A Defect Model of Reliability

C. Glenn Shirley
Intel Corp.

1995 International Reliability Physics Symposium

Outline

- Reliability Statistics
- Defect Reliability
  - Relationship between yield and reliability
- Accelerated Stressing and Burn-In
- Analysis of Reliability Data
  - Test Flows
  - Model Extraction
- Reliability Prediction
  - Effect of Die Area
  - Effect of Defect Density
  - Effect of Burn In
  - Standard Reliability Indicators

Reliability Statistics

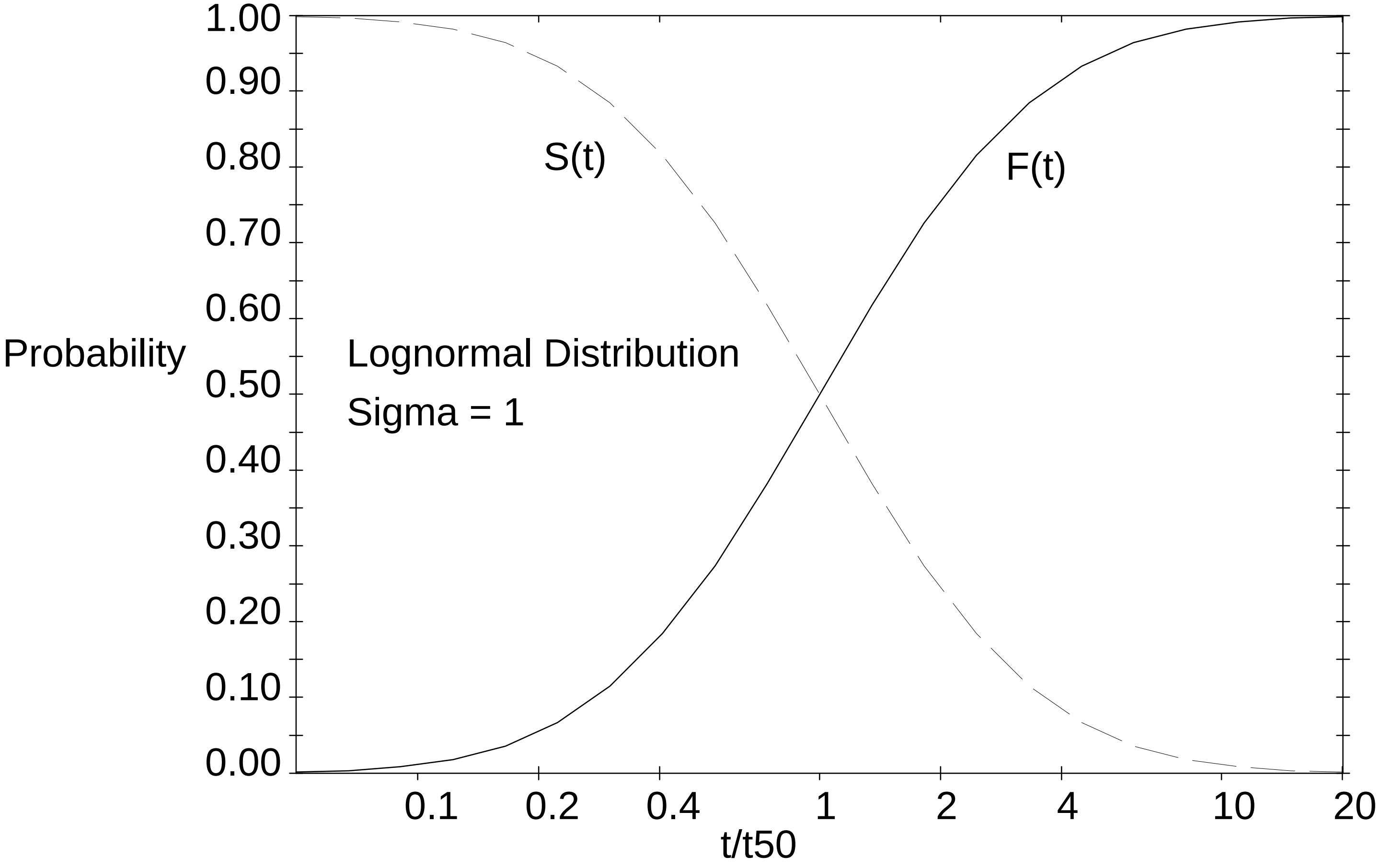
- Several mathematical functions are used to describe the evolution of a population.

- Cumulative distribution function $F(t)$:
  - Probability that a unit from original population fails by time t
  - $F(t=0) = 0$, $F(t=\text{infinity}) = 1$, $F(t)$ increases monotonically, $F(t)$ undefined for $t < 0$. $0 < F(t) < 1$.

- Survival function $S(t) = 1 - F(t)$:
  - Probability that a unit from original population survives to time t.
  - $S(t=0) = 1$, $S(t=\text{infinity}) = 0$, $S(t)$ decreases monotonically, $S(t)$ undefined for $t < 0$. $0 < S(t) < 1$.

Reliability Statistics

Cumulative Distribution Function and Survival Function



Reliability Statistics

- Probability density function, *f*(*t*)

$$f(t) = \frac{\text{Number of failures in } dt}{dt} \times \frac{1}{\text{Initial Population}}$$

$$f(t) = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt}$$

$$F(t) = \int_0^t f(t)\,dt$$

  – Of theoretical interest only.

Reliability Statistics

- Instantaneous Failure Rate, $h(t)$

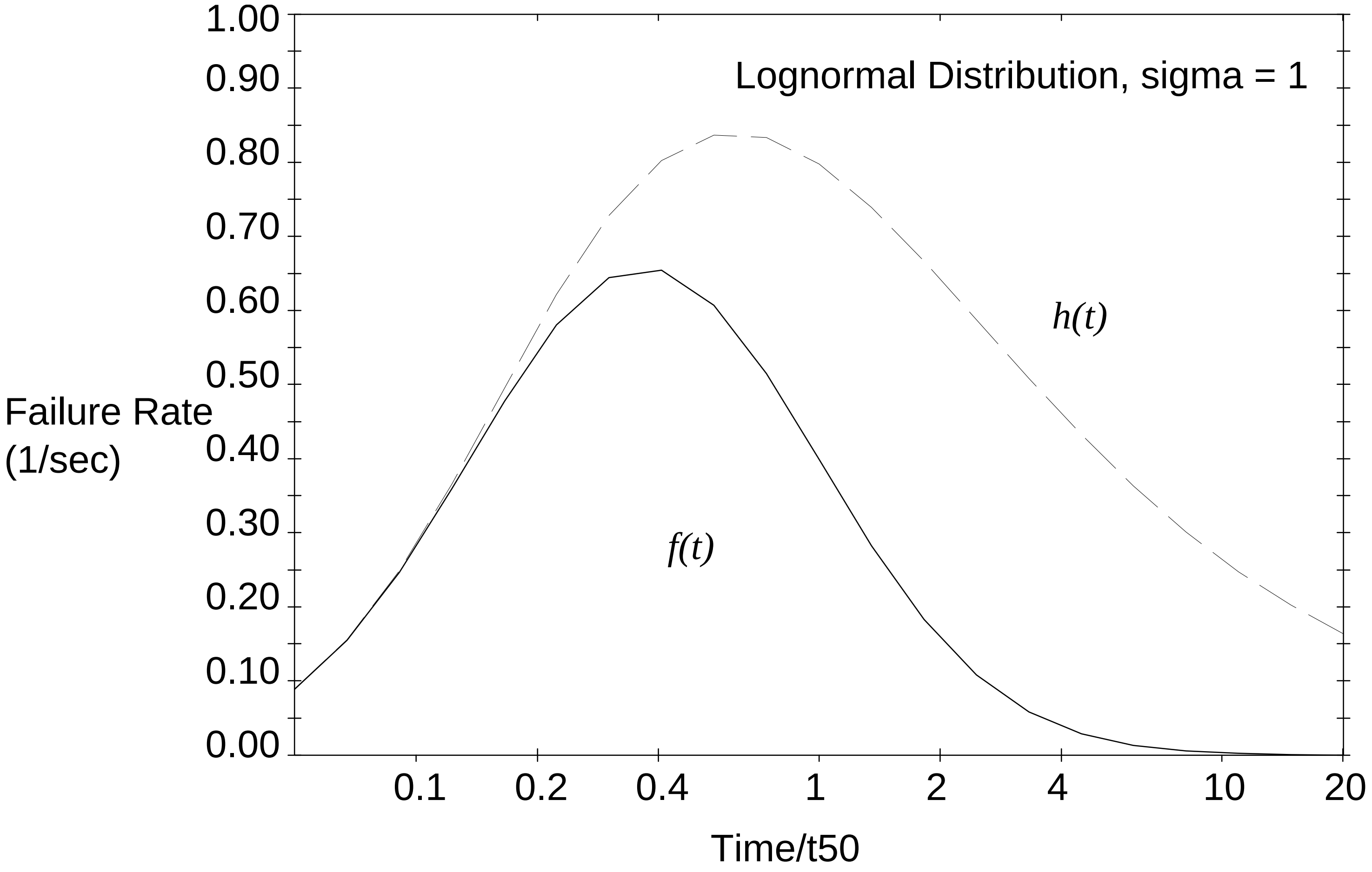
$$h(t) = \frac{\text{Number of failures in } dt}{dt} \times \frac{1}{\text{Population at time } t}$$

$$h(t) = \frac{f(t)}{S(t)} = -\frac{1}{S(t)}\frac{dS(t)}{dt} = -\frac{d\ln S(t)}{dt}$$

  - $h(t)$ can increase or decrease and have any positive value, that is, $h(t) > 0$.

Reliability Statistics

Probability Density Function *f(t)*, and Instantaneous Failure Rate *h(t)*



Reliability Statistics

- Cumulative Hazard Function, $H(t)$
- Defined by

$$H(t) = \int_0^t h(t)dt$$

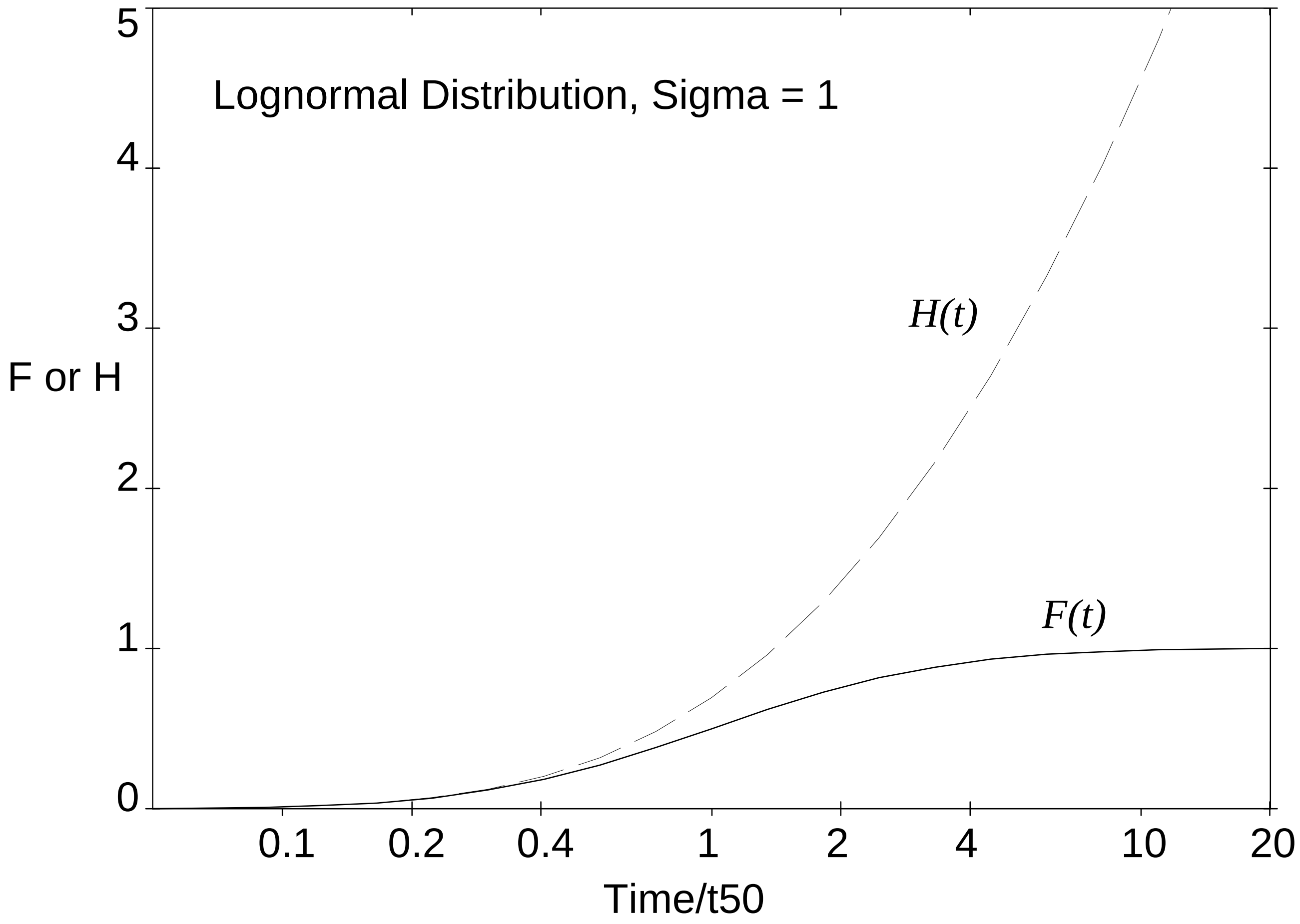
$$S(t) = \exp[-H(t)]$$

$$F(t) = 1 - \exp[-H(t)]$$

  - $H(t)$ is dimensionless, like a probability, but can have any positive value.
  - $H(t)$ increases monotonically with time.
  - $H(t)$ is useful in analysis of “censored” data in which removals or multiple failure mechanisms occur.

Reliability Statistics

Cumulative Hazard, $H(t)$, and Cumulative Distribution Function, $F(t)$



Reliability Statistics

- The functions $F(t)$, $S(t)$, $f(t)$, $h(t)$, $H(t)$ are all interrelated. Given one, the others can be derived.
- No assumptions about the specific distribution (Weibull, Lognormal, etc. have been made).
- A program for extracting models from censored data is
  - Plot $H(t)$ from censored data
  - Determine $F(t)$ via $F(t) = 1 - \exp[-H(t)]$
  - Fit parametric distribution to $F(t)$
  - Use parametric $S(t) = 1 - F(t)$ to calculate predictions.

Reliability Statistics

- Average Failure Rates: A common reliability indicator
  - The average failure rate between times t_1 and t_2

$$\mathrm{AFR}(t_1, t_2) = \frac{\int_{t_1}^{t_2} h(t)dt}{t_1 - t_2} = \frac{H(t_1) - H(t_2)}{t_1 - t_2}$$

$$= \frac{\ln S(t_1) - \ln S(t_2)}{t_1 - t_2}$$

  - For t_1 and t_2 in hours, multiply AFR by 10^9 to get units of Fits.
  - For t_1 and t_2 in hours, multiply AFR by 10^5 to get units of %/1khr.

Reliability Statistics

- Cumulative Fraction Failed: Another indicator
  - Fraction failing between t_1 and t_2

$$\text{Cum Fail} = F(t_2) - F(t_1) = S(t_1) - S(t_2)$$

  - If t_1 = 0 then

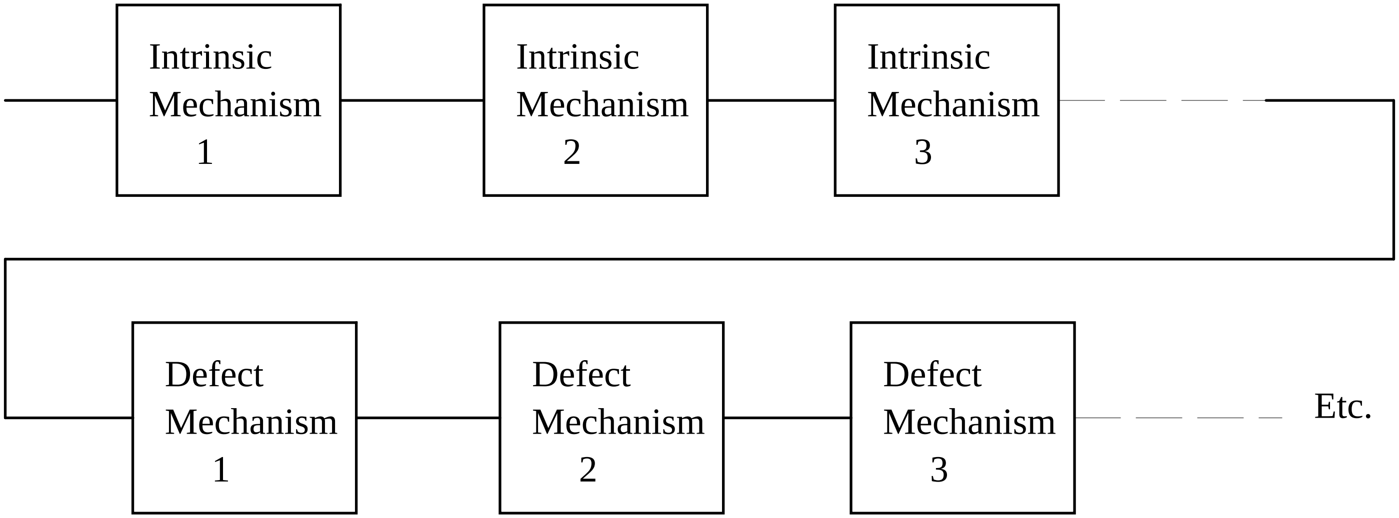
$$\text{Cum Fail} \;=\; F(t_2) = 1 - S(t_2)$$

  - Multiply Cum Fail by 10^6 to get DPM (Defects per Million)

- All indicators can be expressed in terms of the Survival Function.

