Appeared in *IEEE DFT'95*, Louisiana, Nov. 1995

The Effect of Spot Defects on the Parametric Yield of Long Interconnection Lines

Israel A. Wagner *
Computer Science Department
Technion, Haifa 32000, Israel

Israel Koren †
Department of Electrical and Computer Engineering
University of Massachusetts, Amherst, MA 01003

Abstract

The effect of non-catastrophic *(or* soft*) defects (i.e., neither* short *nor* open*) on long interconnection lines is analyzed and an estimate is derived for the frequency-dependent critical area for such lines. The analysis is based on a transmission-line model of interconnection lines, and the reflections caused by the defect are taken into account. This analysis results in an estimated prediction of the parametric yield, and a practical recommendation for a better jog insertion in VLSI routing.*

1 Parametric Yield

In conventional (catastrophic) yield prediction, the main question is: 'How many fault-free chips will be manufactured?' whereas for parametric yield prediction the question is 'How many chips will have an acceptable performance?'. The term "acceptable performance" may have different interpretations, e.g., frequency, power consumption, MTTF, etc. We will focus in this paper on the frequency interpretation, that is: we will try to answer the question: How many chips will be operational at a desired frequency in the presence of spot defects (assuming a known defect distribution)? In this manuscript we restrict our discussion to interconnection lines.

In most previous publications (e.g., [13], [14], [15]) it was assumed that significant parametric yield loss can only be caused by global process faults (e.g., mask misalignment and line width variations), while local process defects (pinholes, missing/extra metal) can only result in catastrophic faults. A detailed survey of existing methods in parametric yield analysis can be found in [17]. A preliminary analysis of the effect of spot defects on the circuit performance was made in [7], but the model used there does not take into account the effect of the discontinuities in the transmission lines caused by spot defects; hence the resulting reflections are not considered. Such reflections, which may have been negligible when CPU clock frequencies were below 100Mhz, have become more crucial with recent progress in technology and higher processing rates.

Our goal here is to investigate the effect of local process defects on the parametric yield of metal lines by analyzing the increase in the (signal propagation) delay of such lines due to spot defects. We extend the well-known notion of *critical area* and estimate the *delay-increase-dependent yield* $Y(\sigma)$, which is defined as the percentage of chips that will

*On temporary leave from IBM Israel - Science and Technology, Matam, Haifa 31905, Israel.
†Supported in part by NSF under contract MIP-9305912.

have their lines delayed by no more than σ times the predicted delay of defect-free lines. According to this definition, as σ goes to ∞, $Y(\sigma)$ approaches the traditional catastrophic-fault yield.

2 The Effect of an Extra Metal Defect on the Delay

Consider two parallel metal lines of length L separated by a distance S (see Figure 1(a)) and assume that an extra (missing) metal defect of size $2r \times 2r$ falls at distance x from the driver and its center is at height y from the lower line. Note that we model the defect as a square of side $2r$ rather than a circle of radius r. This defect touches line 2 but not line 1. Hence, this defect is not "catastrophic" in the regular sense since it does not cause a short (open) circuit. However, it does increase the coupling between the lines (the resistance of the line) and hence the delay, which may result in performance degradation, the amount of which depends on the criticality of the line and the parameters of the defect (size and location). The following is a preliminary analysis that is intended to give the flavor of this dependency, and is divided into two subsections. In the first one, the effect of a defect on the resistance and capacitance of a distributed RC line is estimated, while in the second subsection we derive an expression for the extra delay caused by a defect.

2.1 The Impact of a Defect on the RC characteristics

Any defect along an interconnection line is a discontinuity caused by the RC characteristics of the defect which are different from those of the rest of the line. We assume that an extra metal defect causes extra capacitance while a missing metal defect causes extra resistance. This assumption is justified by the following observation: the line resistance depends on the length-to-width ratio which, for a soft extra-metal fault, is always less than 2 (assuming line width which is approximately equal to the line-to-line spacing), while the line capacitance depends on the length-to-spacing ratio, which can grow arbitrarily as the spacing is decreased by the defect. On the other hand, in the case of missing metal, the line width (hence $1/R$) can decrease arbitrarily, while the line-to-line spacing (and $1/C$) will not even be doubled.

