A Differential Burn-in Policy Considering Nonhomogeneous Distribution of Spatial Defects in Semiconductor Manufacturing

Tao Yuan
Dept. of Industrial and Systems Engineering
Ohio University
Athens, Ohio, USA
yuan@ohio.edu

Yuan Chen
Dept. of Industrial and Systems Engineering
Ohio University
Athens, Ohio, USA
yc258117@ohio.edu

Yue Kuo
Dept. of Chemical Engineering
Texas A&M University
College Station, Texas, USA
yuekuo@tamu.edu

***Abstract*—This paper proposes a differential burn-in policy that considers the spatial nonhomogeneous distribution of defects in semiconductor manufacturing. Due to the nonhomogeneous distribution of spatial defects, devices at different locations on a semiconductor wafer may exhibit different probabilities of being defective. Unlike conventional burn-in policies, which subject all devices to the same burn-in test, the differential burn-in policy can take different actions for different devices, i.e., acceptance without burn-in, rejection without burn-in, or burn-in with a certain duration. A mixed integer nonlinear programming model is developed to find the cost-optimal decisions. A numerical example is used to demonstrate the potential application of the proposed burn-in policy.**

***Index Terms*—burn-in, defects, integrated circuits, optimization, semiconductor manufacturing, mixture distribution**

I. Introduction

Infant mortality has been widely recognized as a major issue in the semiconductor industry [1]. Defects generated in the complex fabrication processes can cause early failures of defective devices. Burn-in is an effective procedure to identify weak devices by running all devices under certain conditions for a suitable duration [2]. Devices that survive the burn-in test are shipped to customers. Because the burn-in test is costly, it needs to be carefully designed and optimized.

The determination of an optimal burn-in duration has been the subjects of numerous studies [1]–[3]. Various criteria, e.g., maximum mission reliability, maximum mean residual life, and minimum expected total cost, have been used to optimize the burn-in duration [2]. Burn-in populations are usually heterogeneous due to the existence of subpopulations, e.g., subpopulations for weak devices with defects and strong devices without defects, respectively. Mixture distributions have been frequently used as lifetime distributions for burn-in populations [4].

Conventional burn-in policies generally test all devices for the same duration. However, it has been commonly observed that defects on semiconductor wafers are not homogeneously distributed; instead they tend to cluster. Consequently, the defective devices are not homogeneously distributed, and the probability that a device being defective varies with the device's location on a wafer. This motivates the differential burn-in policy proposed in this paper. Under the proposed burn-in policy, devices at different locations may be subject to different burn-in decisions. This differential burn-in policy may be a more cost effective alternate to the conventional burn-in policies.

The remainder of this paper is organized as follows. Section II presents the proposed burn-in policy and the optimization model to find the cost-optimal burn-in decisions. Section III uses a numerical example to demonstrate the potential application of the proposed burn-in policy. Finally, Section IV concludes this paper and states possible future research work.

This research was partially supported by National Sciences Foundation projects 1633500 and 1633580.

II. Methodology

This section presents the proposed differential burn-in policy and an optimization model to find the cost-optimal burn-in decisions. For the purpose of comparison, we first establish a traditional cost-optimal burn-in model that tests all devices with a common burn-in duration, t_b. Then, we modify the first model to incorporate the spatial nonhomogeneous distribution of defects. The burn-in decisions may be different for devices at different locations on a wafer.

A mixture distribution involving two Weibull distributions is used to describe the failure-time distribution of a heterogeneous burn-in population with two subpopulations, i.e., a weak subpopulation consisting of devices with defects and a strong subpopulation consisting of devices without defects. The probability density function (PDF) and reliability function of this mixture distribution are, respectively, given by [5]

$$f(t) = p[f_E(t)R_I(t) + f_I(t)R_E(t)] + (1-p)f_I(t) \quad (1)$$

and

$$R(t) = R_I(t) - pR_I(t)(1 - R_E(t)), \quad (2)$$

978-1-7281-7102-9/20/$31.00 ©2020 IEEE

978-1-7281-7102-9/20/$31.00 ©2020 IEEE

where p is the probability of a device being defective, the subscripts "E" and "I" denote extrinsic and intrinsic failure modes, respectively, and $f_I(t)$ ($f_E(t)$) and $R_I(t)$ ($R_E(t)$) are the Weibull PDF and reliability function to describe the failure-time distribution of the intrinsic (extrinsic) failure mode. The Weibull reliability functions are

$$R_I(t) = \exp\left(-\left(\frac{t}{\theta_I}\right)^{\beta_I}\right) \text{ and } R_E(t) = \exp\left(-\left(\frac{t}{\theta_E}\right)^{\beta_E}\right) \quad (3)$$

for the two failure modes, where θ and β are, respectively, the Weibull scale and shape parameters. Extrinsic failures are early failures caused by defects; while intrinsic failures are caused by intrinsic device wearouts. Therefore, $0 < \beta_E < 1$, $\beta_I > 1$, and $\theta_I > \theta_E$ are expected. This mixture distribution assumes that a defective device in the weak subpopulation has both extrinsic and intrinsic failure modes competing to determine the failure time but a defect-free device in the strong subpopulation only has the intrinsic failure mode. It can be shown that this mixture distribution has a bathtub shaped hazard rate function [5].

