yield Statistics," *Solid-State Electron.* **22,** 205–211 (1979).
7. G. E. Moore, "What Level of LSI is Best for You?" *Electronics* **43,** 126–130 (Feb. 1970).
8. O. Paz and J. R. Lawson, Jr., "Modification of Poisson Statistics: Modeling Defects Induced by Diffusion," *IEEE J. Solid-State Circuits* **SC-12,** 540–546 (1977).
9. R. B. Seeds, "Yield and Cost Analysis of Bipolar LSI," presented at the IEEE International Electron Devices Meeting, Washington, DC, Oct. 1967.
10. C. H. Stapper, "Defect Density Distribution for LSI Yield Calculations," *IEEE Trans. Electron. Devices* **ED-20,** 655–657 (1973).
11. C. H. Stapper, "On a Composite Model to the IC Yield Problem," *IEEE J. Solid-State Circuits* **SC-10,** 537–539 (1975).
12. 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.* **24,** 398–409 (1980).
13. R. M. Warner, "Applying a Composite Model to the IC Yield Problem," *IEEE J. Solid-State Circuits* **SC-9,** 86–95 (1974).
14. T. Yanagawa, "Yield Degradation of Integrated Circuits Due to Spot Defects," *IEEE Trans. Electron Devices* **ED-19,** 190–197 (1972).
15. J. E. Price, "A New Look at Yield of Integrated Circuits," *Proc. IEEE* **58,** 1290–1291 (1970).
16. D. Stoyan, *Qualitative Eigenschaften und Abschätzungen Stochastischer Modelle,* Oldenburg, München-Wien, 1977.
17. B. Harris, "Determining Bounds on Expected Values of Certain Functions," *Ann. Math. Statist.* **33,** 1454–1457 (1962).
18. H. O. Hartley and H. A. David, "Universal Bounds for Mean Range and Extreme Observations," *Ann. Math. Statist.* **25,** 85–99 (1954).
19. W. Hoeffding, "The Extrema of the Expected Value of a Function of Independent Random Variables," *Ann. Math. Statist.* **26,** 268–275 (1955).
20. D. Stoyan, "Bounds for the Extrema of the Expected Value of a Convex Function of Independent Random Variables," *Studia Sci. Math. Hungarica* **8,** 153–159 (1973).
21. R. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing—Probabilistic Models,* Holt, Rinehart, and Winston, Inc., New York, 1975.

Received March 17, 1983; revised July 15, 1983

Bernd Werner Meister *IBM Research Division, Säumerstrasse 4, CH-8803 Rüschlikon, Switzerland.* Dr. Meister is a research staff member at the Zurich research laboratory, working on performance evaluation of local area networks in the Communications and Computer Science Department. He joined IBM in 1964 at the Zurich laboratory. Since then, he has worked at the Heidelberg Scientific Center from 1972 to 1974, in the Department for Mathematics and Computer Science, University of Stuttgart, as a visiting professor for six months during 1977, and in the same department at the University of Stuttgart as a lecturer since 1977. Prior to joining IBM, Dr. Meister worked as a scientific assistant at the Institute for Applied Mathematics and Mechanics at the University of Freiburg im Breisgau, Federal Republic of Germany, from 1962 to 1964. He received his M.S. in mathematics from Humboldt University, Berlin, Democratic Republic of Germany, in 1958 and his Ph.D. in mathematics from the University of Freiburg im Breisgau in 1962. He has published numerous papers on hydrodynamic stability theory, numerical analysis, queueing theory and applications, and performance evaluation.