From now on we shall concentrate on the effect of extra metal defects for two reasons: these defects are more likely to occur (as has been found by many chip monitors), and their effect on the parametric yield is much higher. The total capacitance of a line is the sum of three terms

$$C_p + C_f + C_{ll} \tag{1}$$

where

- C_p – "parallel plates" capacitance
- C_f – "fringe" capacitance
- C_{ll} – "line to line" capacitance

In the usual case where the lines are minimally spaced, C_{ll} is the dominant term in (1). The exact value of C_{ll} for a given geometry is not easy to calculate. A detailed study of the characteristics of coupled microstrips ([9],[11]), reveals that when S is small, the most significant component in the coupling capacitance (also termed "gap capacitance") per unit length is

$$C_{ll} = 2\epsilon_{ox} \cdot t/S \tag{2}$$

where ϵ_{ox} is the electric permitivity coefficient of the insulator that separates the metal from the substrate (usually Silicon-Oxide), and t is the height of the metal lines above the substrate (see Figure 2 for an illustration). In [9] and [11] the formulas are much more complicated since they are more general; here we are interested only in the component of the line-to-line capacitance, which is dominant when S is very small; hence we are mostly concerned with the $1/S$ component, as in equation (2). In our specific case, modeling the defect as a square stub, we can use equation (2) to approximate the extra capacitance due to the defect, denoted by C_d, in the following way. The neighborhood of the defect may be considered as a segment of the line with a significantly smaller spacing $S' = S + W - y - r$ (see Figure 1). The length of this segment is $2r$. We therefore obtain

$$C_d(y,r) = \begin{cases} C\left(1 + \frac{2r}{S+W-y-r}\right) & \text{if } r \geq y - W \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where C is the typical coupling (crosstalk) capacitance per unit length, which may be calculated using equation (2). Note that when there are no defects (i.e., $r = 0$) we get $C_d = C$ as expected.

2.2 The Delay Caused by a Defect

As we are concerned with on-chip lines, the inductance is negligible and the line can be modeled as an RC transmission line (see Fig. 1). We consider a defect along a line to be a discontinuity in its impedance. The impedance variation is estimated as a function of the defect size (r) and location (x and y). We then calculate the extra delay caused by reflections from the discontinuity back to the driver.

We assume that the change in the resistance of the line is negligible and focus on the extra capacitance caused by the coupling to the other line. The impedance due to the defect is $Z_d = R + 1/j\omega C_d$ rather than the typical $Z_0 = R + 1/j\omega C$, and the reflection coefficient is (see for example, [8],[12],[16])

$$\Gamma_d = \frac{Z_0 - Z_d}{Z_0 + Z_d}$$

which, for typical values of R and C, may be approximated as

$$\Gamma_d \approx \frac{C_d/C - 1}{C_d/C + 1}$$

and using the approximation for this extra capacitance in equation (3) above we obtain

$$\Gamma_d(y,r) = \begin{cases} \frac{r}{S+W-y} & \text{if } r \geq y - W \\ 0 & \text{otherwise} \end{cases}$$

Note that if $r < y - W$ the defect does not touch the line and the impedance is unchanged, hence there are no reflections at all and $\Gamma_d = 0$. Also note the extreme case of $r = S+W-y$, where the defect touches both lines. In this case we obtain $\Gamma_d(y, S + W - y) = 1$; that is: everything is reflected and nothing can pass through when the defect short-cuts the two wires. Hence, the *short-circuit* defect can be treated as a special case of our reflection analysis.