Reliability Statistics

- Multiple failure mechanisms
  - If the earliest occurrence of a mechanism is fatal, then the device is logically a chain:



  - This is the usual case for semiconductor components. That is, there is no *functional* redundancy.

Reliability Statistics

- Multiple failure mechanisms (cont.)
  - The survival probability for a chain is the product of the survival probabilities of the links:

$$S(t) = S_{\text{mech 1}}(t) \times S_{\text{mech 2}}(t) \times \ldots$$

$$= \prod_{\text{mech } i} \exp[-H_i(t)] = \prod_{\text{mech } i} \exp[-\int_0^t h_i(t')dt']$$

$$= \exp[-\int_0^t \sum_i h_i(t')dt'] \equiv \exp[-\int_0^t h(t')dt'] \equiv \exp[-H(t)]$$

  - All that means is that the total instantaneous failure rate is the sum of instantaneous failure rates for each mechanism.

$$h(t) = \sum_{\text{mechanisms } i} h(t_i)$$

Intrinsic versus Defect Mechanisms

- Intrinsic mechanisms are due to non-defect-related manufacturing or design errors.
  - – Typically associated with gross areas of the wafer.
- The total survival function may be written

$$
\begin{aligned}
S(t) = {} & S_{\text{intrinsic mech 1}}(t) \times S_{\text{intrinsic mech 2}}(t) \times \ldots \\
& \times S_{\text{defect mech 1}}(t) \times S_{\text{defect mech 2}}(t) \times \ldots
\end{aligned}
$$

- The focus in this tutorial is on defect-related mechanisms.
  - – These are the main concern in the manufacturing environment.

Defect Reliability

- Factory production reliability issues are dominated by *defects*.
- The same kinds of defects that degrade yield, degrade reliability.
  - Yield is measured before any stress: At “Sort” (wafer-level functional test) and pre-burn-in class test.
  - Reliability is measured by post-burn-in class test.
- Since the “yield” and “reliability” defects are from the same source, yield and defect reliability are related.
- Yield is routinely measured - it can be used to predict reliability.
- Yield fallout is easier to measure than reliability fallout: It is larger.

Defect Size Distribution

- Establish distribution by visual counting and classifying particles and other defects in the factory.
- $D(x)$ is the *observed* number of defects per unit area with dimension (eg. diameter) between x and $x+dx$
  - For example, Stapper's model*

$$D(x) = \overline{D} \times (x / x_0^2) \quad \text{for} \quad x \le x_0$$

$$D(x) = \overline{D} \times (x_0^2 / x^3) \quad \text{for} \quad x > x_0$$

  - x_0 is a characteristic length << lithographic resolving power. (Operators can't see very small defects.)
  - $\overline{D}$ is the defects per unit area of defects of all sizes.

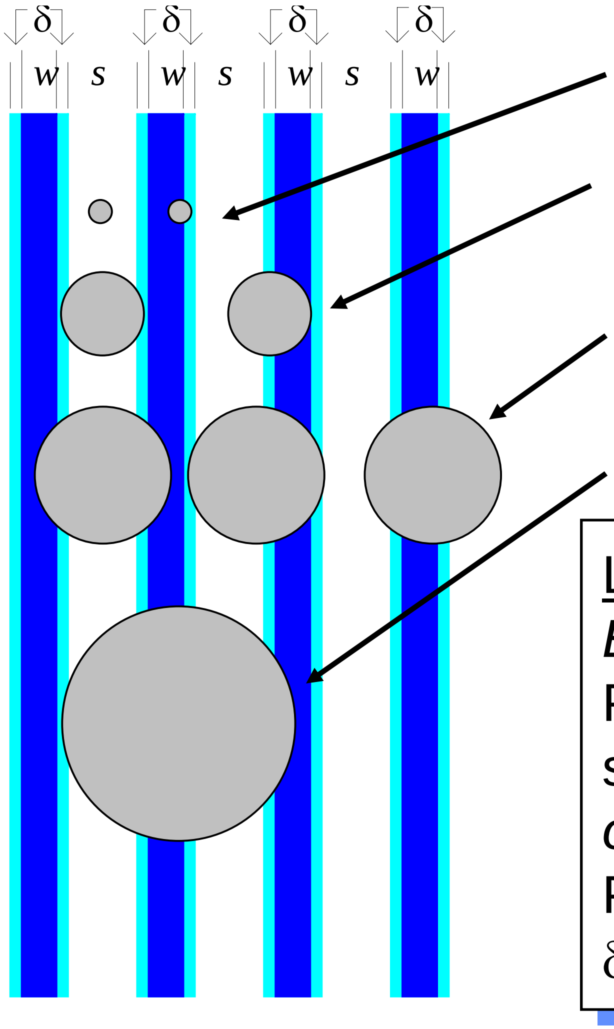
* C. H. Stapper, "Modeling of Integrated Circuit Defect Sensitivities", IBM J. Res. Develop. Vol. 27, pp 549-557 (1983)

Probability of “Yield” and “Reliability” Defects.

- “Yield” defects prevent operation of the device at before any stress ($t = 0$).
- Latent “reliability” defects will eventually kill the device (that is, at $t > 0$).
- In simple cases, the probability of occurrence of a given defect type can be *calculated* as a function of defect size, assuming random spatial distribution of defects.*
- We’ll calculate the probability of “Yield” defects and “Reliability plus Yield” defects falling on a metal comb.

* See, for example, C. H. Stapper, “Modeling of defects in integrated circuit photolithographic patterns.” IBM J. Res. Develop. Vol. 28, pp 461-475 (1984)

Probability of “Yield” and “Reliability” Defects



OK, Never a yield or reliability issue.

Sometimes a latent reliability defect.

Sometimes a yield defect, sometimes a latent reliability defect, sometimes OK.

Always a yield defect.

Latent Reliability Defect:
Either:
Particle does not touch conductors, but both sides are within δ of the conductor.
or:
Particle touches one conductor and is within δ of its neighbor.

Probability of “Yield” and “Reliability” Defects

- Calculation of proportion of yield and reliability defects assuming random distribution of defects of diameter x.
- $P_{\text{yield}}(x)$ is the proportion of “yield” defects.

$$\begin{aligned} P_{\text{yield}}(x) &= 0, && \text{for } x < s \\ &= \frac{x-s}{s+w}, && \text{for } s \le x < 2s+w \\ &= 1, && \text{for } x \ge 2s+w \end{aligned}$$

- $P_{\text{yield \& latent rel}}(x)$ is the proportion of “reliability” and “yield” defects. ($s => s - 2\delta$ and $w => w + 2\delta$)

$$\begin{aligned} P_{\text{yield \& latent rel.}}(x) &= 0, && \text{for } x < s-2\delta \\ &= \frac{x-s+2\delta}{s+w}, && \text{for } s-2\delta \le x < 2s+w-2\delta \\ &= 1, && \text{for } x \ge 2s+w-2\delta \end{aligned}$$

Yield and Reliability Defect Densities

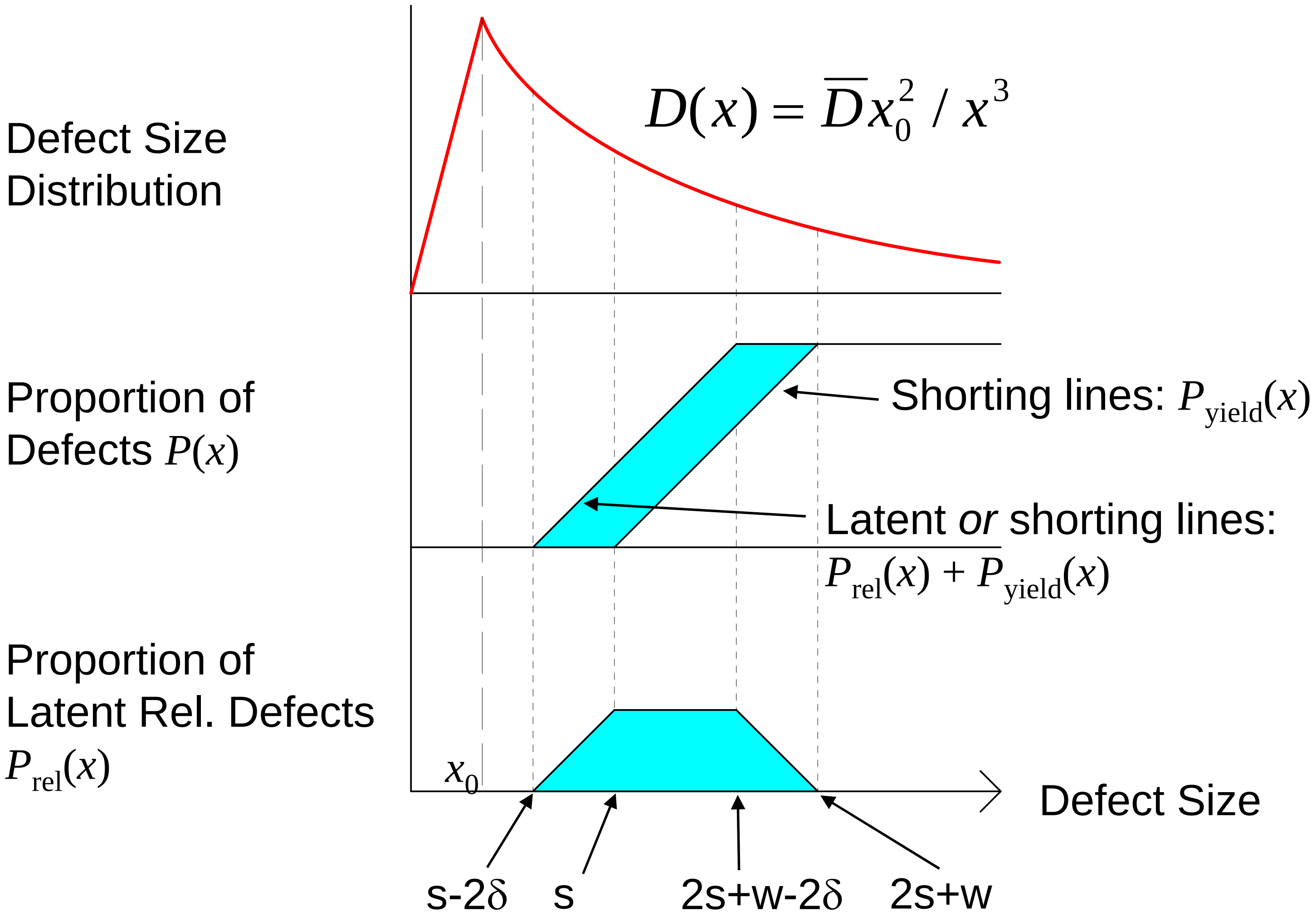
- Combine
  - – Defect size distribution.
  - – Probability of type of defect vs defect size.
- Calculate the defect density of defects
  - – which are fatal to device at $t = 0$:

$$D_{\text{yield}} = \int_0^\infty D(x) P_{\text{yield}}(x) dx = \frac{\overline{D} x_0}{2s(w+2s)}$$

  - – and those which are latent reliability defects:

$$D_{\text{rel}} = \int_0^\infty D(x) P_{rel}(x) dx = \frac{\overline{D} x_0}{2} \left[\frac{1}{(s-2\delta)(w+2s-2\delta)} - \frac{1}{s(w+2s)} \right]$$

Yield and Reliability Defect Densities



$$D(x) = \overline{D} x_0^2 / x^3$$

Relationship Between Yield and Reliability Defect Densities

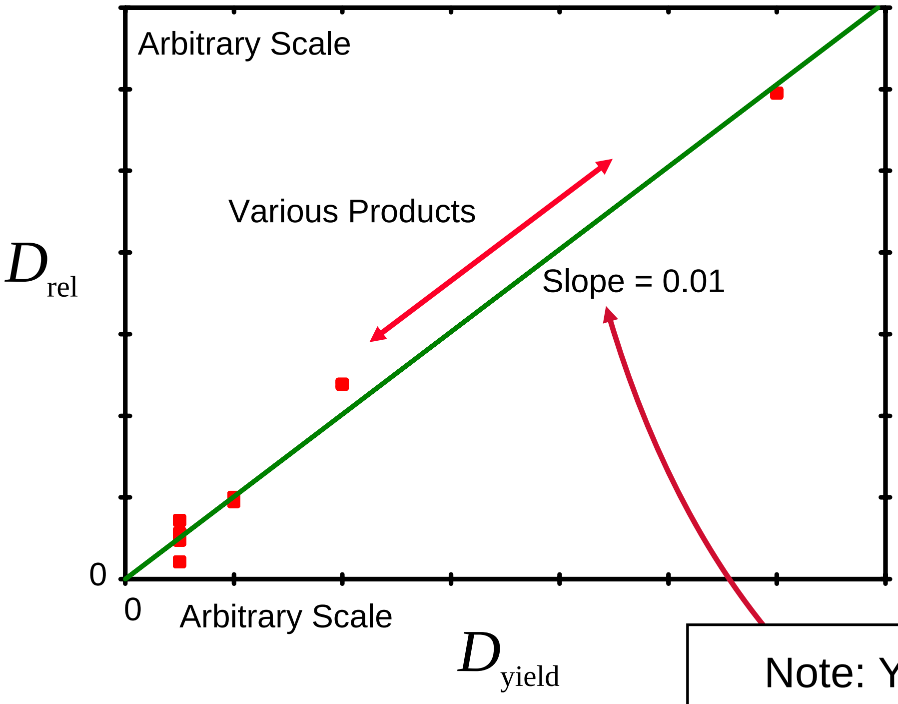
- Reliability and yield defect densities are proportional.

$$\frac{D_{\text{rel}}}{D_{\text{yield}}} = \delta \times \frac{2(w+3s)}{s(w+2s)} + \text{ higher order terms in } \delta$$

- The ratio of latent reliability defect density to yield defect density depends on
  - The *shape* of the defect size distribution.
  - The *pattern* on which the defects fall (layout sensitivity)
  - The definition of “latency” (the value of δ).
  - An assumption of non-interacting, randomly distributed defects.

Relationship Between Yield and Reliability Defect Densities

- The ratio D_{rel}/D_{yield} can be measured...



Reliability defect density as measured by 6 hour burn-in fallout, versus yield for various products.

$$F(t) = \text{ Proportion failing at time } t$$

$$D_{rel} = -\ln\{F(t = 6 \text{ hours})\} / \text{Area}$$

$$D_{yield} = -\ln\{\text{Yield}\} / \text{Area}$$

Note: Yield defect density is 100X larger than reliability defect density. So it's easier to measure.

Simulation of Defect Reliability

- In general, analytical calculation of reliability and yield defectivities is complex because of
  - Complex defect size distributions.
  - Non-circular defects with orientation distributions.
  - Complex substrate patterns.
- Often it is easier to use Monte Carlo methods to evaluate defectivities by *simulation*.
- We'll discuss simulation more when we look at an assembly-related example a bit later.

