Sandia National Laboratories

Exceptional service in the national interest

IMPROVED YIELD ESTIMATION FOR HETEROGENEOUSLY INTEGRATED COMPONENTS

For Continuous and *In Situ* monitoring

Andrew Ferris, Matthew Jordan

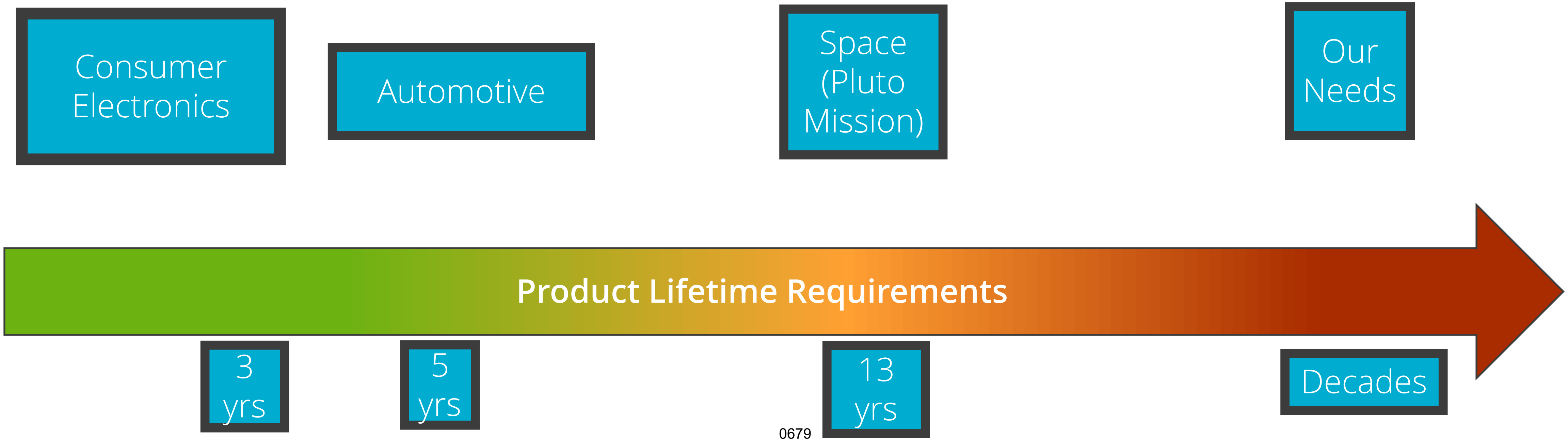
IMAPS Device Packaging Conference 2024,

March 20th 2024, Fountain Hills, AZ, USA

SAND2024-03292C

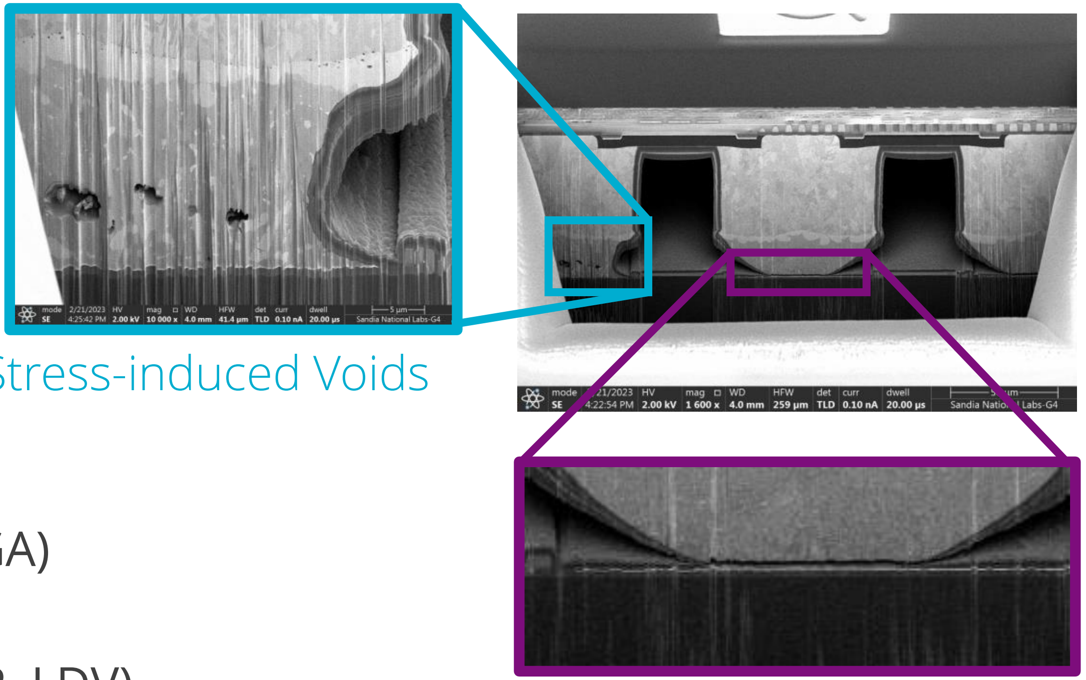
MOTIVATION: RELIABILITY REQUIREMENTS

- We're interested in extremely long lifetime components (decades)
- Low Volume, High Consequence manufacturing:
  - Limited development lots.
  - Require high confidence in a process with very few development lots.



MOTIVATION: SINGLE POINTS OF FAILURE & DCS

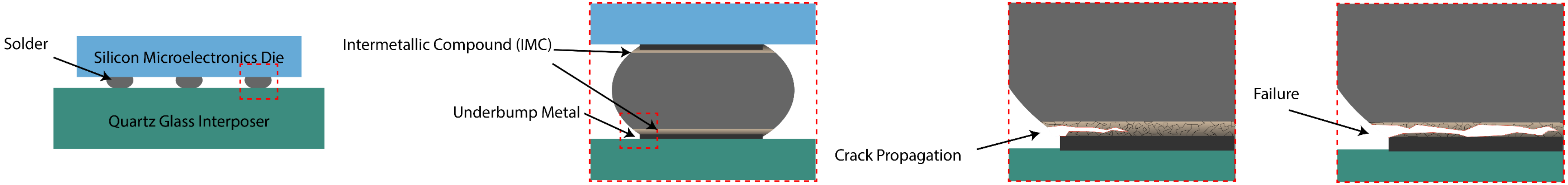
- Mechanisms of Failure:
  - Fatigue
  - Stress
  - Voids
  - Whiskering
  - Electromigration



Stress-induced Voids

Oxidization = Failure

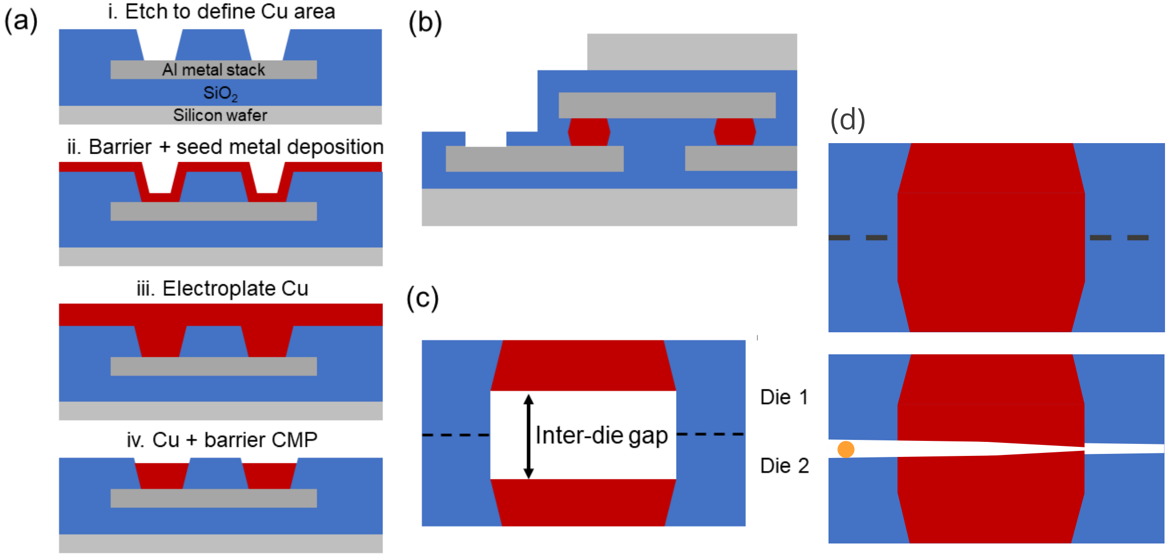
- Failure Analysis (FA) Methods :
  - Scanning Optical Probe (TIVA, LIGA)
  - EM Analysis (PSA, MFI)
  - Thermomechanical Probes (FDTR, LDV)
  - Destructive Testing (HAST, Cross-sectioning)