In order to estimate the effect of the spot defect on the delay, we should first address a more general question: *How does a discontinuity with reflection coefficient* Γ *affect the delay?* To answer this, one should first note that when a voltage wave V arrives at the barrier, only $(1-\Gamma)\cdot V$ of it is transferred to the receiver, while the other $\Gamma\cdot V$ is reflected back to the driver. This reflected wave travels the distance of $2x$ to the driver and back, and is then divided again between $(1-\Gamma)\Gamma\cdot V$ that go to the receiver and $\Gamma^2\cdot V$ that is reflected back, etc (see Figure 1). V_n, the receiver voltage after n reflections, is at most the sum of voltages that have arrived so far (we say "at most" since the precise value depends on the phase differences between the arriving waves). Hence the receiver voltage after the nth reflection is bounded from above by

$$\begin{aligned} V_n &\leq V_{n-1} + (V_{in} - V_{n-1})(1-\Gamma) = \Gamma V_{n-1} + (1-\Gamma)V_{in} \\ &\leq \Gamma^2 V_{n-2} + (1+\Gamma)(1-\Gamma)V_{in} \\ &\cdots \\ &\leq \Gamma^n V_0 + (1+\Gamma+\cdots+\Gamma^{n-1})(1-\Gamma)V_{in} = V_{in}(1-\Gamma^n) \end{aligned}$$

where V_{in} is the full voltage sweep of the pulse. Note that V_0, the receiver voltage before any reflection, is equal to zero. If the receiver is a standard CMOS circuit, we can then assume that in order to switch the receiver, its input voltage should be at least $V_{in}/2$. The (minimal) required number of reflections is determined by the inequality $V_n \geq V_{in}/2$ and depends on the size of the defect and its vertical position y. This number is denoted by $n(y,r)$ and is given by

$$n(y,r) = \left(\log\left(\Gamma_d^{-1}(y,r)\right)\right)^{-1}\log 2 = \frac{\log 2}{\log\left(\frac{S+W-y}{r}\right)} \tag{4}$$

Each reflection causes a "trip" to the driver and back. The time of such a trip can be approximated ([16], p. 199) by (line resistance) $\times$ (line capacitance). Since R and C are the typical line resistance and capacitance, respectively, we get that one trip from the defect to the driver and back takes about $(Rx)(Cx) = RCx^2$. The total delay caused by these reflections is approximately $T_d = 2\cdot n(y,r)\cdot RCx^2$ where x is the distance of the defect from the driver's end of the line and RCx^2 is the propagation delay from 0 to x. Note our assumption that the defect has about the same sheet resistance and real/fringe capacitance as the line. This is not necessarily true, especially for defects caused by embedded particles. But, as can be seen from equation (4), the dependence of the delay on Z_d, the impedance of the defect, is only logarithmic, the only exception being if $Z_d << Z_0$, when the denominator is close to zero. One may conclude that the sensitivity of the delay to the defect impedance is quite small as long as Z_d is not very small compared to Z_0. To see this, observe that T_d depends on $\left(\log\frac{Z_0+Z_d}{Z_0-Z_d}\right)^{-1}$ and let $Z_d = \alpha Z_0$ where α satisfies $0 \leq \alpha \leq 1$. Then we have $T_d = T_0/\log\frac{1+\alpha}{1-\alpha}$ where T_0 is a constant and hence

$$\frac{\partial T_d}{\partial \alpha} = \frac{2T_0}{(1-\alpha^2)\log^2\left(\frac{1+\alpha}{1-\alpha}\right)} \approx -\frac{2T_0(1-\alpha)}{4\alpha^2(1+\alpha)}$$

Denote by σ the *delay increase coefficient*, that is: the ratio of the extra delay caused by a defect to the propagation delay of the defect-free line, denoted by T_p

$$\sigma = \frac{T_d}{T_p} \approx \frac{2\cdot n(y,r)\cdot RCx^2}{RCL^2} = 2n(y,r)\cdot(x/L)^2$$