Relationship Between Yield and Reliability Defect Densities

- Reliability and yield defect densities can be modeled, simulated, or measured.
- But for reliability prediction we don't care what the value of D_{rel}/D_{yield} is.
  - We only care that they proportional.
- The model we derive requires that κ_i be a constant for each mechanism and substrate pattern, i:

$$\kappa_i = \frac{D_{rel}(i)}{D_{yield}(i)}$$

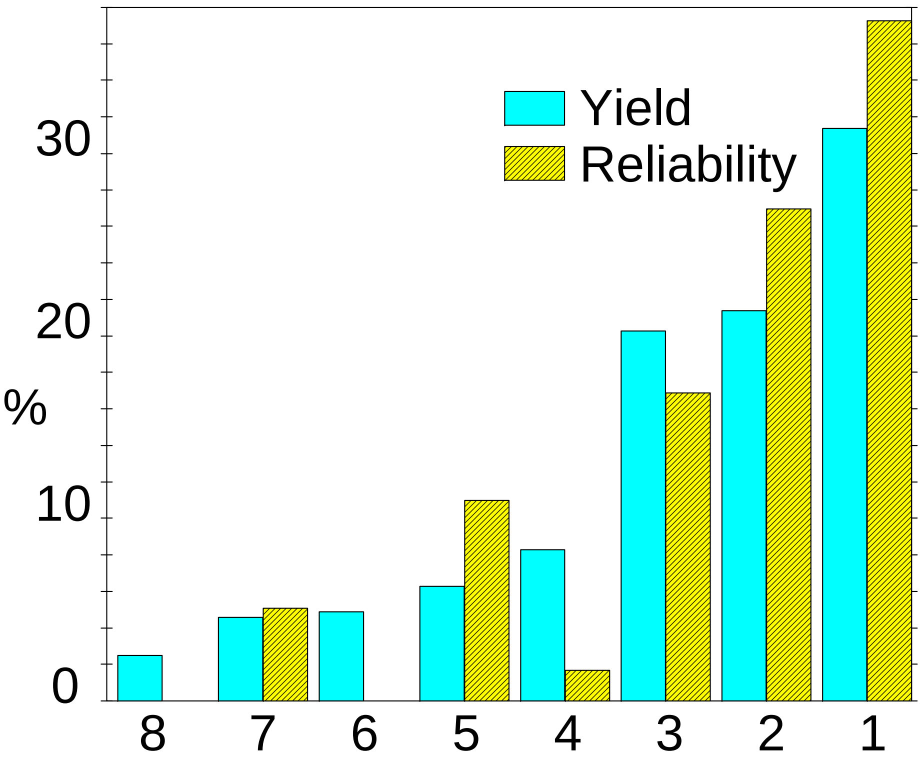
- *This is not a law of nature* - it depends on a constant defect size distribution *shape*, ie. a process under statistical control.

Relationship Between Yield and Reliability Defect Densities

- Yield and reliability are proportional for all defects, especially the most frequently occurring ones.

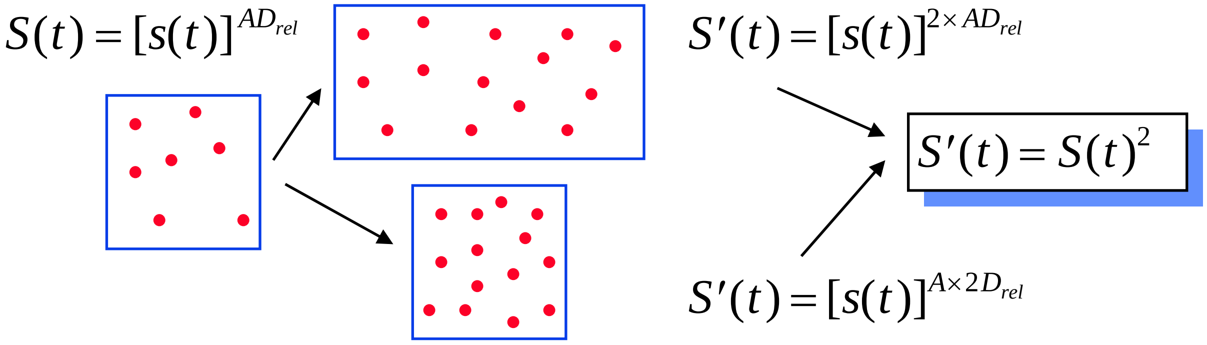
<u>Yield and Reliability Defect Pareto</u>

1. Metal Defects and Particles
2. W Defects and Particles
3. NVD
4. Plug Defects
5. Other Defects
6. Spacer Defects
7. Poly Defects and Particles
8. Diffusion Defects



Scaling of Defect Reliability

- Assume
  - Each defect has a survival function $s(t)$, and the density is D_{rel} (defects/cm^2).
  - Random, non-interacting defects.
  - $S(t)$ is the survival probability of a die of area *A*.
- Consider 2 cases
  - Double the area, keep the defect density the same.
  - Double the defect density, keep the area the same.



$$S(t) = [s(t)]^{AD_{rel}}$$

$$S'(t) = [s(t)]^{2\times AD_{rel}}$$

$$S'(t) = [s(t)]^{A\times 2D_{rel}}$$

$$\boxed{S'(t) = S(t)^2}$$

Scaling of Defect Reliability

- For one mechanism *i*.
- For one circuit layout pattern.
- At one condition of temperature and bias.

Case 1 (“Reference”)

$S_i(t)$ = known
A = known area
$D_{rel}(i)$ = “unknown”

→ **?**

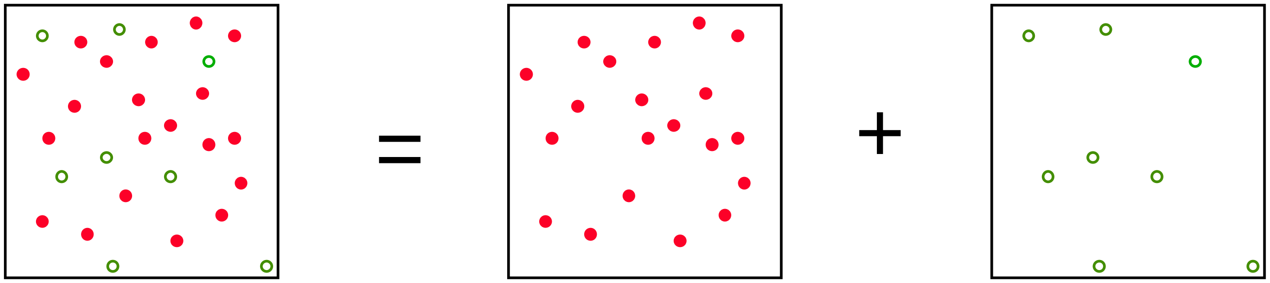
Case 2 (“Product”)

$S'_i(t)$ = UNknown
A’ = known area
$D'_{rel}(i)$ = “unknown”

$$S_i'(t) = S_i(t)^{\frac{D'_{rel}(i)A'}{D_{rel}(i)A}}$$

Concept of Reliability Defect Density

- Concept in this tutorial:

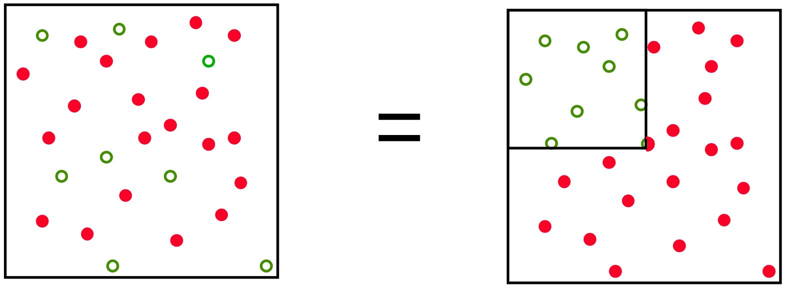


Total Defect Density

Yield Defect Density

Reliability Defect Density

- Concept used in other* work:



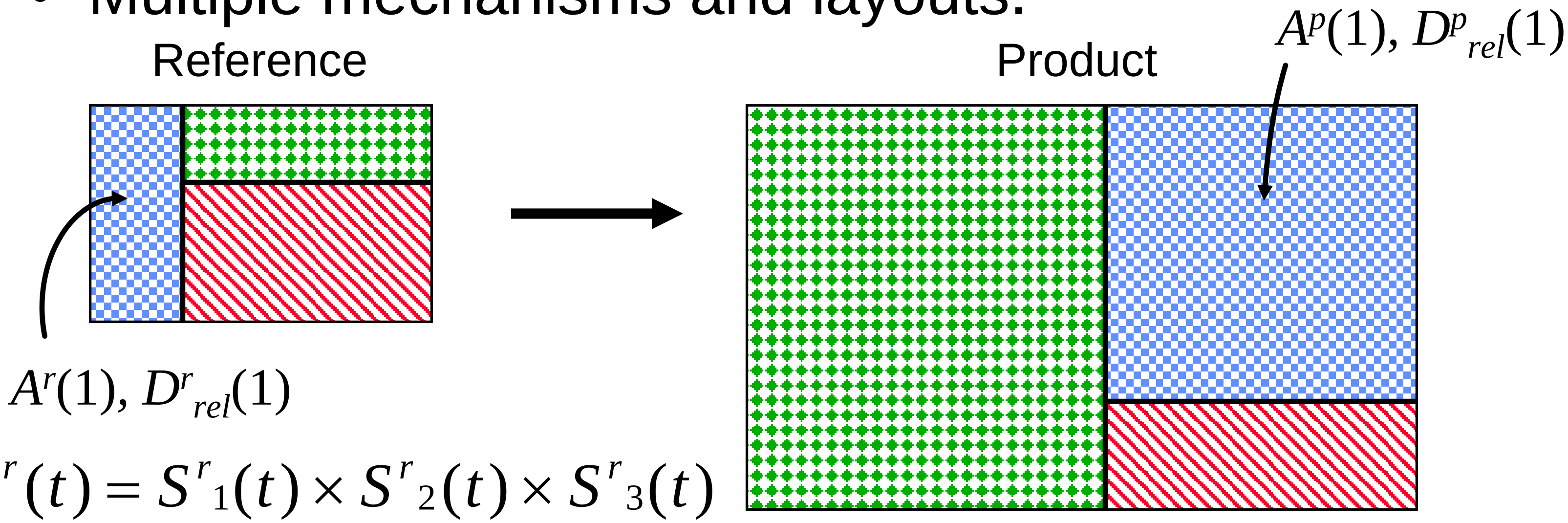
Constant defect density, “critical areas” for yield and reliability.

Problem: It’s easy to confuse physical die area (and subarea) scaling with the abstract concept of “critical area”.

* H.H. Huston, and C.P. Clarke, “Reliability Defect Detection and Screening during Processing - Theory and Implementation”, IRPS 1992, pp 268-274.

Scaling of Defect Reliability

- Consider a process reference monitor “*r*” (eg. an SRAM), and a product “*p*”.
- Multiple mechanisms and layouts.



$$S^r(t) = S^r{}_1(t) \times S^r{}_2(t) \times S^r{}_3(t)$$

$$S^p(t) = S^r{}_1(t)^{\frac{D^p{}_{rel}(1)\times A^p(1)}{D^r{}_{rel}(1)\times A^r(1)}} \times S^r{}_2(t)^{\frac{D^p{}_{rel}(2)\times A^p(2)}{D^r{}_{rel}(2)\times A^r{}_{rel}(2)}} \times S^r{}_3(t)^{\frac{D^p{}_{rel}(3)\times A^p(3}{D^r{}_{rel}(3)\times A^r(3}}$$

Scaling of Defect Reliability

The critical relationship...

$$\frac{D^p_{\text{rel}}(i)\times A^p(i)}{D^r_{\text{rel}}(i)\times A^r(i)} = \frac{\kappa_i(\text{product})\times D^p_{\text{yield}}(i)\times A^p(i)}{\kappa_i(\text{reference})\times D^r_{\text{yield}}(i)\times A^r(i)} \cong \frac{D^p_{\text{yield}}(i)\times A^p(i)}{D^r_{\text{yield}}(i)\times A^r(i)}$$

If this ratio is unity, *then* this is true.

The ratio is unity when the *shape* of the defect size distribution is a constant. This will be true for a process which is in statistical control.

Scaling of Defect Reliability

- Reliability defect densities, D_{rel}, are not well known and are small, but D_{yield} are related to production indicators and are larger.
- Appeal to constancy of $D_{\mathrm{rel}}/D_{\mathrm{yield}}$ for each mechanism/subdie to write

$$S^{p}(t) \cong S_{1}(t)^{\frac{D^{p}{}_{yield}(1)\times A^{p}(1)}{D^{r}{}_{yield}(1)\times A^{r}(1)}} \times S_{2}(t)^{\frac{D^{p}{}_{yield}(2)\times A^{p}(2)}{D^{r}{}_{yield}(2)\times A^{p}{}_{rel}(2)}} \times S_{3}(t)^{\frac{D^{p}{}_{yield}(3)\times A^{p}(3)}{D^{r}{}_{yield}(3)\times A^{r}(3)}}$$

- In general:

$$S^{p}(t) = \prod_{i} [S_{i}^{r}(t)]^{R_{i}(p|r)}$$

$$R_{i}(p|r) = \frac{D^{p}_{\mathrm{yield}}(i) \times A^{p}(i)}{D^{r}_{\mathrm{yield}}(i) \times A^{r}(i)}$$

This is not yet in a form corresponding to the usual yield statistics acquired by the factory...

Yield Statistics

- Assuming Poisson statistics, the yield for the compound die is given by

$$Y^p = Y^p_{\text{intrinsic}} \times \exp[-D_{\text{yield}}{}^p(1)A^p(1)] \times \exp[-D_{\text{yield}}{}^p(2)A^p(2)]$$
$$\times \exp[-D_{\text{yield}}{}^p(3)A^p(3)]$$

$$Y^p = Y^p_{\text{intrinsic}} \times \exp\left(\sum_j -D_{\text{yield}}{}^p(j)A^p(j)\right) = Y^p_{\text{intrinsic}} \times \exp(-D_{\text{yield}}{}^p \times A^p)$$

$$\boxed{D_{\text{yield}}{}^p \times A^p = -\ln\left(\frac{Y^p}{Y^p{}_{\text{intrinsic}}}\right)}$$

Subdie area-weighted defect density.

Total die area.

$$D^p \equiv \frac{\sum_i D_{\text{yield}}{}^p(j)A^p(j)}{A^p} \qquad A^p \equiv \sum_j A^p(j)$$

Scaling of Defect Reliability

- Some manipulation shows that

$$R_i(p|r) = \frac{D^p_{\text{yield}}(i) \times A^p(i)}{D^r_{\text{yield}}(i) \times A^r(i)} = \frac{P^p(i) \times D^p_{\text{yield}} \times A^p}{P^r(i) \times D^r_{\text{yield}} \times A^r}$$

where the Pareto (proportion of all defects attributable to mechanism i) is defined by

$$P^p(i) \equiv \frac{D^p_{\text{yield}}(i) \times A^p(i)}{\sum_j D^p_{\text{yield}}(j) \times A^p(j)}$$

- So if $Y^p_{intrinsic} = 1$ (as usual), then

$$R_i(p|r) = \frac{P^p(i) \times \ln(Y^p)}{P^r(i) \times \ln(Y^r)}$$

Scaling of Defect Reliability: Practical Formulae

- So, in terms of the usual Pareto and yield indicators acquired as factory yield indicators at sort test:

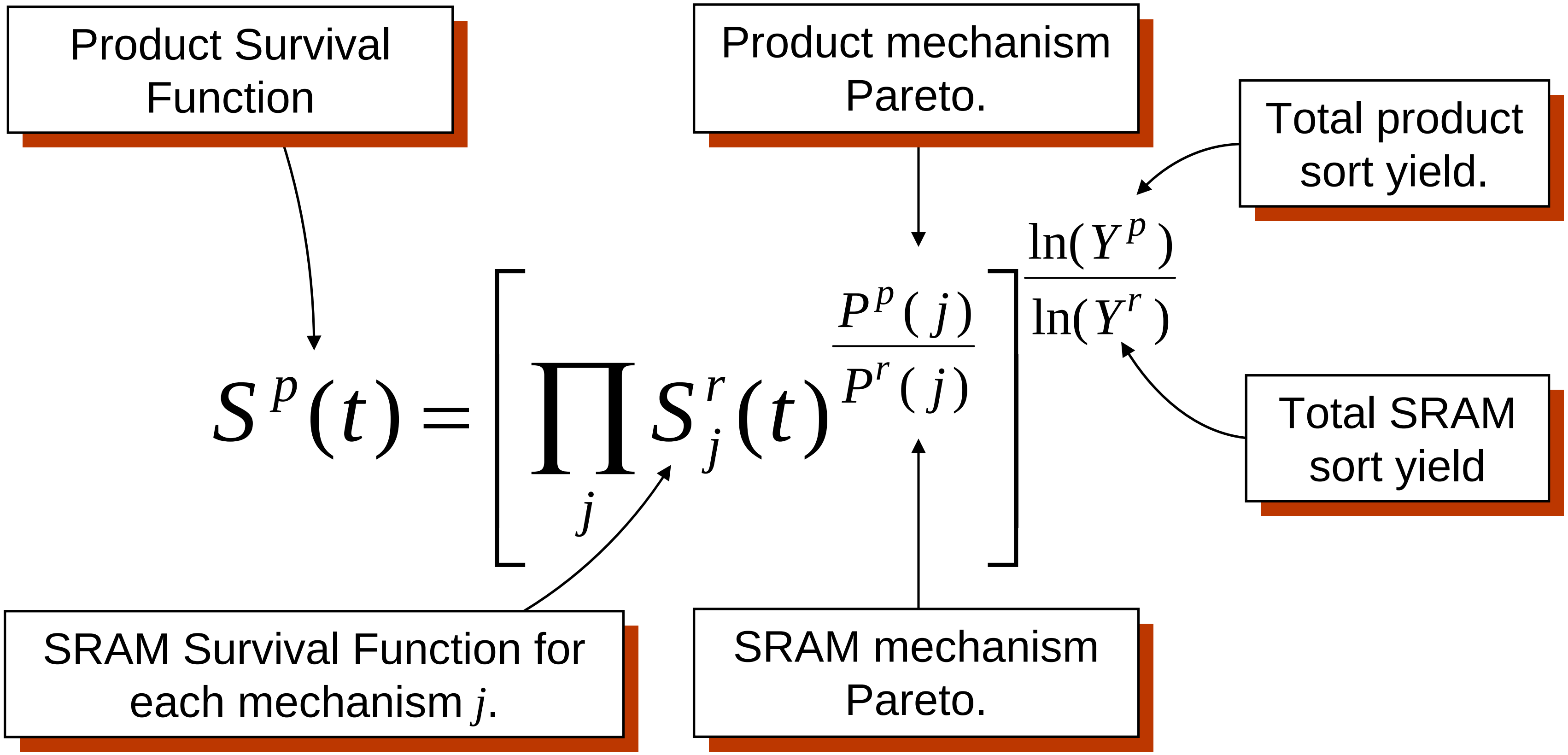
$$S^p(t) = \left[S^r{}_1(t)^{\frac{P^p(1)}{P^r(1)}} \times S^r{}_2(t)^{\frac{P^p(2)}{P^r(2)}} \times S^r{}_3(t)^{\frac{P^p(3)}{P^r(3)}} \right]^{\frac{\ln(Y^p)}{\ln(Y^r)}}$$

or, in general

$$S^p(t) = \left[\prod_j S^r_j(t)^{\frac{P^p(j)}{P^r(j)}} \right]^{\frac{\ln(Y^p)}{\ln(Y^r)}}$$

where we have assumed $Y^p{}_{intrinsic} = 1$ as is usual.