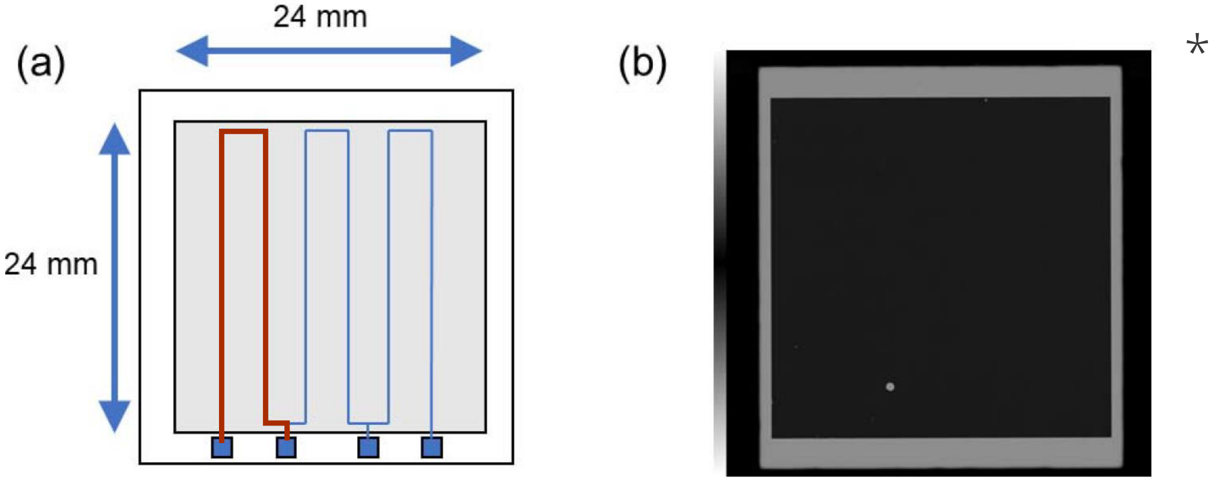
0680

MOTIVATION: RANDOM FAILURES.

- Not fully random:
  - Incremental stress is focused on corners (fatigue)
- Random Failures:
  - Processing Defects: particles, PR variation, etc.



- Hybrid Bonding: Particle contamination prevents bonding
  - 10,000:1 failure size: 1μm particle = 1cm void.
  - A recent example:



- 1/6 Daisy Chains Failed : Y = .99994%
- CSAM reveals one particle ($\sim 50\text{nm}$) resulting in 0.20mm^2 void: ~ 70 failed interconnects. 0.998%

* J.P.Mudrick, et al (2019), used with permission.

USAGE OF DAISY CHAINS (DC) IN UNCERTAIN YIELD

Ullan (2006) Approximation: $Y = \frac{1}{L}\log\left(1 - \frac{M}{w}\right) + 1$

Fewer, Longer Chains:

- Effective for high-yield processes
- Limited use while Yield is Low

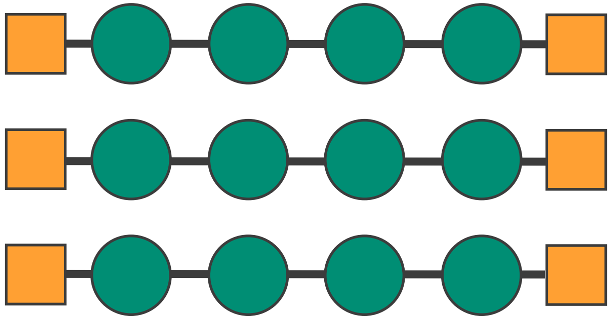
More, Shorter Chains:

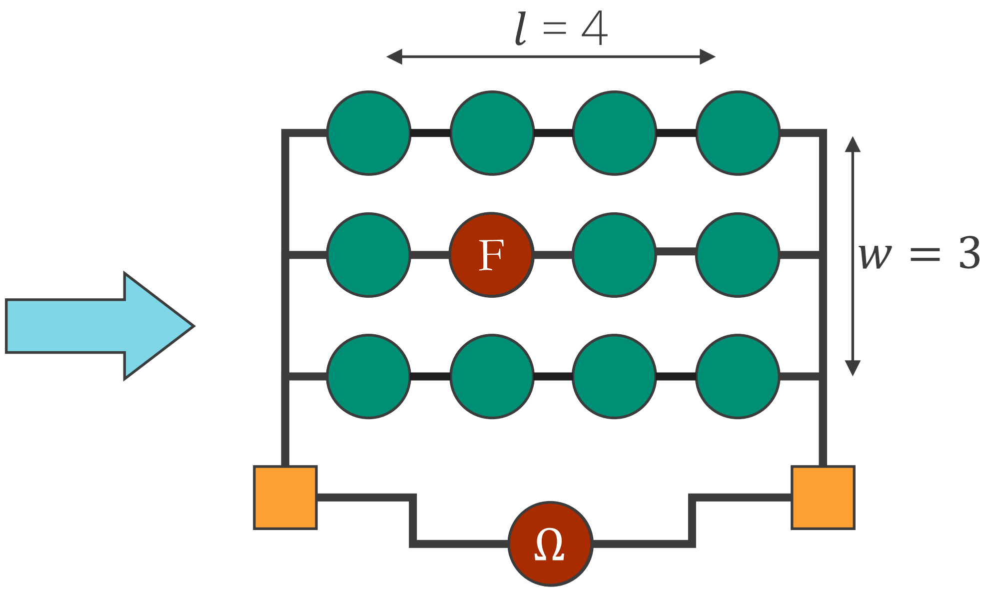
- More effective for low-yield processes
- Decreased Statistical Relevance
- Increased Testing Requirements.

Limitations on both

- Information is Lost due to open circuit.,
- Can only take one data point per experiment – requires many tests to get statistically relevant results from multiple process conditions.