We wish σ to be as small as possible. The delay increase coefficient must be smaller than some maximum value σ_m in order to allow the chip to operate at the desired frequency. The relation between the maximum operational frequency f_m and the maximum delay increase coefficient σ_m depends on the specific context. If interconnection line is on a critical path, we have

$$f_m = \frac{1}{t_m} = \frac{1}{(1+\sigma)t_0} = \frac{f_0}{1+\sigma}$$

where $f_0 = 1/t_0$ is the maximum possible operating frequency of the defect-free circuit. σ must therefore satisfy the inequality

$$\sigma = 2n(y,r)(x/L)^2 = \frac{2(x/L)^2 \log 2}{\log\left(\frac{S+W-y}{r}\right)} \leq \sigma_m$$

or

$$\log\left(\frac{S+W-y}{r}\right) > \frac{1.386(x/L)^2}{\sigma_m}$$

To satisfy this condition the size of the defect must satisfy the following inequality:

$$r < r_{max}(x,y,\sigma_m) = (S+W-y)e^{-\frac{1.386(x/L)^2}{\sigma_m}} \tag{5}$$

Note that inequality (5) has an interesting corollary, which can be expressed as the following layout design rule:

> "If you *must* introduce a discontinuity (e.g., a jog) in a long wire, preferably do it near the driving end (smaller value of x) rather than near the receiving end."

Hence the wiring in Figure 3(a) is better than the one in Figure 3(b).

Based on (5), we can now define the *delay dependent critical area*, denoted by $A_c(\sigma)$, as the area in which a defect will cause an extra delay (measured by σ) which may prevent the chip from operating at the desired frequency. This critical area is defined as follows:

$$A_c(\sigma) = \int_{x=0}^{L} \int_{y=0}^{2W+S} \int_{r=r_{max}(x,y,\sigma)}^{\infty} D(r)\, dr\, dy\, dx \tag{6}$$

where $r_{max}(x,y,\sigma)$ is defined in (5) above, and $D(r)$ is the density function of the defect size. Experimental data on defects in many wafers lead to the following formula ([1], [2], [3], [5]):

$$D(r) = \begin{cases} r/2r_0^2 & \text{if } 0 \leq r \leq r_0 \\ r_0^2/r^3 & \text{if } r_0 \leq r \leq \infty \end{cases} \tag{7}$$

where r_0 is the peak of this distribution.

Following the analysis presented in [6], we can rewrite equation (6) to obtain

$$A_c(\sigma) = \int_{x=0}^{L} S(x,\sigma)\, dx$$

where $S(x,\sigma)$ is the defect sensitivity at point x along the line and is defined as

$$S(x,\sigma) = \int_{y=0}^{2W+S} \int_{r=r_{max}(x,y,\sigma)}^{\infty} D(r)\, dr\, dy \tag{8}$$

This integral is not easy to calculate analytically, hence we do it by numerical integration. See Figure 4 for $S(x,\sigma)$ vs. x for several values of σ. The following values were used for the parameters: $R = 0.02\Omega/\mu$, C=0.1fF/μ, $r_0 = 0.2\mu$, $W = 1.0\mu$ and $L = 5000\mu$. From this figure one may conclude that in order to reduce the defect sensitivity of critical nets (i.e., nets for which a small value of σ is desirable, e.g., clock trees), it may be worthwhile to increase the line-to-line spacing for the second half of the line (if $\sigma_m = 0.02$) or even at a distance of 600μ from the driver (in case σ_m is as small as 0.01). Such a "tapering" may reduce the sensitivity of the net to *soft* manufacturing defects.