An optimization model is developed to find the optimal burn-in duration that minimizes the expected total cost per unit given by

$$C_0(t_b) = c_{st} + c_{vt}t_b + c_{ft}[1 - R(t_b)] + c_{fw}[R(t_b) - R(t_b + t_w)], \quad (4)$$

where the burn-in duration t_b is the decision variable, and c_{st}, c_{vt}, c_{ft} and c_{fw} are, respectively, the fixed burn-in cost per unit, variable burn-in cost per unit per unit time, failure cost during burn-in, and failure cost during field operation within the warranty period t_w. For a given set of Weibull shape and scale parameters and cost coefficients, the optimal burn-in duration t_b^* certainly depends on the defective probability p.

It has been widely observed that defects are not uniformly distributed on a semiconductor wafer; instead, they tend to cluster. Therefore, it is reasonable to assume that the probability that a device has defects, p, in the mixture distribution (1) varies according to the device's location on the wafer. Let $\mathbf{z}_k$ denote the location of the kth device on the wafer and denote $p_k \equiv p(\mathbf{z}_k)$ the probability of the kth device being defective. Then the reliability function of the kth device is

$$R(t; \mathbf{z}_k) = R_I(t) - p_k R_I(t)(1 - R_E(t)). \quad (5)$$

The conventional burn-in cost model (4) may be modified to

$$C_1(t_b) = \frac{1}{K}\sum_{k=1}^{K}\{c_{st} + c_{vt}t_b + c_{ft}[1 - R(t_b; \mathbf{z}_k)] + c_{fw}[R(t_b; \mathbf{z}_k) - R(t_b + t_m; \mathbf{z}_k)]\}, \quad (6)$$

where K is the total number of devices.

Due to the spatial nonhomogeneous distribution of defects and defective devices, a differential burn-in policy, which chooses different burn-in actions for different devices, may be a more cost effective alternate to the conventional burn-in policies that test all devices with the same duration [6]. A proposed differential burn-in policy attempts to make one out of the following three decisions for each device: (1) rejection without burn-in; (2) acceptance without burn-in; and (3) burn-in with a duration $t_b > 0$. For each device, let us define two binary decision variables

$$x_k = \begin{cases} 0, & \text{rejection without burn-in} \\ 1, & \text{otherwise} \end{cases} \quad (7)$$

and

$$y_k = \begin{cases} 0, & \text{acceptance without burn-in} \\ 1, & \text{burn-in with the duration } t_b \end{cases} \quad (8)$$

for $k = 1, 2, \ldots, K$. The optimal burn-in policy will be the solution that minimizes the following expected total cost per unit

$$\begin{aligned} C_2(\mathbf{x}, \mathbf{y}, t_b) = \frac{1}{K}\sum_{k=1}^{K}\{ & c_{st}x_k y_k + c_{vt}t_b x_k y_k \\ & + c_{ft}[1 - R(t_b; \mathbf{z}_k)]x_k y_k \\ & + c_{fw}[R(t_b; \mathbf{z}_k) - R(t_b + t_w; \mathbf{z}_k)]x_k y_k \\ & + c_{fw}[1 - R(t_w; \mathbf{z}_k)]x_k(1 - y_k) \\ & + (1 - x_k)c_{ft}\} \end{aligned} \quad (9)$$

subject to the constraint

$$x_k - y_k \geq 0, \quad k = 1, 2, \ldots, K. \quad (10)$$

The constraint (10) ensures that $y_k = 0$ when $x_k = 0$. When a device is discarded without burn-in, i.e., $x_k = 0$, the total cost is the burn-in failure cost; when a device is accepted without burn-in, i.e., $x_k = 1$ and $y_k = 0$, the total cost includes the field failure cost; and when a device is tested with the duration t_b, i.e., $x_k = 1$ and $y_k = 1$, all four cost components are included in the expected total cost. The GEKKO Python package is used to solve this nonlinear mixed integer programming problem.