PARALLELIZATION



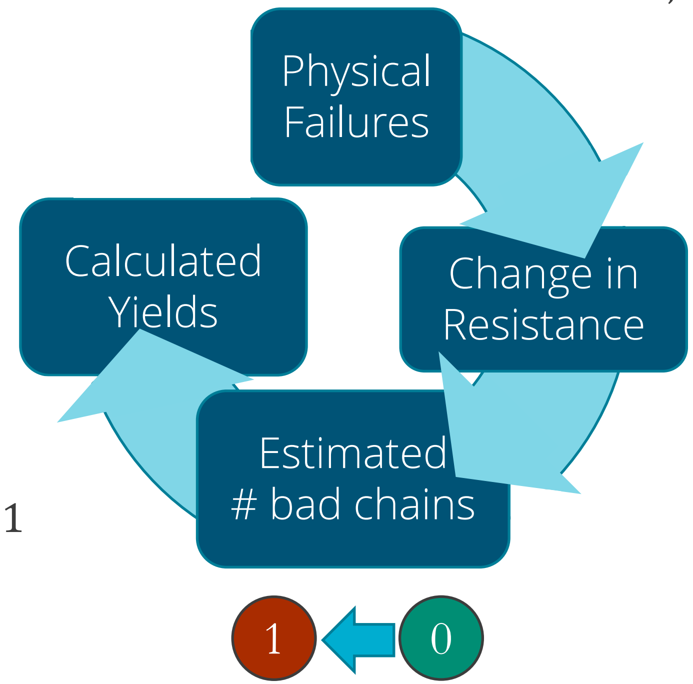


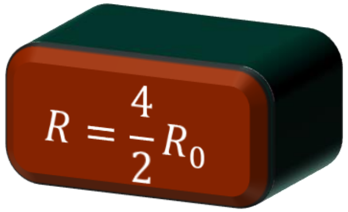
Neglecting wire resistance

$$R = \left(\sum_w \frac{1}{\sum_l R_{l,w}}\right)^{-1}$$

$$R_{l,w} = \begin{cases} R_0, & N_f \in l = 0 \\ \infty, & SPoF \end{cases}$$

- Failures are often single points of failure (SPoF) in DCs.
- Can monitor from a single output resistance
- Measurement simplicity enables continuous monitoring:
- Lose ability to identify precise position of failures.

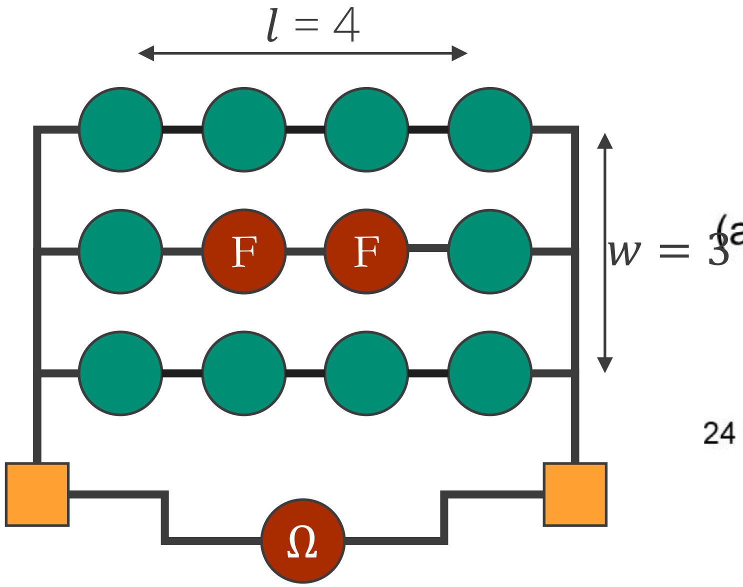




$$Y = \frac{1}{L}\log\left(1 - \frac{M}{w}\right) + 1$$

PARALLELIZATION: OPEN CIRCUIT INFORMATION LOSS

$$R = \frac{4}{2} R_0$$

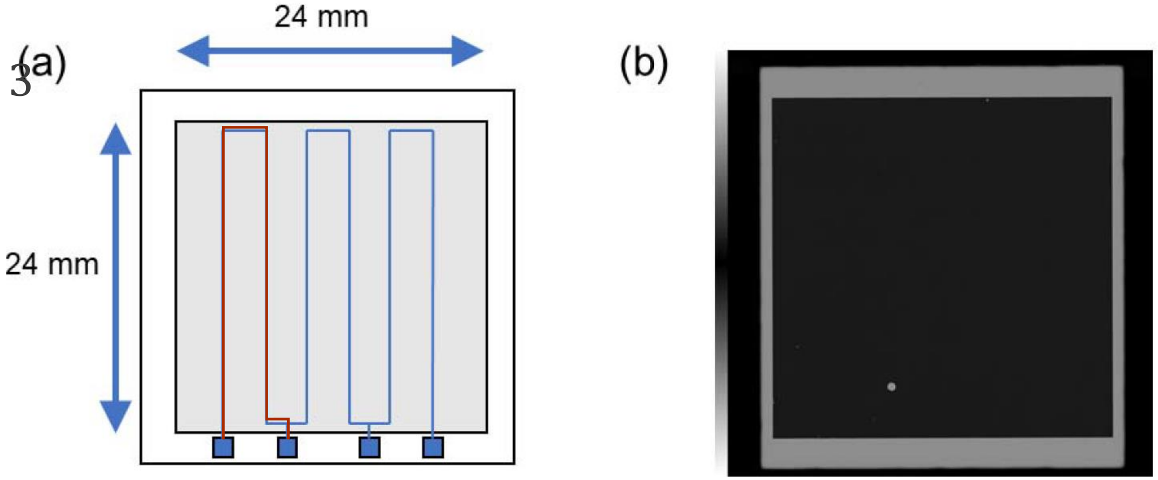


- DC Failures produce open circuits.
- Unable to collect further info from chain → loss of precision.