Another function of interest is the The estimated yield is $Y(f) \approx e^{-D \cdot Ac(\sigma(f))}$ and is plotted as a function of the frequency in the left side of Figure 5, assuming that the maximum working frequency of the interconnection line is $f_0 = 500$Mhz, with 3 different spacing values. From this figure we conclude that the separation between long lines should be much larger for frequencies beyond $0.4f_0$, while in lower frequencies one may stick to the minimum spacing allowed by the technology ground-rules. In the right side of Figure 5 we show how the estimated yield depends on σ for various values of C_d/C, the ratio of the defect characteristic capacitance to that of the line. It can be seen that highly-capacitive defects (e.g. those caused by embedded particles) reduce the yield significantly.

It is interesting to note the two asymptotic lines: the first is the predicted yield for unlimited delay increase, $Y(\sigma \to \infty)$, which is the traditional yield. The other asymptotic line is $\sigma = 0$, i.e., no increase in delay is allowed and the yield goes to 0.

Note also that our analysis refers to lines which are implemented in one layer; in practice, this is not always the case and multi-layer nets are quite common. In some cases it may make sense to consider using upper metal layers to reduce the defect sensitivity for critical nets.

3 Summary and Future Research Directions

The effect of a spot defect on the delay of interconnection lines has been analyzed for the case of long parallel lines. The main effect of such a defect is the increase in the crosstalk capacitance between the lines. This effect was shown to be stronger near the receiving end of the line, using reflection analysis. An expression was derived for $r_{max}(x,y,\sigma)$, the maximum defect-size that can be tolerated at point (x,y) if the delay increase coefficient should not exceed σ. Using numerical integration, the functions $S(x,\sigma)$ (defect sensitivity at distance x from the driver) and $Y(\sigma)$ (estimated yield) were calculated and plotted. This yield estimate is based on a novel generalization of the classical critical area A_c that has been extended to be a function $A_c(\sigma)$ of the desired performance (maximum allowed delay).

In addition to a parametric, delay-dependent, yield estimate for this case, our analysis may lead to a possible cure for this extra sensitivity, by modifying the layout of interconnection lines from drivers to receivers. From our analysis it can be seen that a straight line is not always the best way to connect two points, once yield is taken into account. An interesting problem is thus to find the best layout for two adjacent interconnection lines for minimum delay.

The analysis reported in this paper also shows the advantage of using tapered clock trees, and suggests that other critical nets should use tapered interconnects in order to increase their yield. The significance of the yield is clearly highly dependent on the parameters of the technology. As the feature size (and the wiring pitch) decreases and the frequency

increases, the effect of potential spot defects on the parametric yield of the circuit increases.

Our analysis is also related to the problem of preserving the signal integrity. This problem will become more and more crucial as technology progresses and interconnection lines become narrower than ever. To manage these signal problems, several CAD tools have been suggested (see [10] for a popular introduction). However, as far as we know, no tool has been developed that takes into account the reflective effect of spot defects along interconnection lines. Our analysis may be a first step in this direction.

References

[1] A.V. Ferris-Prabhu, "Defect Size Variations and Their Effect on the Critical Area of VLSI Devices," *IEEE Journal of Solid State Circuits,* vol. sc-20, pp. 878-880, 1985.

[2] I. Koren, "The Effect of Scaling on the Yield of VLSI Circuits," *Yield Modeling and Defect Tolerance in VLSI Circuits,* W.R. Moore *et al.* (Eds.), pp. 91-99, Adam Hillger Ltd., 1988.

[3] I. Koren and A.D. Singh, "Fault Tolerance in VLSI Circuits," *IEEE Computer Magazine,* pp. 73-83, July 1990.

[4] C.H. Stapper, "Modeling of Integrated Circuit Defect Sensitivities," *IBM J. Res. Dev.,* vol. 27, no. 6, pp. 549-557, 1983.

[5] C.H. Stapper, "Modeling of Defects in Integrated Circuits Photolithographic Patterns," *IBM J. Res. Dev.,* vol. 28, no. 4, pp. 461-475, 1984.

[6] I.A. Wagner and I. Koren, "An Interactive VLSI CAD Tool for Yield Estimation," *IEEE Transactions on Semiconductor Manufacturing,* vol. 8, Special Issue on Defect, Fault, and Yield Modeling, pp. 130-138, May 1995.