Scaling of Defect Reliability: Practical Formulae



$$S^{p}(t) = \left[\prod_{j} S_{j}^{r}(t)^{\frac{P^{p}(j)}{P^{r}(j)}} \right]^{\frac{\ln(Y^{p})}{\ln(Y^{r})}}$$

Product Survival Function

Product mechanism Pareto.

Total product sort yield.

Total SRAM sort yield

SRAM Survival Function for each mechanism *j*.

SRAM mechanism Pareto.

Eg. Reference Product: SRAM. Product of Interest: Microprocessor

Scaling of Defect Reliability: Practical Formulae

- If the defect paretos are the same for reference and “unknown” product, then

$$P^{p}(i) = P^{r}(i), \quad \text{for each } i$$

Usually a good approximation since only one or two mechanisms dominate.

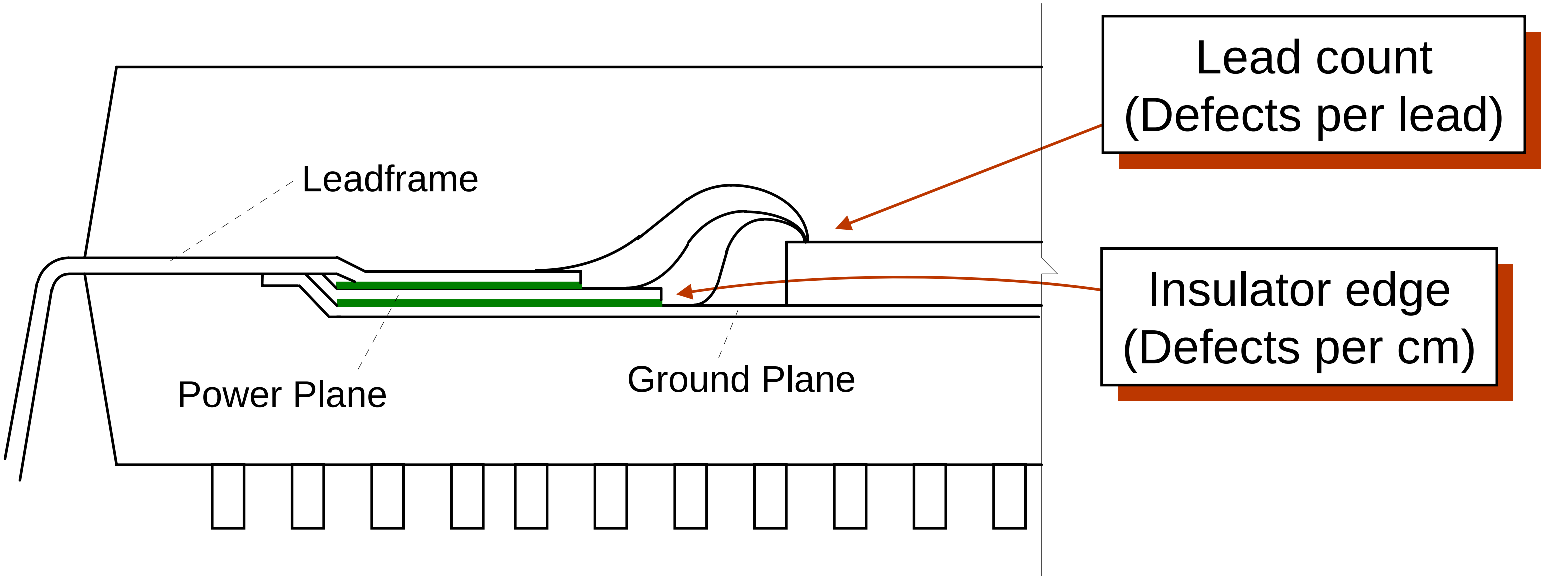
so

$$\boxed{S^{p}(t) = S^{r}(t)^{\frac{\ln(Y^{p})}{\ln(Y^{r})}}}$$

where the total reference (usually SRAM) survival function is $S^{r}(t) = \prod_{j} S_{j}^{r}(t)$

Extensions to Non-Die Area-Related Mechanisms

- Defect count per device may scale with other extensive properties of the product.
  - Die Area => Lead count, perimeter of dielectric edge in package, etc.
  - Areal defect density => defects per lead, defects per length of perimeter in package, etc.



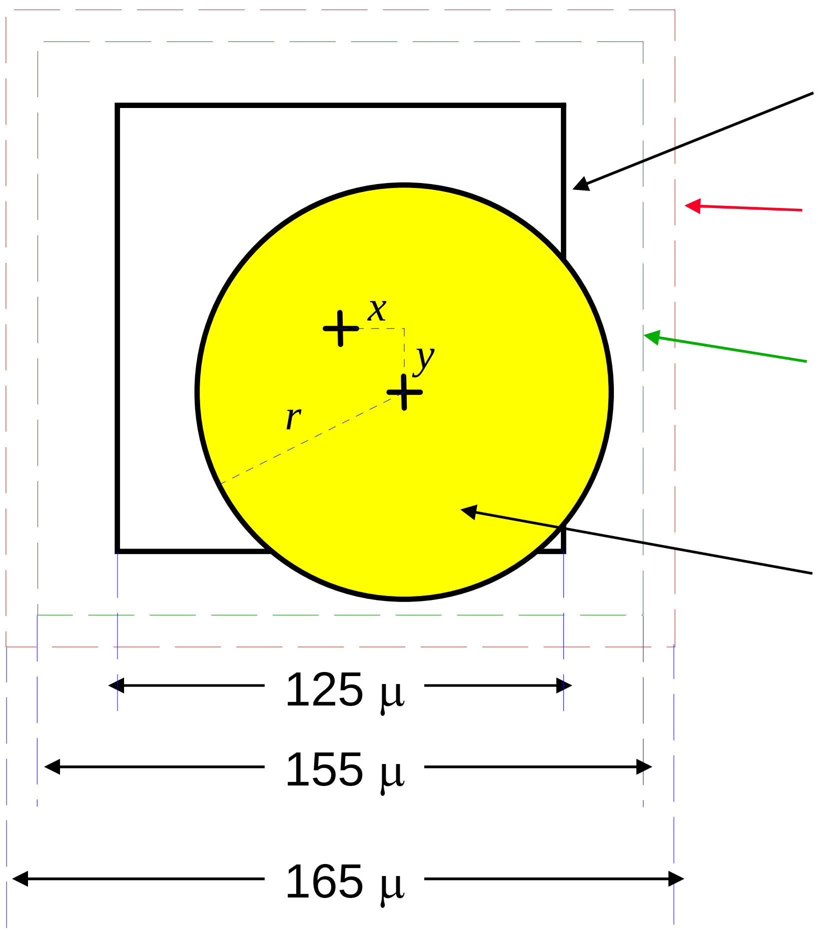
Lead count
(Defects per lead)

Insulator edge
(Defects per cm)

Yield/Reliability Simulation for Wire Bonding

- Measure physical process capability.
  - Make measurements of bond location and ball size.
  - Use a sample of about 200.
  - Determine distribution of bond center (x,y), and ball diameter, r.
    - Shape (normal, etc.), Mean, Variance.
    - Determine whether x, y, r are correlated.
- Decide on yield and reliability specification limits.
- Calculate yield and latent reliability DPM.
  - Assume that process is in statistical control.
  - Analytical calculation - difficult, not general.
  - Simulate the process using fitted distribution parameters.

Yield/Reliability Simulation for Wire Bonding



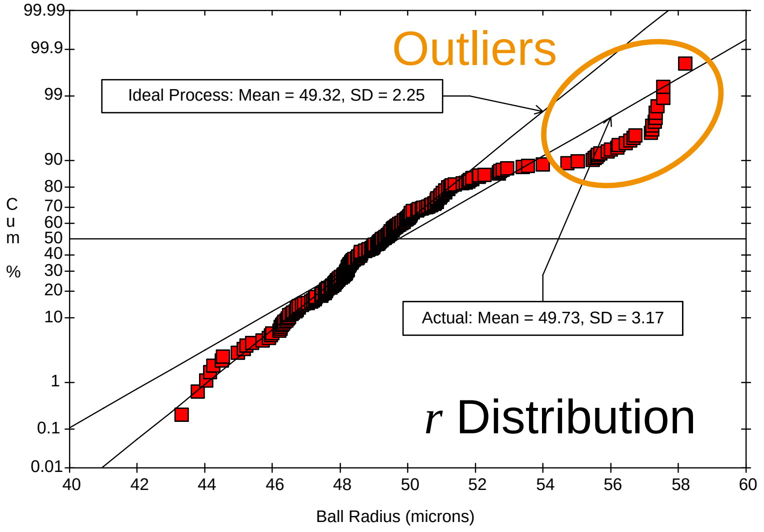
Bond pad opening.

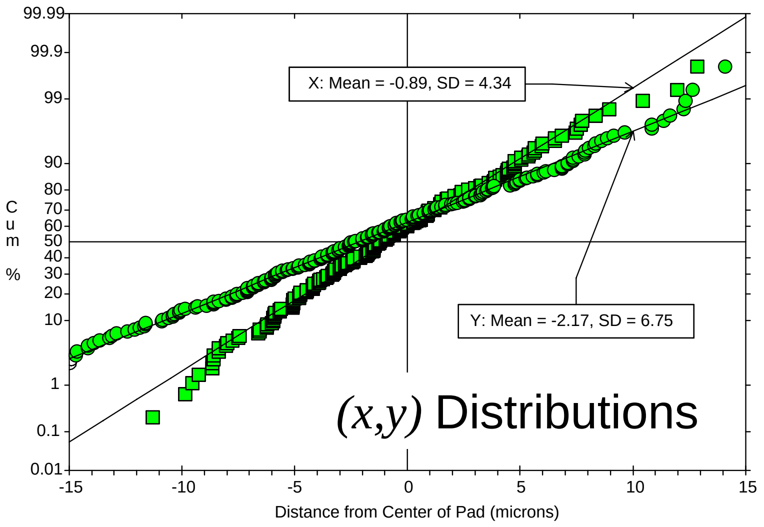
Ball can overlap adjacent metal (reliability jeopardy).

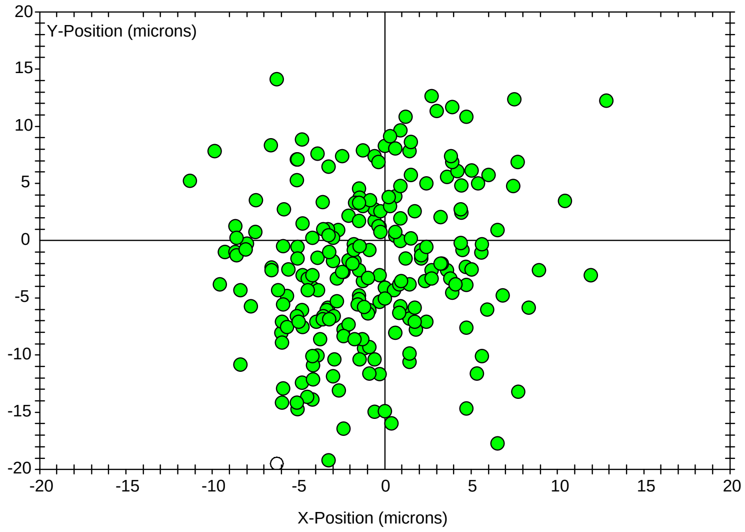
Ball can overlap adjacent pad opening - dead short (yield issue)

Ball bond.

Yield/Reliability Simulation for Wire Bonding





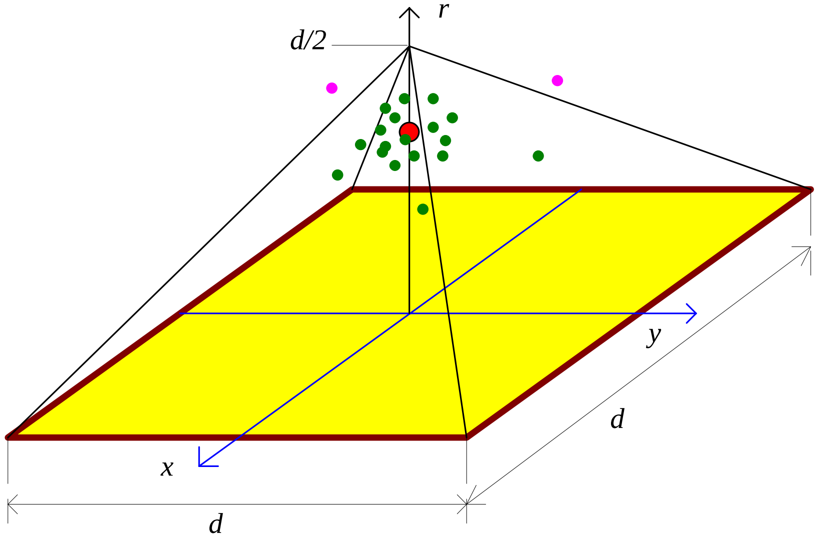


- x, y, and r distributions are normal and uncorrelated.
- The parameters of the process in "statistical control" are

x Mean	x SD	y Mean	y SD	r Mean	r SD
0	4.34	0	6.75	49.3	2.25

Yield/Reliability Simulation for Wire Bonding

- Individual bonds are points clustering around the target.
- Bonds inside pyramid pass the criterion.
- Bonds outside the pyramid fail the criterion.
- Integrate an elipsoidal probability function centered on the target over the volume intersected by the pyramid to get DPM. *Difficult* to do in general. OR..
- Use random number generator to simulate millions of bonds using distribution parameters determined from 200-unit experiment. This is *easy*!



Yield/Reliability Simulation for Wire Bonding

Number of Bonds to Simulate

```
procedure(n);
/* BONDSIM - n is the number of iterations */
    {
    bnd_tbl = "@bndplace1@bond_param";
    bnd_results = "@bndplace1@bond_results";
    xmean = bnd_tbl [1, 1];
    xsd = bnd_tbl [1, 2];
    ymean = bnd_tbl [1, 3];
    ysd = bnd_tbl [1, 4];
    dmean = bnd_tbl [1, 5];
    dsd = bnd_tbl [1, 6];
    do m = 1 to n;
        {
        xcen = xmean + normdev() * xsd;
        ycen = ymean + normdev() * ysd;
        rad = 0.5 * (dmean + normdev() * dsd);
        ball_right = xcen + rad;
        ball_left = xcen - rad;
        ball_top = ycen + rad;
        ball_bott = ycen - rad;
        do p = 1 to lastcol(bnd_results );
            {
            lright = ball_right > 0.5 * bnd_results [1, p];
            lleft = ball_left < - 0.5 * bnd_results [1, p];
            ltop = ball_top > 0.5 * bnd_results [1, p];
            lbott = ball_bott < - 0.5 * bnd_results [1, p];
            lfail = lright OR lleft OR ltop OR lbott;
            if lfail then
                bnd_results [2, p] = bnd_results [2, p] + 1;
            }
        }
    }
```

Distribution Parameters

X-Mean	X SD	Y-Mean	Y-SD	Dia.-Mean	Dia-SD
0	4.34	0	6.75	98.64	4.5

Simulate Normal Deviate (mean = 0, SD = 1).