$Y_{act} = 92\%$ → $Y_{act} = 83\%$

$Y_{est} = 90\%$ (-2%) → $Y_{est} = 90\%$ (+7%)

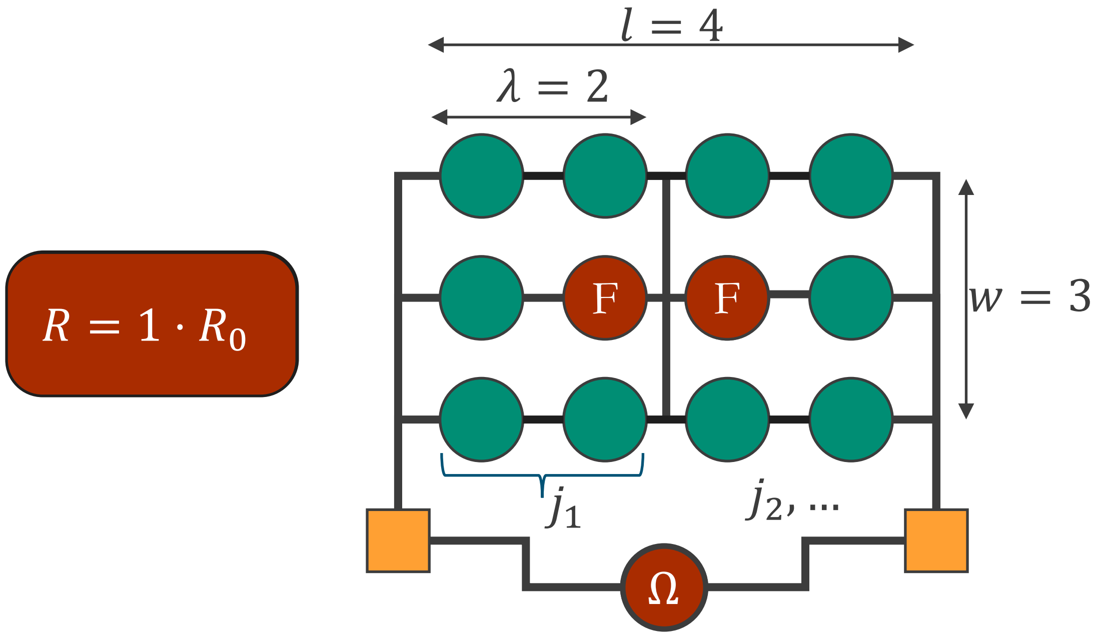
A recent example: *



- Open Circuit = $1 < X < 6600$ failures.
- Resolving scope of failures requires incorporating other tools such as CSAM, indicating $N_f \sim 70$.

* J.P.Mudrick, et al (2019), used with permission.

PARALLELIZATION: SHORTING CONNECTIONS TO MITIGATE LOSS



- By Shorting DCs together between elements into columns of chains j, we mitigate information loss.
- Greater Precision

$Y_{act} = 92\%$ → $Y_{act} = 83\%$

$Y_{est} = 91\%$ (+1%) → $Y_{est} = 80\%$ (-3%)