TABLE I
PARAMETERS IN THE NUMERICAL EXAMPLE

Distribution Parameters		Cost Coefficients	
β_E	0.5	c_{st}	\$10
θ_E	1,000 hours	c_{vt}	\$0.05
β_I	2	c_{ft}	\$200
θ_I	100,000 hours	c_{fw}	\$1,000

III. NUMERICAL EXAMPLE

This section presents a numerical example to compare the proposed differential burn-in policy with the conventional policy. Table I lists the parameters used in the numerical example. In addition, $t_w = 10{,}000$ hours. Under the assumed distribution parameters, $R_E(t_w) = 4\%$ and $R_I(t_w) = 99\%$. Therefore, it is necessary to weed out the weak units before sending the products to customers.

We first minimize the conventional cost model (4) to examine the effect of the defective probability p on the optimal burn-in decision. Figs. 1 and 2 show, respectively, the optimal burn-in duration and expected total cost per unit for different

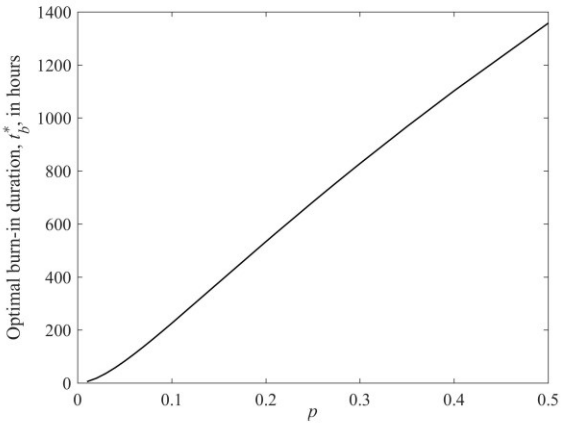

Fig. 1. Optimal burn-in duration t_b^* for different p.

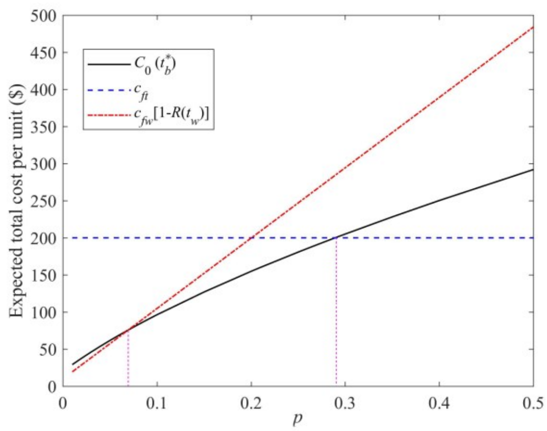

Fig. 2. Expected total cost $C_0(t_b^*)$ for different p.

values of p, ranging from 1% to 50%. The optimal burn-in duration and expected total cost increase as the defective probability, p, increases. If a device is discarded without burn-in the expected total cost is the burn-in failure cost, c_{ft}; on the other hand, if a device is accepted without burn-in the expected total cost is the field failure cost, $c_{fw}[1 - R(t_w)]$. Fig. 2 also depicts those two expected total costs for comparison. As shown in Fig. 2, when $p < 0.06$ the expected total cost per unit with burn-in, $C_0(t_b^*)$, is higher than the expected total cost per unit if the device is accepted without burn-in. Hence, a more economic decision would be the acceptance without burn-in when p is small. On the other hand, when p is large, i.e., $p > 0.28$, the expected total cost per unit with burn-in is higher than the cost if the device is discarded without burn-in, and a reasonable decision would be the rejection without burn-in. When p is between 0.06 and 0.28, burn-in would be the reasonable action.

Next, we use the wafer shown in Fig. 3 to illustrate the proposed differential burn-in policy. The diameter of the wafer is 20 cm and there are $K = 473$ dies (i.e., devices) fabricated

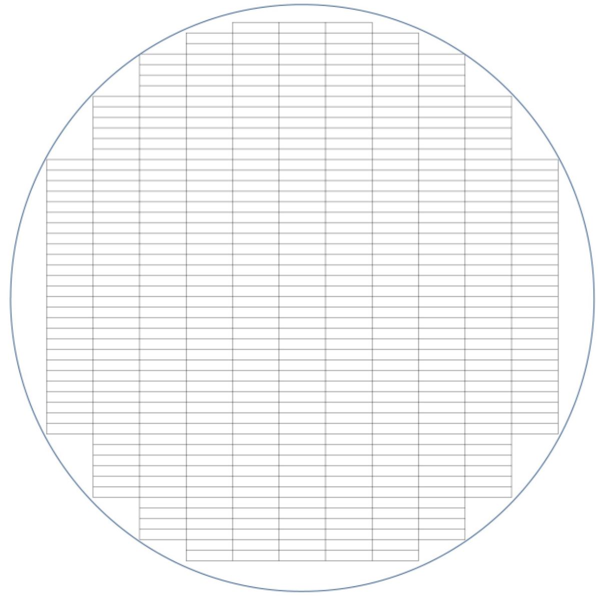

Fig. 3. The wafer used in the numerical example.