"Numerical Recipes" by W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, Cambridge UP (1986), p203.

Results

	Total Bonds	Pad Opening	Overlap Metal	Dead Short
Criterion	0	125	155	165
Counts	1000000	70232	68	3

"Reliability"

"Yield"

Scaling of Defect Reliability: Summary

- Yield and reliability defect densities may be calculated, simulated, or measured, but...
- The model requires an assumption (or null hypothesis) of
  - Random, non interacting defects.
  - An invariant *ratio* of Yield to Reliability defect densities.
- The defect-related part of the survival function scales with the density of latent reliability defects and die (or affected subdie) area.
- By hypothesis, yield and reliability defect densities are proportional, so the defect part of the survival function ALSO scales with yield defect density.
- The "practical" form of the model involves sort yield and sort Pareto data. Simplification obtains if yield defect Paretos are invariant.

Accelerated Stressing And Burn-In

- What is an acceleration factor?
  - Start with the same population
  - Case 1: Temperature T_1, voltage V_1, time interval dt_1, *a certain proportion* fails.
  - Case 2: Temperature T_2, voltage V_2, it takes dt_2 for the *same proportion* of the population to fail.
  - The acceleration of case 2 relative to case 1, for mechanism i is

$$\frac{dt_1}{dt_2} = AF_i(2|1) \quad \text{(instantaneous)}$$

$$t_1 = AF_i(2|1)t_2 \quad \text{(constant acceleration)}$$

Accelerated Stressing and Burn-In

- The survival function for mechanism *i* at environmental condition 2 is related to the survival function at environmental condition 1 by:

$$S_i(2|t) = S_i\{1|AF_i(2|1)t\}$$

- We'll use an acceleration factor function given by:

$$AF_i(2|1) = \exp\left\{ \frac{Q_i}{k}\left[\frac{1}{T_1} - \frac{1}{T_2} \right] + C_i(V_2 - V_1) \right\}$$

Accelerated Stressing and Burn-In

- For all mechanisms, the survival probability of an unknown product p at T_2 and V_2 may be calculated from the mechanism survival probabilities of a “known” reference product r at T_1 and V_1:

$$S^p(2|t) = \prod_i [S_i^r\{1|AF_i(2|1)t\}]^{R_i(p|r)}$$

$$R_i(p|r) = \frac{D_{\text{yield}}^p(i) \times A^p(i)}{D_{\text{yield}}^r(i) \times A^r(i)} = \frac{P^p(i) \times \ln(Y^p)}{P^r(i) \times \ln(Y^r)}$$

Accelerated Stressing and Burn-In

- Consider a device undergoing t_B hours of burn-in at environmental condition “B”, followed by t hours of “use” at environmental condition “2”.

Probability of surviving t_B hours of burn-in at condition “B” AND t hours of use at condition “2”

= Probability of surviving t_B hours of burn-in at condition “B”

X **Probability of surviving t hours of use at condition “2” GIVEN THAT the device has survived t_B hours of burn-in at condition “B”.**

This is what the end-user sees.

or, symbolically

$$\tilde{S}^p = S^p(B|t) \times S'^p(2|t)$$

so

$$S'^p(2|t) = \frac{\tilde{S}^p}{S^p(B|t)}$$

Accelerated Stressing and Burn-In

- Effect of Burn-in
  - The proportion of the initial population which survives burn-in for time t_B at T_B and V_B is

$$S^p(B|t_B) = \prod_i [S_i^r\{1|AF_i(B|1)t_B\}]^{R_i(p|r)}$$

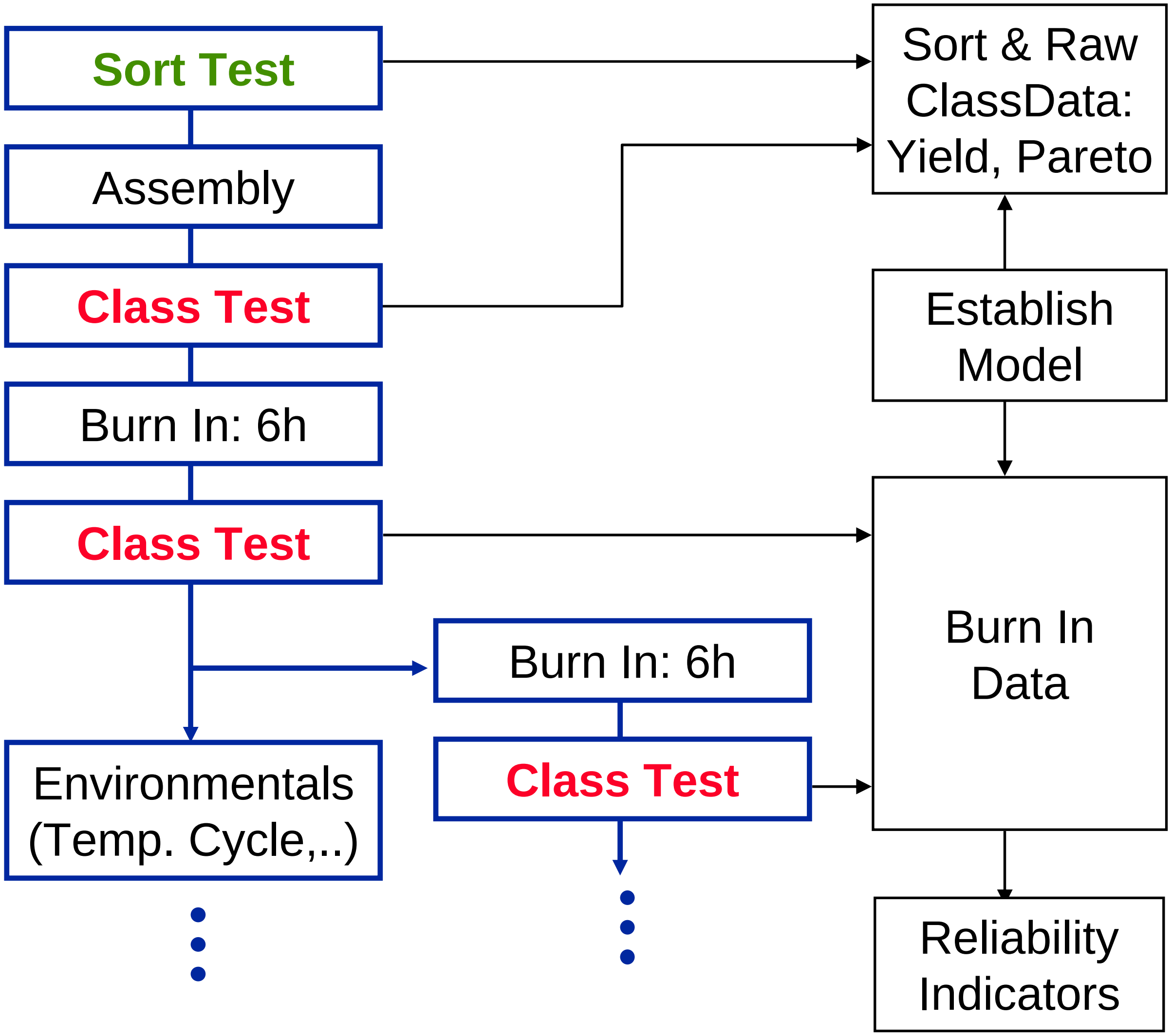
  - after additional time t at T_2 and V_2 the proportion surviving is

$$\tilde{S}^p = \prod_i [S_i^r\{1|AF_i(2|1)t + AF_i(B|1)t_B\}]^{R_i(p|r)}$$

  - so the probability of surviving t at T_2 and V_2, *given that a unit has survived burn* in is :

$$S'^p(2|t) = \frac{\tilde{S}^p}{S^p(B|t_B)} = \prod_i \left[\frac{S_i^r\{1|AF_i(2|1)t + AF_i(B|1)t_B\}}{S_i^r\{1|AF_i(B|1)t_B\}} \right]^{R_i(p|r)}$$

Technology Development Test Flows


Sort Test

Assembly

Class Test

Burn In: 6h

Class Test

Burn In: 6h

Class Test

Environmentals
(Temp. Cycle,..)

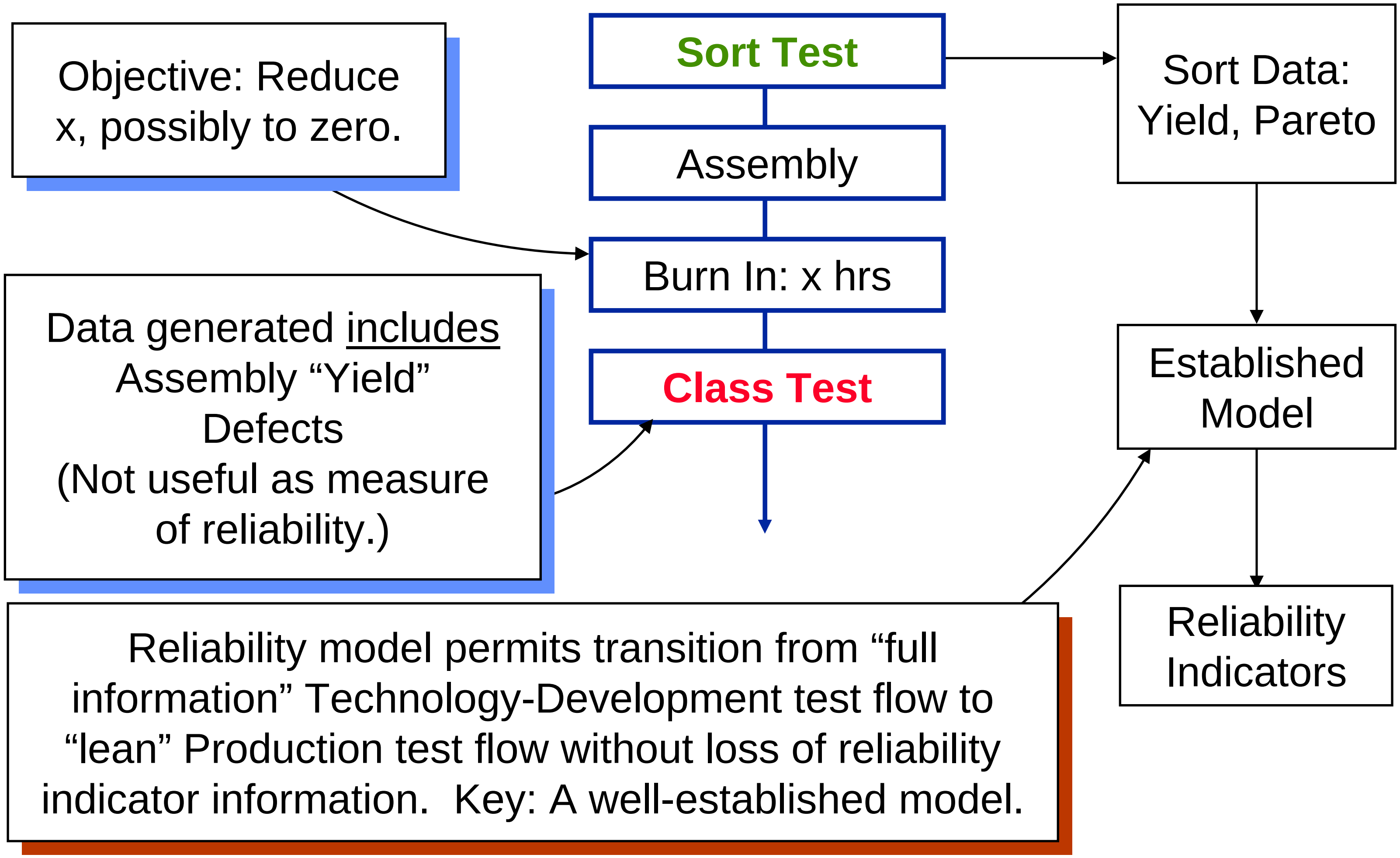
Sort & Raw
ClassData:
Yield, Pareto

Establish
Model

Burn In
Data

Reliability
Indicators

Production Test Flows



Objective: Reduce x, possibly to zero.

Sort Test

Assembly

Burn In: x hrs

Class Test

Sort Data: Yield, Pareto

Established Model

Reliability Indicators

Data generated includes
Assembly "Yield"
Defects
(Not useful as measure of reliability.)

Reliability model permits transition from "full information" Technology-Development test flow to "lean" Production test flow without loss of reliability indicator information. Key: A well-established model.

Test Programs

- Sort test is a wafer-level room temperature test.
- Class test is a unit level test using temperature-controlled hander.
- Sort and Class tests *<u>can stress</u>* units, particularly the high-voltage test.
  - Nominal temperature/volts is done last.
  - The stress in the test must be taken account of in low voltage burn-in (for acceleration studies).

Typical Sort Test

<u>Temperatures:</u>
Room
<u>Voltages:</u>
Low
High

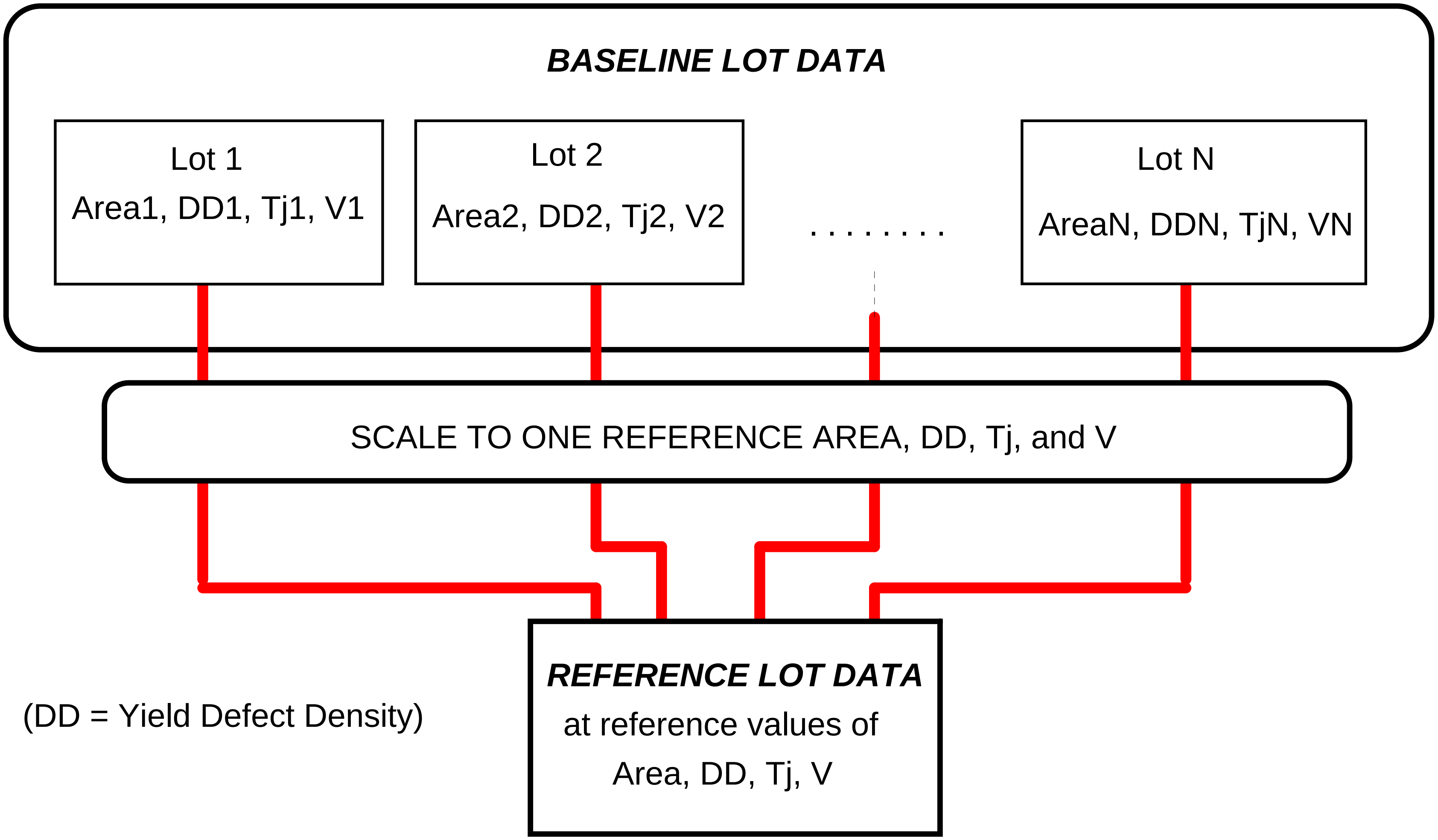
Typical Class Test

<u>Temperatures:</u>
Hot: 90 C
Cold: -10 C
Room
<u>Voltages:</u>
Low
High
Nominal

Analysis of Reliability Data

- How do we get the “reference” survival function?
- Production burn-in data and extended life test data from a variety of products fabricated using a specific process are accumulated. This body of data is the “baseline lot” reliability data.
- Baseline data is consolidated using “known” acceleration models and defect scaling to produce a “reference lot” reliability data.
- Reference lot data is calculated at single reference values of defect density, die area, temperature and bias.
- Parametric fits to reference lot data gives “reference model distributions”.