- By Solving the following for all $N_{f,j}$

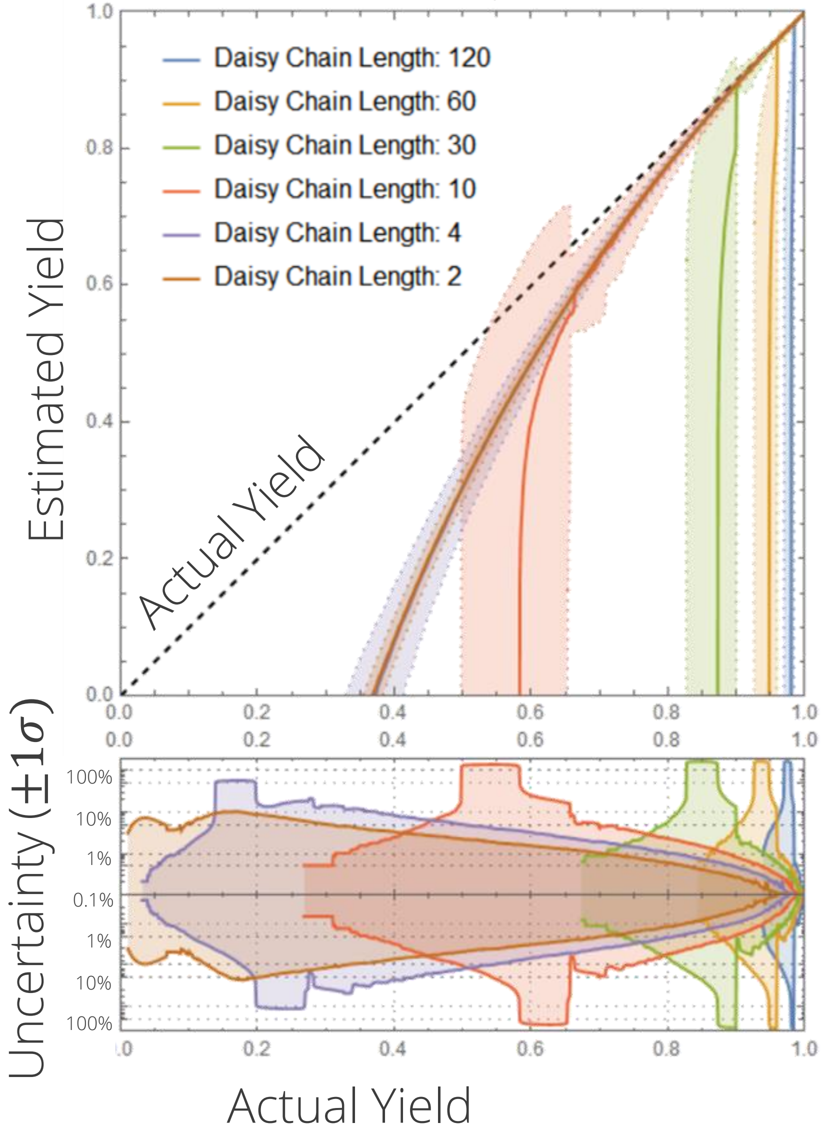
$$\Omega = \sum_{j=0}^{wid} \frac{N_{f,j} \cdot R_0}{wid - j}$$

- Subject to the following constraints:
  - $N_{f,j} \in \mathbb{Z}^+$ (SPoF are total failures)
  - $\sum_j N_{f,j} = l$ (#sets = length)
  - $\sum_j j \cdot N_{f,j} \leq w * l$ (#failed bumps ≤ #bumps)
  - If $R < R_{co}$, then $N_{f,j=w} = 0$ (finite resist = no open circuit)

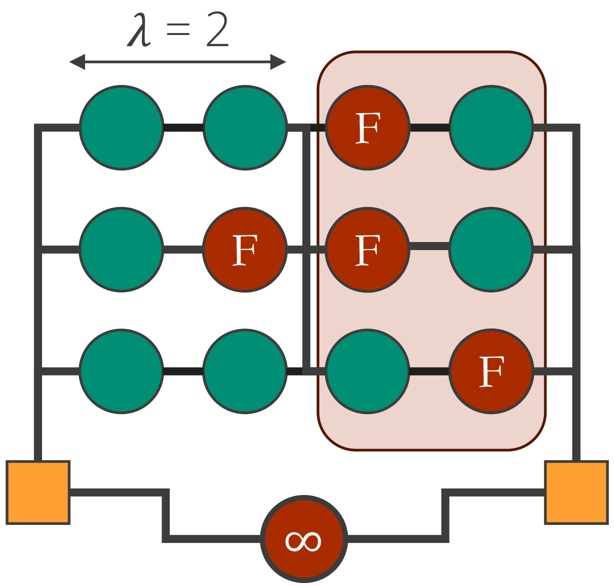
PARALLELIZATION: PRECISION COMPARISON

- Precision and utility domain both improve as the system is increasingly parallelized
- In the limit $\lambda \rightarrow 1$, Precision → Exact Yield from a single measurement.
- With two caveats:
  - If a column of parallelized bumps have all failed, the system still fails → minimum yield for a given configuration.
  - There are degenerate solutions, weakening precision.