[7] J. Zubairi, G.L. Craig, "A Bottom-Up Methodology to Characterize Delay Faults," *Proc. of DFT'91 - International Workshop on Defect and Fault Tolerance in VLSI Systems,* pp. 183-186, 1991.

[8] A.K. Goel, *High-Speed VLSI Interconnections,* John Wiley & Sons, New York, 1994.

[9] R. Garg and I.J. Bahl, "Characteristics of Coupled Microstriplines," *IEEE Transactions on Microwave Theory and Techniques,* vol. MTT-27, no. 7, July 1979.

[10] R. Goyal, "Managing Signal Integrity," *IEEE Spectrum,* March 1994.

[11] E.T. Lewis, "An Analysis of Interconnect Line Capacitance and Coupling for VLSI Circuits," *Solid-State Electronics,* vol. 27, nos. 8/9, pp. 741-749, 1984.

[12] R.E. Matick, *Transmission Lines for Digital and Communication Networks,* McGraw-Hill, New-York, 1969.

[13] D.M.H. Walker and S.W. Director, "VLASIC: A Catastrophic Fault Yield Simulator for Integrated Circuits," *IEEE Transactions on CAD,* vol. 5, no. 4, pp. 541-556, October 1986.

[14] D.E. Hocevar, P.F. Cox and P. Yang, "Parametric Yield Optimization for MOS Circuit Blocks," *IEEE Transactions on CAD,* vol. 7, no. 6, pp. 645-658, June 1988.

[15] D.M.H. Walker, "Tolerance of Delay Faults," *Proc. of DFT'92 - International Workshop on Defect and Fault Tolerance in VLSI Systems,* pp. 207-216, 1992.

[16] H.B. Bakoglu, *Circuits, Interconnections, and Packaging for VLSI,* Addison Wesley, 1990.

[17] R. Spence and R.S. Soin, *Tolerance Design of Electronic Circuits,* Addison Wesley, 1988

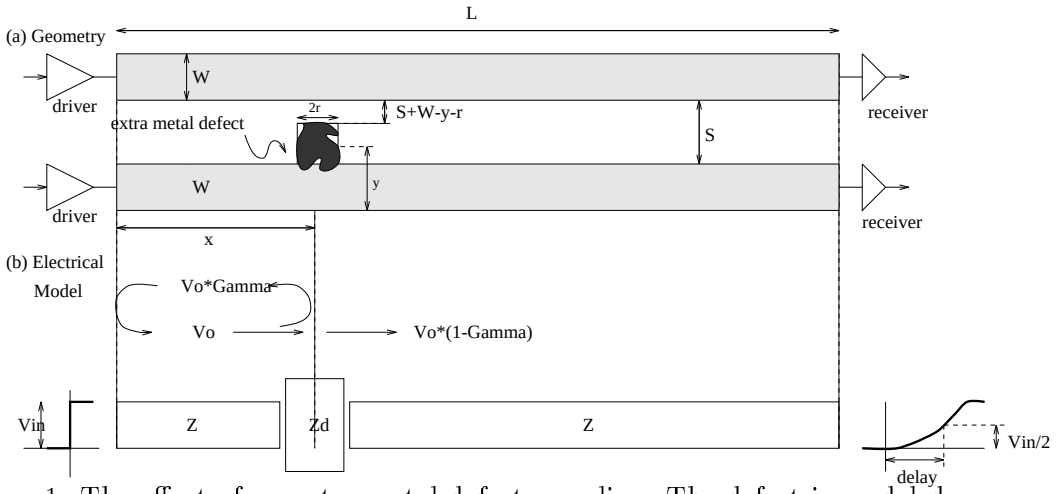

Figure 1: The effect of an extra metal defect on a line. The defect is modeled as a square of side $2r$.