Analysis of Reliability Data



BASELINE LOT DATA

Lot 1
Area1, DD1, Tj1, V1

Lot 2
Area2, DD2, Tj2, V2

. . . .

Lot N
AreaN, DDN, TjN, VN

SCALE TO ONE REFERENCE AREA, DD, Tj, and V

REFERENCE LOT DATA
at reference values of
Area, DD, Tj, V

(DD = Yield Defect Density)

Typical Minimum Data Requirements for Determination of Process Reference Models

- 4 lots of SRAM, 4000 units at V = 140% of nominal, and 125C.
- Several lots at other bias/voltage conditions to determine acceleration parameters.
  - nominal bias, room temperature
  - sometimes assign Q, C based on "known" mechanism.
- All lots Class tested before burn-in ("clean burn-in")
- Readouts at 6, 48, 168, 500, 1000, 2000 hours.
- Known Yield and Defect Pareto for each lot.
- All failures validated, all failure signatures traceable to a physically analyzed failure.
  - "A Q and a C for every failure".

Analysis of Reliability Data

- Example of one lot of *Baseline Lot Data*.

Hours	6	12	24	48	168	500	1k	2k
Pass Defect (PD)	-	-	1	0	0	0	0	1
Fab Defect (FD)	-	-	3	0	0	2	0	0
Bake Recov. (BR)	-	-	0	0	0	0	0	0
Junct. Spike (JS)	-	-	0	0	0	0	0	0
Sample Size (SS)	-	-	2748	2744	2743	2293	2290	2290

Mechanism	Q_i(eV)	C_i(1/volts)
PD	0.3	1.8
FD	0.5	2.0
BR	1.0	0.0
JS	1.0	0.6

SRAM at V = 5.5 volt and T_j = 131C, die area = 36160 mils^2, D_{yield} = 1 (arbitrary units)

Note: Model predictions and data in this tutorial are examples only and are not representative of Intel products.

Analysis of Reliability Data

- Example of *Reference Lot Data* combined from from multiple lots of various products fabricated using the process.

Hours	6	12	24	48	168	500	1k	2k
PD	0	0	1.6	0	0	0	3.2	6.2
SS /PD	22642	1609	38305	51551	45212	5480	11808	5297
FD	105.7	0	18.6	54.0	53.9	19.1	24.8	20.4
SS /FD	21056	1407	34973	48604	42288	4304	10409	4207
BR	0	0	7.3	4.6	0	0	7.7	0
SS /BR	18281	1059	29629	47932	39302	3798	9383	3632
JS	0	0	2.9	0	27.7	7.5	0	9.6
SS/JS	18281	1059	29155	45472	37964	3015	8616	2958

Scaled to a Reference Condition of
V = 7 volts, T_j =160 C, Area = 268,686 mils^2,
D_{yield} = 0.21 (arbitrary units).

Analysis of Reliability Data

REFERENCE LOT DATA
at reference values of
Area, Defect Density, Tj, V

↓

Hazard Analysis using Kaplan-Meier-Greenwood Algorithm

↓

Fit Lognormal Distributions to Best Estimate and x% UCL KMG data points
x = 60%, 90%, 95%, 99%

↓

REFERENCE MODEL DISTRIBUTIONS
Model Parameters for Each Mechanism
at reference values of
Area, Defect Density, Tj, V

Statistical Interlude: Hazard Analysis

- Data produced by burn-in and life-test flows is nearly always censored (has removals).
  - Because material is diverted into other stresses in TD.
  - Because failures are often invalidated.
  - Because of multiple failure mechanisms.
- A simple method of analysis. For each mechanism:
  - Calculate instantaneous hazard.
  - Find cumulative hazard.
  - Use $F = 1\text{-}\exp(-H)$ to find cumulative failures.
  - Plot F vs time on log probability plot.

Statistical Interlude: Hazard Analysis

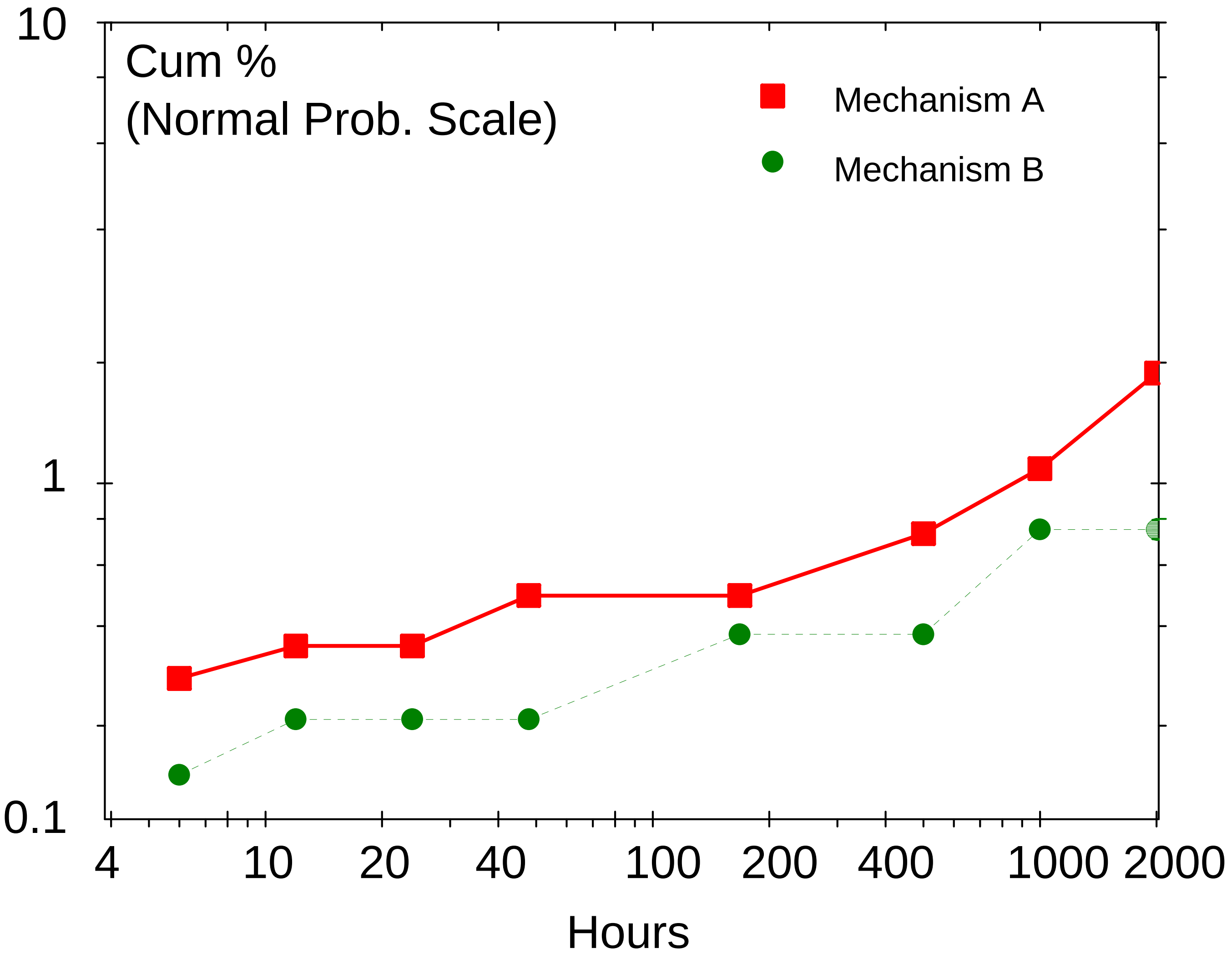
Two mechanisms with removals.

Removals: 0 0 1 839 150 149 147

Hours	***6***	***12***	***24***	***48***	***168***	***500***	***1000***	***2000***
SS	1423	1417	1415	1414	573	422	272	123
N(A)	4	1	0	2	0	1	1	1
N(B)	2	1	0	0	1	0	1	0
$h_i(A)=N(A)/SS$	0.0028	0.0007	0.0000	0.0014	0.0000	0.0024	0.0037	0.0081
$H_i(A)=\sum h_i(A)$	0.0028	0.0035	0.0035	0.0049	0.0049	0.0073	0.0110	0.0191
A: $F_i=1-exp(-H_i)$	***0.0028***	***0.0035***	***0.0035***	***0.0049***	***0.0049***	***0.0073***	***0.0109***	***0.0189***
$h_i(B)=N(B)/SS$	0.0014	0.0007	0.0000	0.0000	0.0017	0.0000	0.0037	0.0000
$H_i(B)=\sum hi(B)$	0.0014	0.0021	0.0021	0.0021	0.0038	0.0038	0.0075	0.0075
B: $F_i=1-exp(-H_i)$	***0.0014***	***0.0021***	***0.0021***	***0.0021***	***0.0038***	***0.0038***	***0.0075***	***0.0075***

Plot cumulative failures (bold italics) on probability plot...

Statistical Interlude: Hazard Analysis



Analysis of Reliability Data

- The Kaplan-Meier-Greenwood (KMG) method handles censored readout data *and provides confidence intervals.* See Nelson*.
- Plot, lognormally, KMG estimates of cum fails.
- Least-squares fit of straight line through KMG plot points provides statistical model parameters.

Inverse Normal Probability Function

$$y_i = \Phi^{-1}(F_i); \quad x_i = \ln(t_i)$$

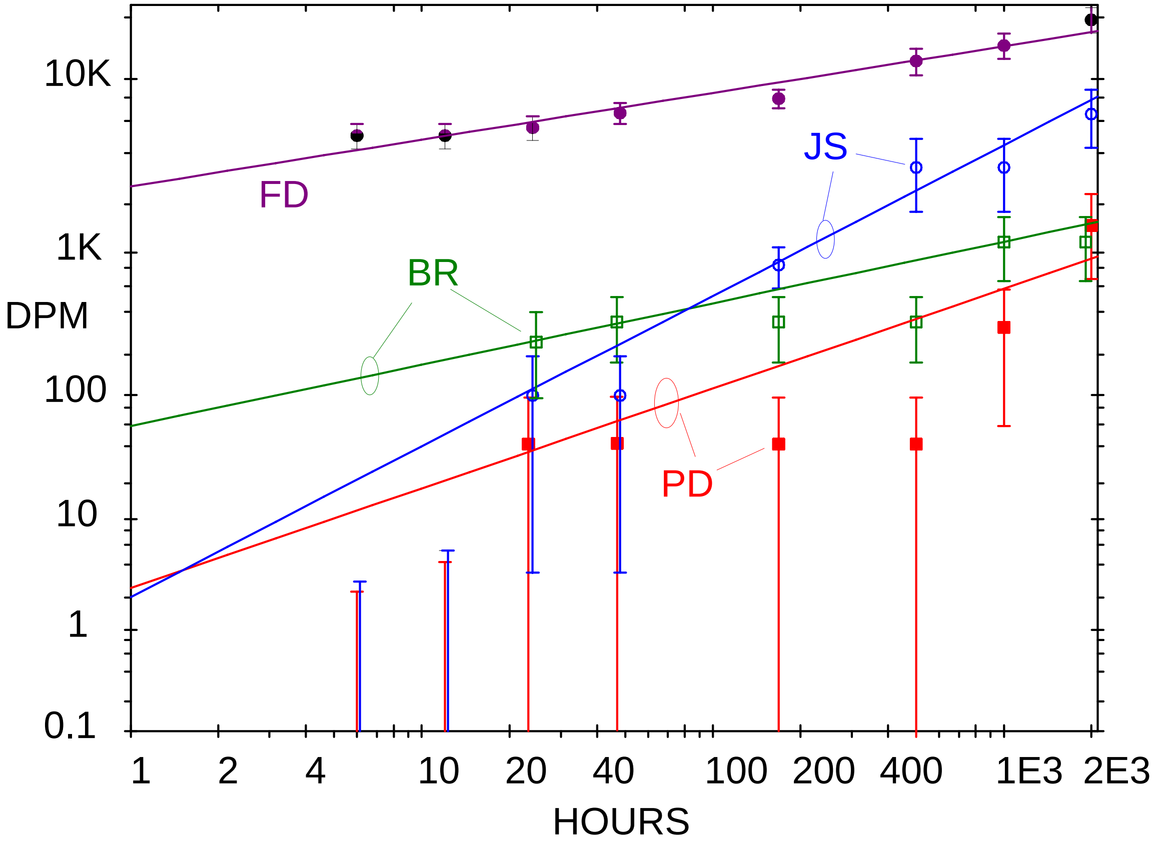
$$\sigma = 1/\text{slope}; \quad \mu = -\sigma \times \text{intercept}$$

$$t_{50} = \exp(\mu)$$

* W. Nelson, "Accelerated Testing," John Wiley & Sons (1989), pp 145-151.

Analysis of Reliability Data

Reference Lot Data at *V* = 7 V, T_j = 160C, *A* = 268,686 mils^2, D_{yield} = 0.21 (arb. units). Plotted using KMG algorithm, and fitted to lognormal time to failure distributions.



Analysis of Reliability Data

- Fitted lines through Best Estimate and the (one-sided) 95% Upper Confidence Limits for each mechanism gives...

Mechanism	σ	μ Best Est.	μ 60% UCL	μ 90% UCL	μ 95% UCL	μ 99% UCL
PD	5.24	23.94	23.76	23.20	23.05	22.79
FD	11.20	31.33	31.24	30.90	30.78	30.57
BR	8.51	32.81	32.63	32.00	31.81	31.49
JS	3.47	16.00	15.92	15.65	15.58	15.44

At reference condition: $V = 7$ V, $T = 160$C, $A = 268686$ mil^2,
$D_{yield} = 0.21$ (arbitrary units)
(The reference condition *must* be specified.)

Analysis of Reliability Data

- Substitution of parameters into the lognormal distribution gives the “reference” survival function at time *t* in environmental condition “2” for the process:

$$S_i^r(2|t) = 1 - \Phi\left(\frac{\ln[AF_i(2|1)t] - \mu_i}{\sigma_i} \right)$$

where μ and σ for the mechanism are known at the reference condition “1”.