on the wafer. The number of defects on the kth devices is assumed to follow the Poisson distribution with a mean μ_k that is dependent on its location of the wafer. As an illustration, we assume [7]

$$\mu_k = \gamma_0 + \gamma_1 r_k + \gamma_2 r_k^2, \qquad (11)$$

for $k = 1, 2, \ldots, K$, where r_k is the distance from the wafer center to the center of the kth device in centimeters, and $\boldsymbol{\gamma} \equiv (\gamma_0, \gamma_1, \gamma_2)$ represents the coefficient vector. Then, the probability that the kth device is defective is given by

$$p_k = 1 - e^{-\mu_k}, \quad k = 1, 2, \ldots, K. \qquad (12)$$

If we assume $\boldsymbol{\gamma} = (0.01, 0.01, 0.001)$, the mean number of defects μ_k ranges from 0.01 to 0.18 and the defective probability p_k is between 0.01 and 0.187. All the p_k values are less than 28%. The optimal differential burn-in decision obtained by solving the proposed differential burn-in optimization model accepts 133 devices near the wafer center without burn-in and tests the remaining 340 devices with a duration of 295.72 hours. The expected total cost C_2 is \$96.17 per device. If all the devices are subject to the same burn-in duration, the optimal burn-in duration can be found by minimizing the expected total cost per unit given by the model (6). The optimal burn-in duration is 233.97 hours with an expected total cost per unit of C_1=\$98.19. The differential burn-in policy results in a lower expected total cost per unit than the conventional burn-in policy. Note that the optimal solution of the conventional burn-in model (6) is a feasible solution of the proposed differential burn-in model (9). The optimal expected total cost per unit of the model (9), therefore, is always less than or equal to that of the model (6).

Next, we modify the μ_k function as

$$\mu_k = \mu_k + 0.25 \boldsymbol{I}_{\{r_k > 8\}}, \qquad (13)$$

to simulate the case where an assignable cause produces a ring-shaped cluster of defects near the wafer edge. Herein $I_{\{.\}}$ is the indicator function. Under this assumption, there are 90 devices with defective probability, p_k, higher than 0.28. The conventional burn-in policy tests all units with a duration of 343.6 hours, resulting in an expected total cost per unit of $120.2. On the other hand, the proposed differential burn-in policy discards 90 devices near the edge of the wafer without burn-in, accepts 115 devices near the wafer center without burn-in, and tests the remaining 268 devices with a burn-in duration of 253.5 hours. The expected total cost per unit under the differential burn-in policy is $109.4, which is again lower than that of the conventional burn-in policy.

IV. Conclusion and Future Work

This paper proposed a differential burn-in policy for semiconductor manufacturing. Due to the spatial heterogeneity of the defective probability, devices at different locations are subject to different burn-in decisions. Numerical results have demonstrated that the proposed burn-in policy may be a cost-saving alternate over the conventional burn-in policies that test all devices for the same duration.

The devices that are subject to burn-in are tested for the same duration. In the future, the burn-in policy may be extended to allow the devices to have different burn-in times. In the current study, only package-level burn-in was considered. In future studies, both package-level and wafer-level burn-in tests can be included.

References

[1] W. Kuo and Y. Kuo, “Facing the headaches of early failures: a state-of-the-art review of burn-in decisions,” Proceedings of the IEEE, vol. 71, pp. 1257–1266, 1983.

[2] W. Kuo, W.-T. K. Chien, and T. Kim, Reliability, Yield, and Stress Burn-in. Norwell, MA: Kluwer Academic Publishers, 1998.

[3] X. Liu and T. A. Mazzuhi, “The optimal burn-in: state of the art and new advances for cost function formulation,”, in Recent Advances in Reliability and Quality in Design, H. Pham, Ed. London: Springer, 2008, pp. 137–182.

[4] G.-A. Klutke, P. C. Kiessler, and M. A. Wortman, “A critical look at the bathtub curve,” IEEE Transactions on Reliability, vol. 52, no. 1, pp. 125–129, 2003

[5] R. Degraeve, Ph. Roussel, G. Groeseneken, and H. E. Maes, “A new analytic model for the description of the intrinsic oxide breakdown statistics of ultra-thin oxides,” Microelectronics Reliability, vol. 36, pp. 1639–1642, 1996.

[6] J. Y. Hwang and W. Kuo, “Differential burn-in and reliability screening policy using yield information based on spatial stochastic processes,” unpublished.

[7] J. Y. Hwang, Spatial Stochastic Processes for Yield and Reliability Management with Applications to Nano Electronics. College Station, TX: Texas A&M University, 2004.