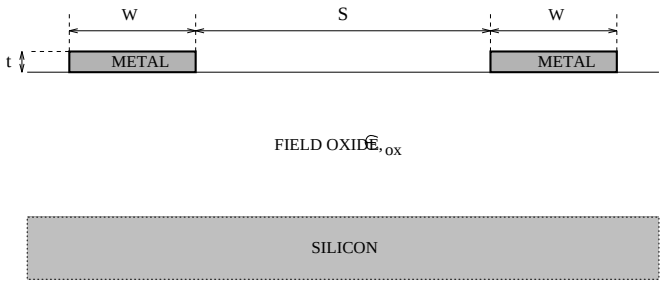

Figure 2: The microstrip parameters as seen in a vertical cut.

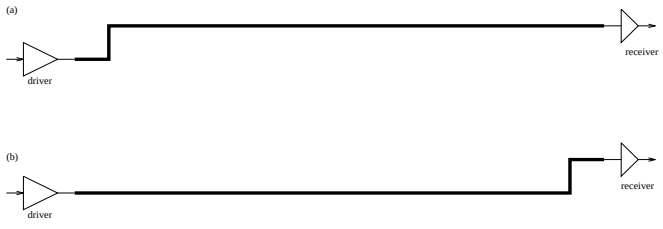

Figure 3: Two ways to add a jog that differ in their predicted yield.

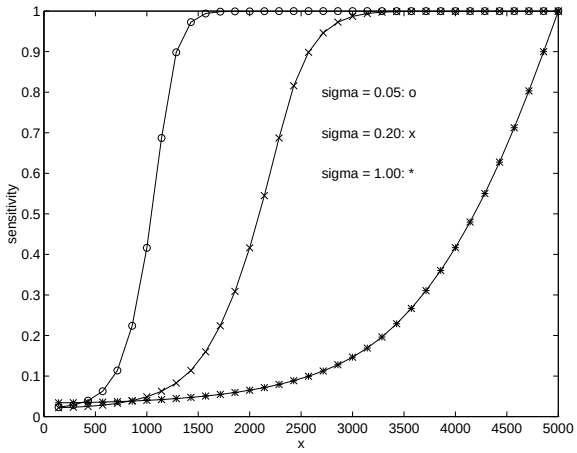

Figure 4: Defect sensitivity, $S(x, \sigma)$, vs. the distance x from the driver, for various values of the delay increase coefficient σ.

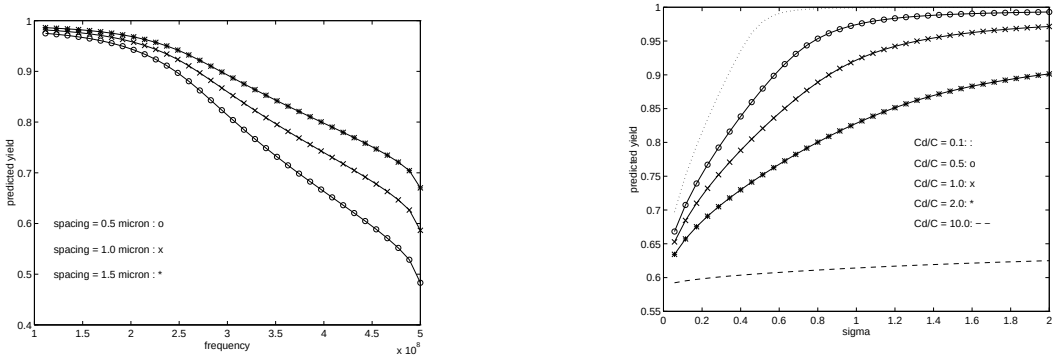

Figure 5: Yield vs. frequency for various values of the inter-line spacing S (left), assuming that f_0 (the maximum working frequency of the line) is 500Mhz. Note the significance of the spacing as the frequency exceeds $0.4f_0$. In the right figure, Yield vs. σ is plotted for various values of C_d/C, the defect capacity variation.