- This would be substituted, for example, into

$$S^p(2|t) = \prod_i [S_i^r\{1|AF_i(2|1)t\}]^{R_i(p|r)} \qquad R_i(p|r) = \frac{P^p(i) \times \ln(Y^p)}{P^r(i) \times \ln(Y^r)}$$

Statistical Interlude: Weibull Analysis

- The reference model can also be fitted to a set of Weibull distributions
  - Characteristic life: α; Shape; β for each mechanism
- Weibull distributions have convenient mathematical properties:

$$W(t,\alpha,\beta) \equiv \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]$$

$$\left[W(t,\alpha,\beta)\right]^{n} = W\left(t, \frac{\alpha}{n^{1/\beta}}, \beta\right)$$

Statistical Interlude: Weibull Analysis

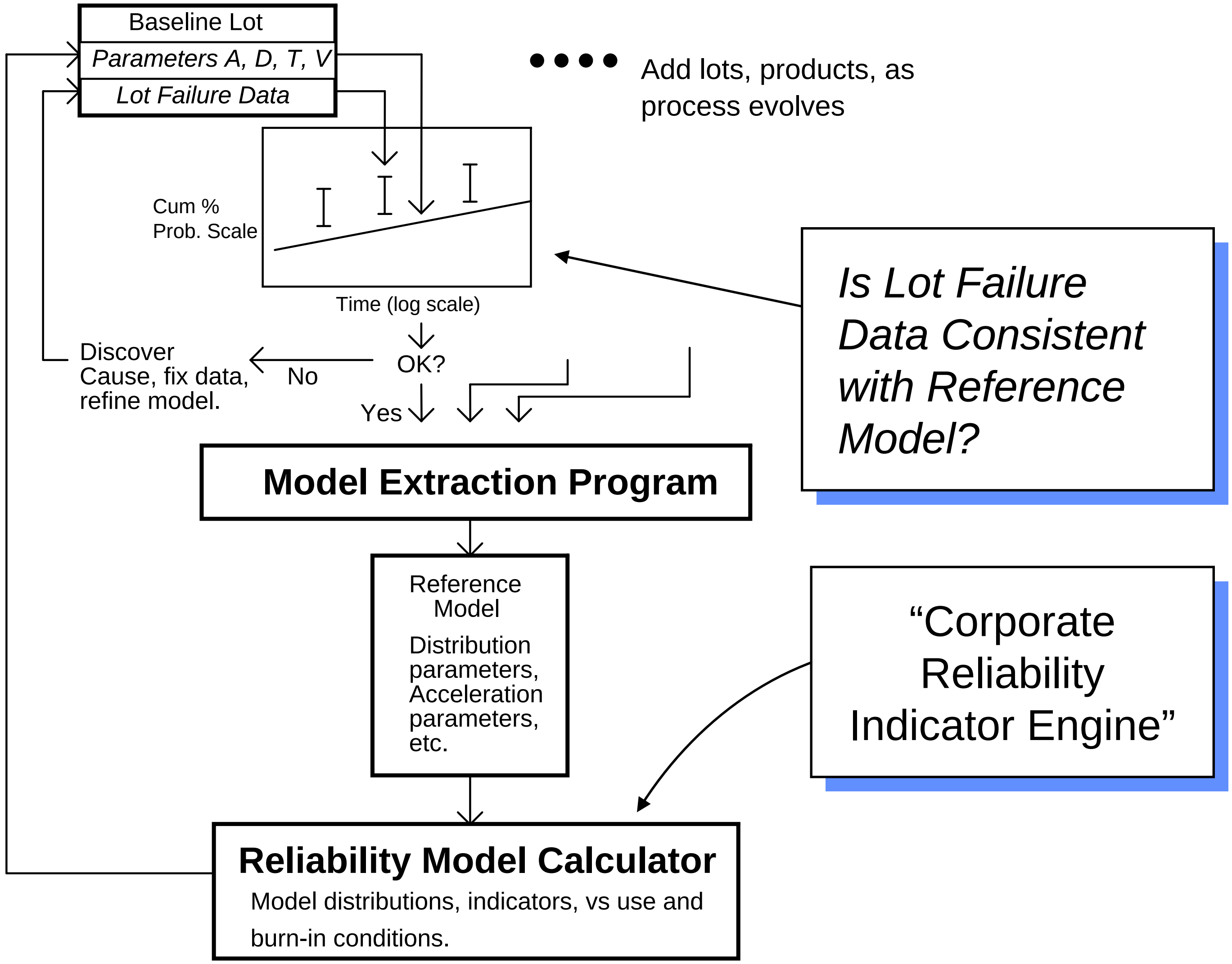
- For example, the product survival function without burn, and for an invariant Pareto, becomes:

Each mechanism has a scaled characteristic life...

$$S^{p}(2|t)=\prod_{i} W\left(AF_i(2|1)t,\ \overbrace{\frac{\alpha_i}{\left(\frac{\ln Y^{p}}{\ln Y^{r}}\right)^{1/\beta_i}}},\ \beta_i\right)$$

...but the same shape parameter as the reference product.

Determining Process Reference Model



Baseline Lot

Parameters A, D, T, V

Lot Failure Data

• • • • Add lots, products, as process evolves

Cum %
Prob. Scale

Time (log scale)

OK?

No

Discover
Cause, fix data,
refine model.

Yes

Is Lot Failure Data Consistent with Reference Model?

Model Extraction Program

Reference Model

Distribution parameters,
Acceleration parameters,
etc.

“Corporate Reliability Indicator Engine”

Reliability Model Calculator

Model distributions, indicators, vs use and burn-in conditions.

Refinement of Process Reference Models

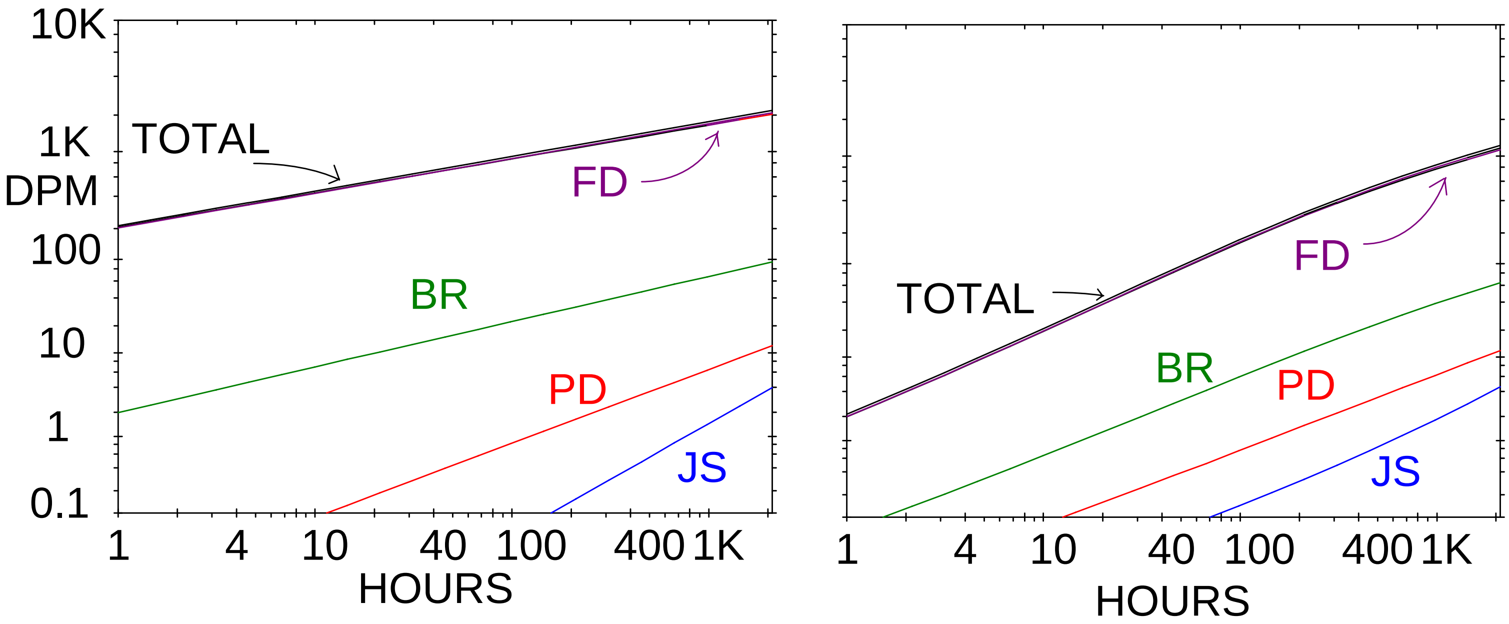
- Add *product* lots to baseline lot set.
  - Reveal mechanisms missed by SRAM model.
- Check for consistency with reference model.
  - Some lots class tested at a single-point (6 hr, 125C, 140%V), full F/A, known lot iso, at a minimum.
  - If failure rates are higher than predicted by model, a “red flag” is indicated.
- Refine the reference model
  - Re-extract using Model Extraction Software.
  - Re-extract and install model
    - Immediately if change is significant.
    - On annual cycle if product is consistent with model.

Reliability Prediction

- Effect of burn in on SRAM reliability.
- Model predictions vs individual lot data from baseline data set.
- Calculation of standard reliability indicators.

Reliability Prediction

- Example: Predicted fallout and effect of burn-in for SRAM (Area = 36160 mil^2, D$_{yield}$ = 1 , V = 5 volts, T_j = 85C)



No Burn-In

After 10 hours 125C, 5.5 volt burn-in

Reliability Prediction

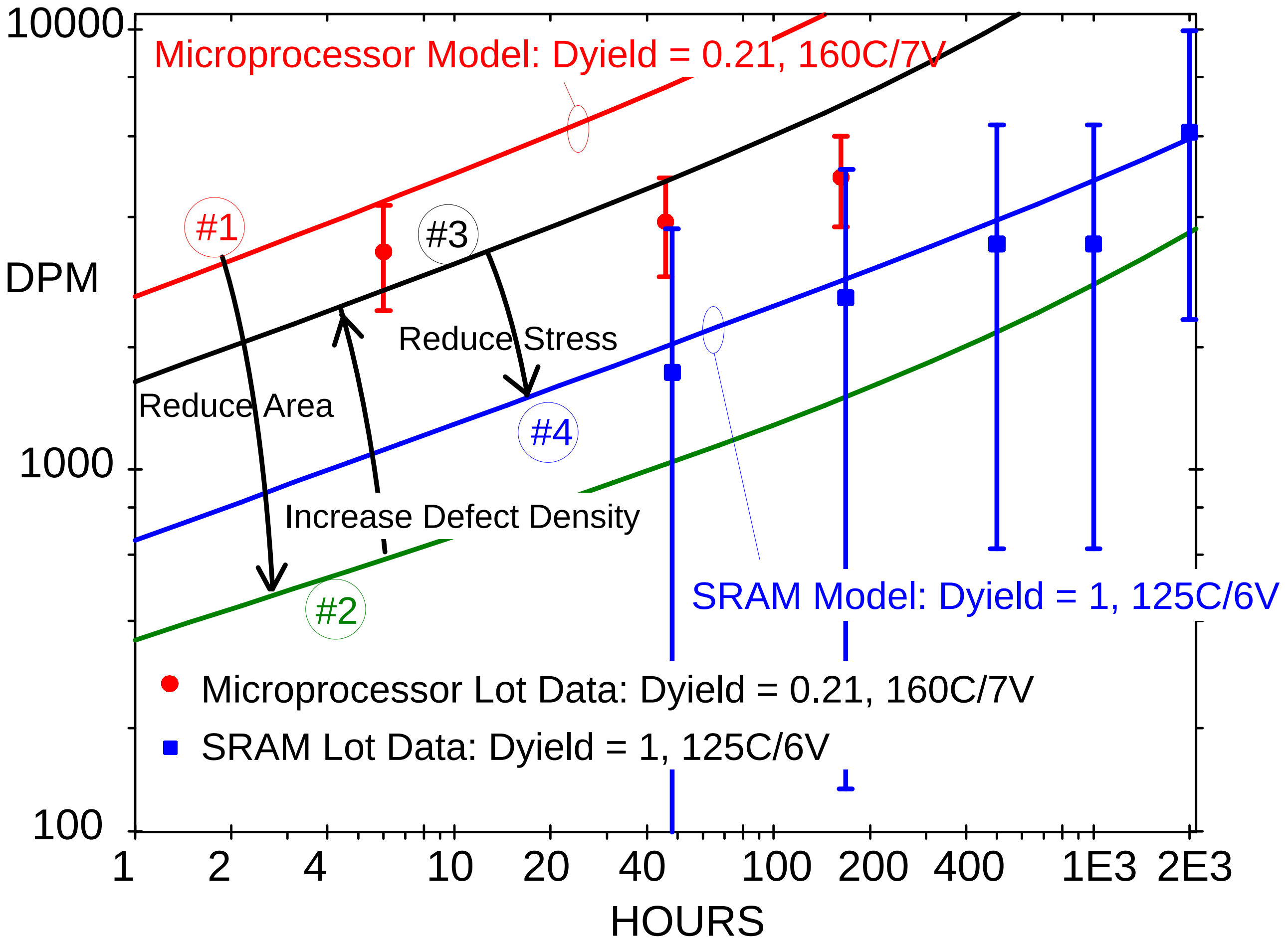
- Model predictions of the reference model based on the entire baseline lot data set *versus* individual data sets selected from the baseline data set.
- A sequence of conditions ranging from conditions of microprocessor data for a particular lot to conditions of SRAM data for a particular lot...

No.	A(mil^2)	Dyld	T(C)	V(volts)	
1	268,686	0.21	160	7	Microprocessor lot data
2	36,160	0.21	160	7	
3	36,160	1.00	160	7	
4	36,160	1.00	125	6	SRAM lot data

Note: Model predictions and data are examples only and are not representative of Intel products.

Reliability Prediction

Model Predictions of Total Failures vs Baseline Lot Data



Reliability Prediction

- Standard reliability indicators
  - Infant Mortality: 0-100 hours at 85C and 5V (DPM)

$$10^6 \times \{1 - S'(t = 100 \text{ hours})\}$$

  - Early Life Mortality: 0 - 1 year at 85C and 5V (DPM)

$$10^6 \times \{1 - S'(t = 8760 \text{ hours})\}$$

  - Early Life Average Failure Rate (AFR): 0-1 year AFR at 85C and 5V (Fits)

$$-10^9 \times \ln[S'(t = 8760 \text{ hours})] / 8760$$

  - Long Term AFR: 1-10 year AFR at 85C and 5V (Fits)

$$10^9 \times \{\ln[S'(t = 8760 \text{ hours})] - \ln[S'(t = 87600 \text{ hours})]\} / 78840$$

Note: Prime indicates “burned-in” survival function.

Reliability Prediction

Reliability Indicators for microprocessor example at T_j = 85C and V = 5 volts, D_{yield} = 0.21, Area = 268,686 mil^2

		CUM FAIL		AFR	
	Mech.	0-100h DPM	0-1yr DPM	0-1yr FIT	1-10yr FIT
No Burn-In	PD	2	69	8	4
	FD	1406	4827	552	48
	BR	39	305	35	6
	JS	0	42	5	6
	Total	1447	5241	600	65
Burn-In: 168 hr at 160C/7V	PD	0.4	35	4	3
	FD	1.6	133	15	13
	BR	0.5	45	5	4
	JS	0.6	52	6	6
	Total	3.1	266	30	25

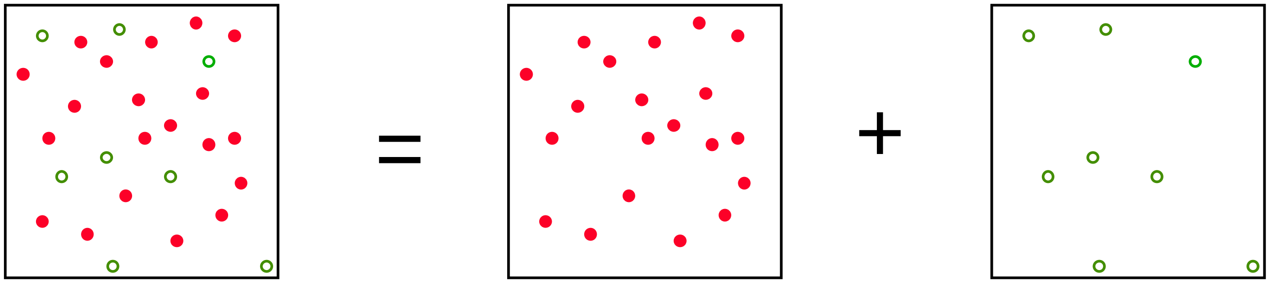
Benefits

- Estimation of the reliability characteristics of any product, including the contributions of various mechanisms.
- Estimation of failure rates of complex products without full reliance on failure analysis, or complete data.
- Estimation of the effect of die area, array area, etc. on the reliability characteristics of any proposed or new product using no, or minimal, data.
- Quantify the reliability benefits of process continuous improvement through defect density reduction.
- Calculate the effect of burn-in.
- Calculate reliability indicators useful to customers, at any desired level of confidence.

Supplementary Slides on Clustering Effects

Effects of Defect Clustering

- Random defects:

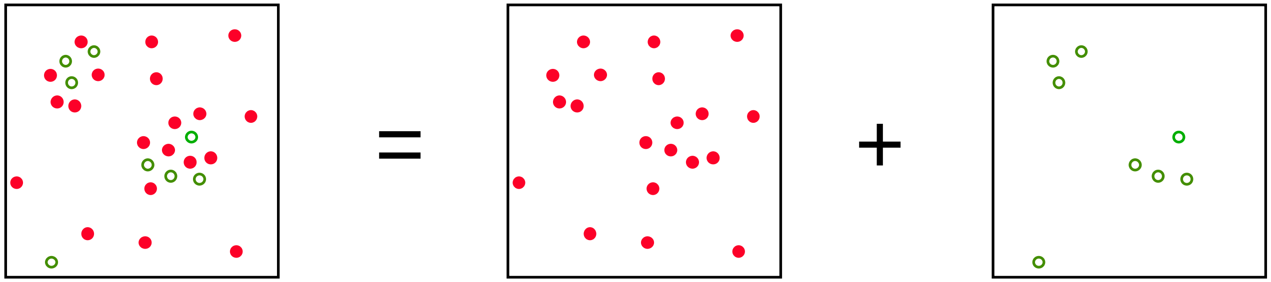


Total Defect Density

Yield Defect Density

Reliability Defect Density

- Clustered defects:



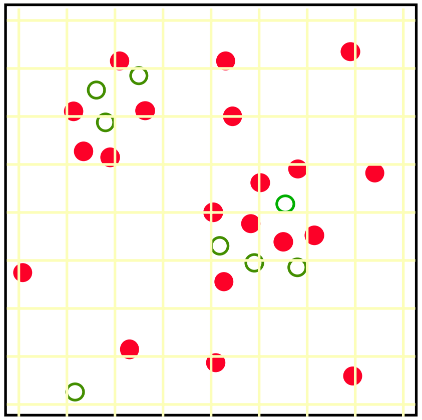
Total Defect Density

Yield Defect Density

Reliability Defect Density

Defect Density Variation

- Clustering can be modeled as a spatial variation of of defect density.
- The clustering can be described by a gamma function distribution:

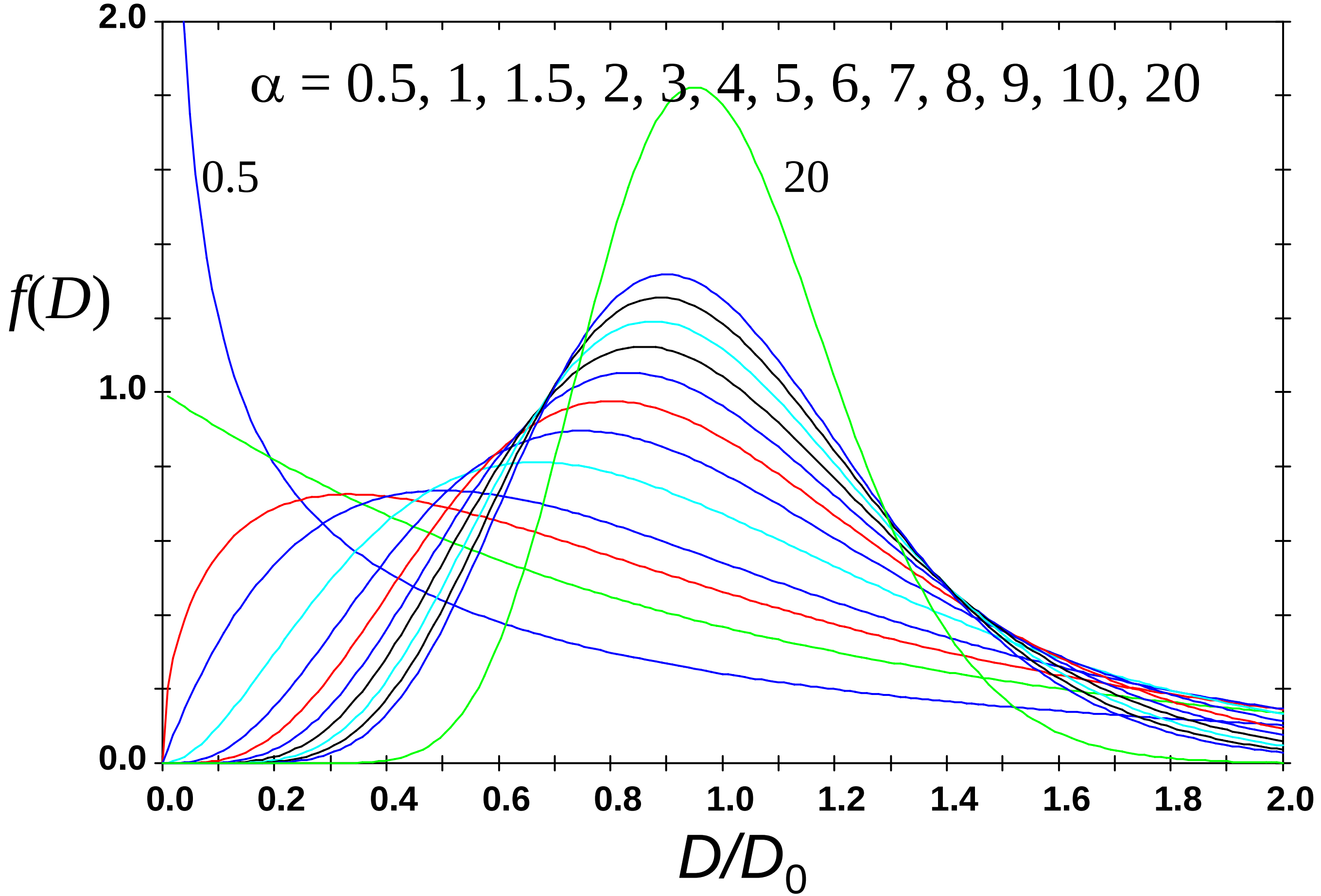


$$f(D) = \frac{\alpha}{D_0 \times \Gamma(\alpha)} \left(\alpha \frac{D}{D_0} \right)^{\alpha - 1} \exp\left(-\alpha \frac{D}{D_0} \right)$$

- The spread in the defect density is described by $\alpha = \mathrm{var}(D)/D_0^2$
- D_0 is the average defect density (defects/cm^2)

Defect Density Variation

- The defect density distribution approaches a delta function as $\alpha \rightarrow \infty$.



Yield Function with Clustering

- The yield function is the probability of occurrence of one defect on a die of area A:

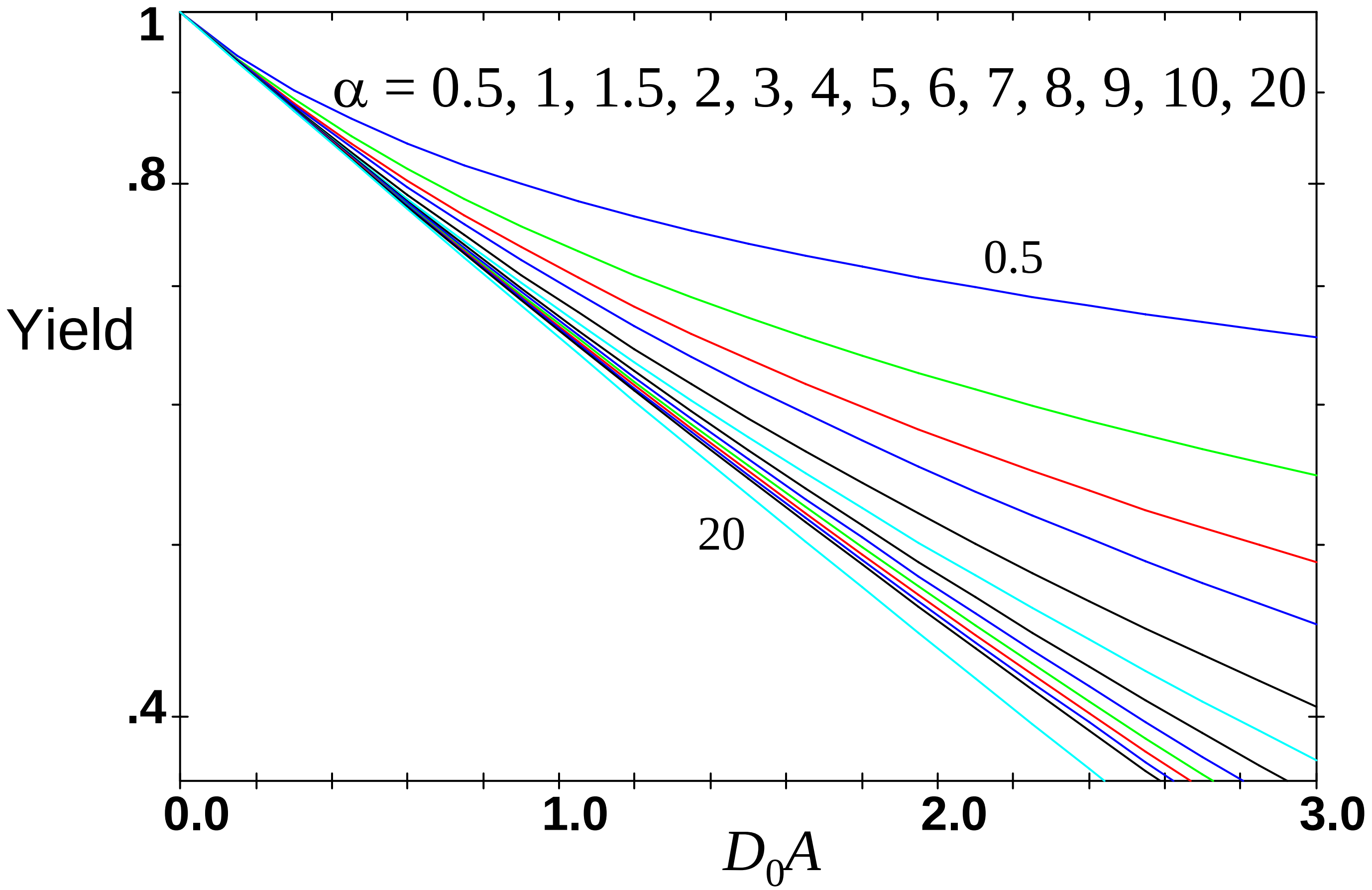
$$Y = \int_0^\infty \exp(-DA) f(D) dD = \frac{1}{\left(1+\frac{D_0 A}{\alpha}\right)^\alpha}$$

- In the limit of no clustering (uniform D), this becomes

$$\frac{1}{\left(1+\frac{D_0 A}{\alpha}\right)^\alpha} \xrightarrow[\alpha\to\infty]{} \exp(-D_0 A)$$

Yield Function with Clustering

- Clustering of defects gives higher yields than predicted by random defect model...



Clustering of Latent Reliability Defects

- How does the chip reliability survival function vary with non-uniform defect density?
- Define $s(t)$, the point defect survival function.
- For uniform defects the chip survival function is $S(t) = [s(t)]^{AD}$, where AD is the number of defects on the chip.
- If defects are clustered..

$$S(t) = \frac{1}{\left(1 - \frac{AD_0 \ln s(t)}{\alpha}\right)^{\alpha}}$$

Scaling of Survival Probability

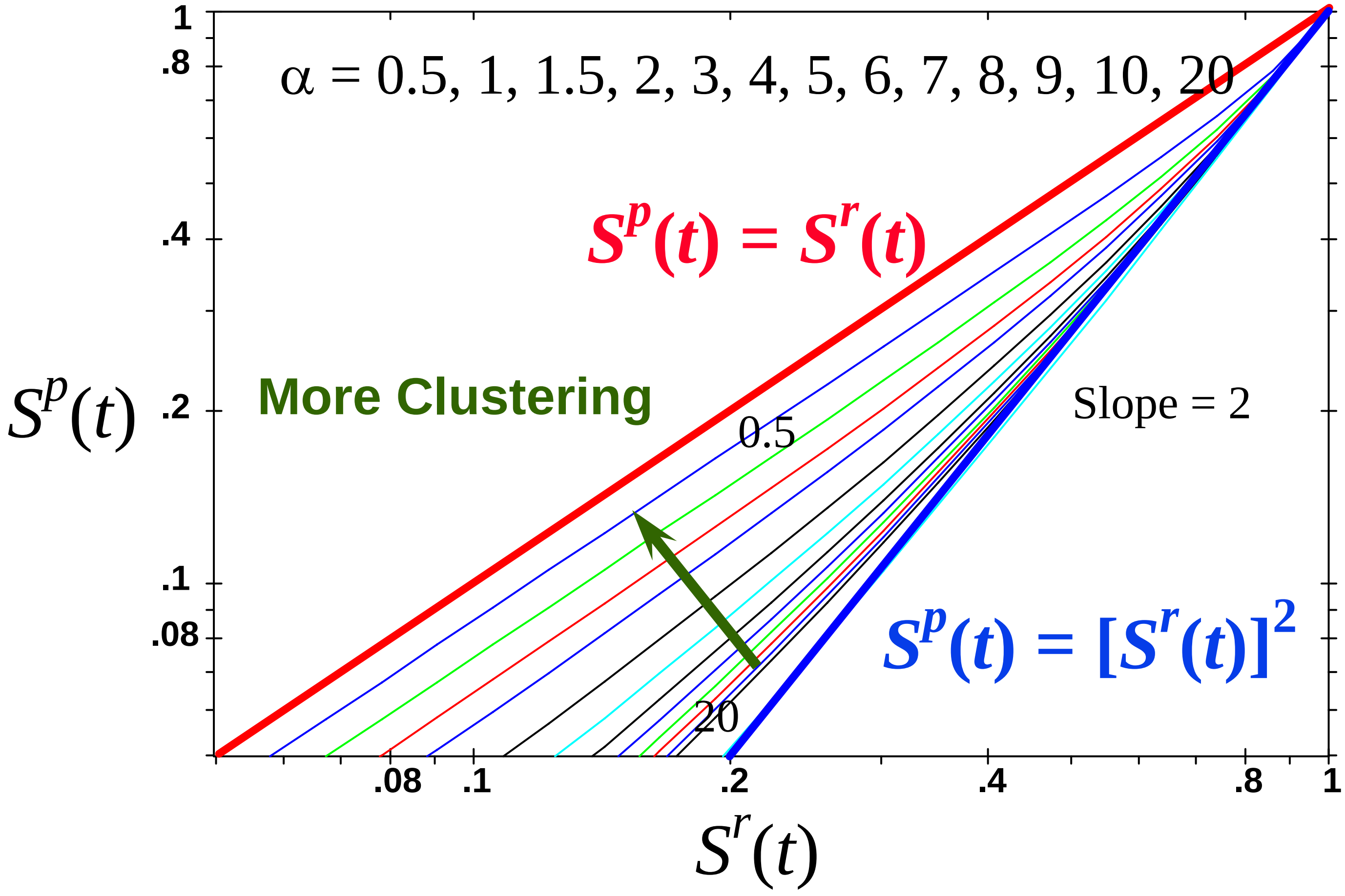
- Consider
  - A product with *unknown* survival probability $S^p(t)$ and,
  - A reference test vehicle with *known* survival probability $S^r(t)$.
  - The defect density variation, α, is the same for both unknown product and reference test vehicle
  - The "scaling ratio" is, in terms of reliability defect density, $R_{\text{rel}}(p|r) = \dfrac{A^p D^p_{\text{rel}\,0}}{A^r D^r_{\text{rel}\,0}}$
- So the unknown survival probability is

$$S^p(t) = \left\{1 + R_{\text{rel}}(p|r) \times \left([S^r(t)]^{-\frac{1}{\alpha}} - 1 \right) \right\}^{-\alpha}$$

$$\xrightarrow[\alpha \to \infty]{} [S^r(t)]^{R(p|r)}$$

Survival Function Scaling

- For the case where product p has twice the area (or avg. defect density) of the reference product:



Yield-Reliability Relationship

- From the yield formulae

$$R_{\text{yield}}(p|r) = \frac{A^p D^p_{\text{yield }0}}{A^r D^r_{\text{yield }0}} = \frac{(Y^p)^{-\frac{1}{\alpha}} - 1}{(Y^r)^{-\frac{1}{\alpha}} - 1} \xrightarrow[\alpha\to\infty]{} \frac{\ln Y^p}{\ln Y^r}$$

- I f we make the fundamental assumption

$$R_{\text{reliability}}(p|r) = \frac{A^p D^p_{\text{reliability }0}}{A^r D^r_{\text{reliability }0}} \cong \frac{A^p D^p_{\text{yield }0}}{A^r D^r_{\text{yield }0}} = R_{\text{yield}}(p|r)$$

- *And* assume the dispersion in reliability and yield defect densities are the same

$$\alpha = \alpha_{\text{reliability}} \cong \alpha_{\text{yield}}$$

Yield-Reliability Relationship

- Then we can calculate the product survival function from yield characteristics

$$S^p(t) = \left\{ 1 + \frac{(Y^p)^{-\frac{1}{\alpha}} - 1}{(Y^r)^{-\frac{1}{\alpha}} - 1} \times \left([S^r(t)]^{-\frac{1}{\alpha}} - 1 \right) \right\}^{-\alpha}$$

$$\xrightarrow[\alpha \to \infty]{} [S^r(t)]^{\frac{\ln Y^p}{\ln Y^r}}$$

For a given defect density, more clustering gives higher yield *and* higher reliability.

Extension to Multiple Mechanisms

$$S^p(t) = \prod_i \left\{ 1 + \frac{(Y_i^p)^{-\frac{1}{\alpha_i}} - 1}{(Y_i^r)^{-\frac{1}{\alpha_i}} - 1} \times \left([S_i^r(t)]^{-\frac{1}{\alpha_i}} - 1 \right) \right\}^{-\alpha_i}$$

$$= \prod_i \left\{ 1 + \frac{P_i^p[(Y^p)^{-\frac{1}{\alpha}} - 1]}{P_i^r[(Y^r)^{-\frac{1}{\alpha}} - 1]} \times \left([S_i^r(t)]^{-\frac{1}{\alpha}} - 1 \right) \right\}^{-\alpha}$$

$$\xrightarrow[\alpha\to\infty]{} \left(\prod_i S_i^r(t)^{\frac{P_i^p}{P_i^r}} \right)^{\frac{\ln Y^p}{\ln Y^r}}$$

Y^p and Y^r are the total yields (all mechanisms).

$P^p{}_i$ and $P^r{}_i$ are yield Paretos.

α is the defect density dispersion parameter (all mechanisms)