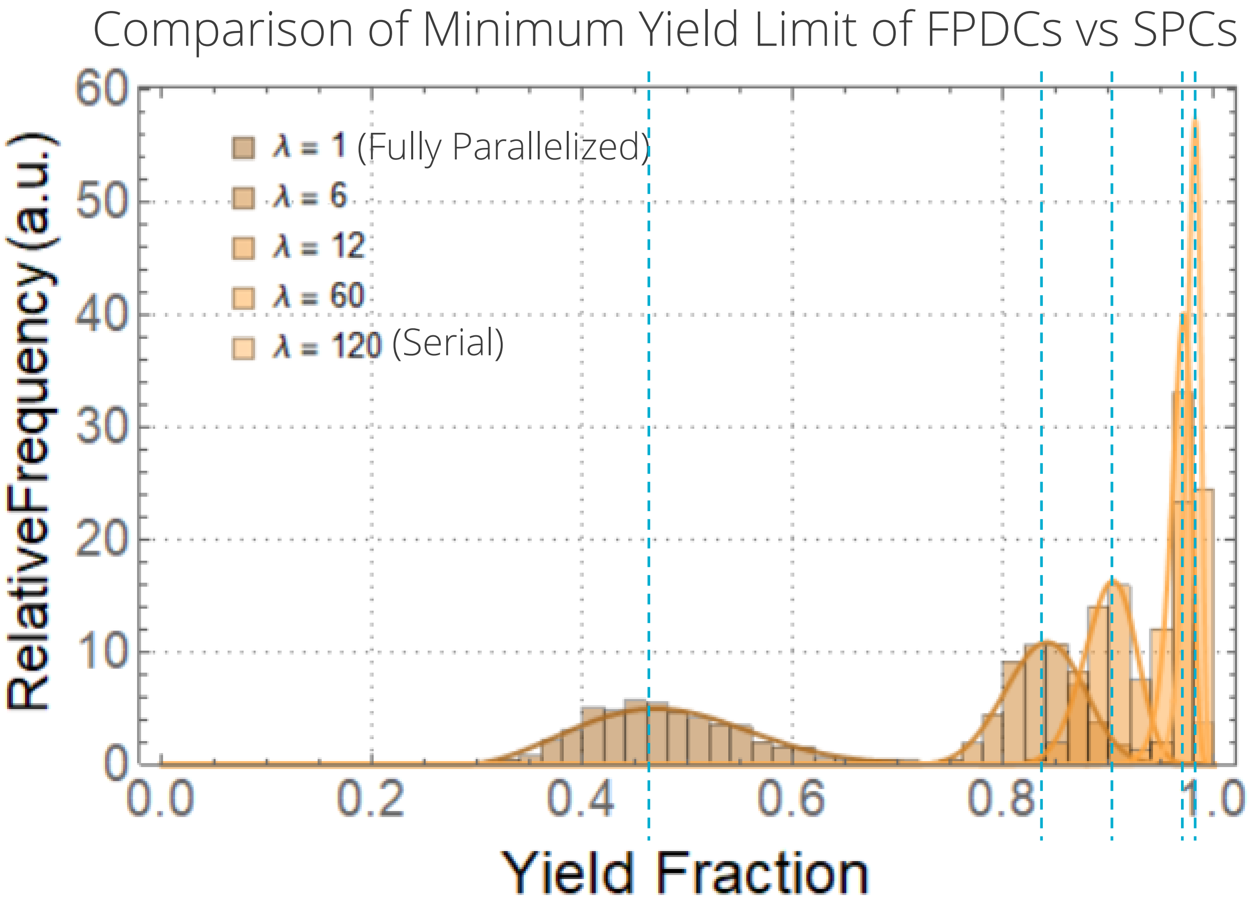
Simulation of $l = 120, w = 8$ Parallelized Daisy Chains



CAVEAT 1: FAILURE DEPTH



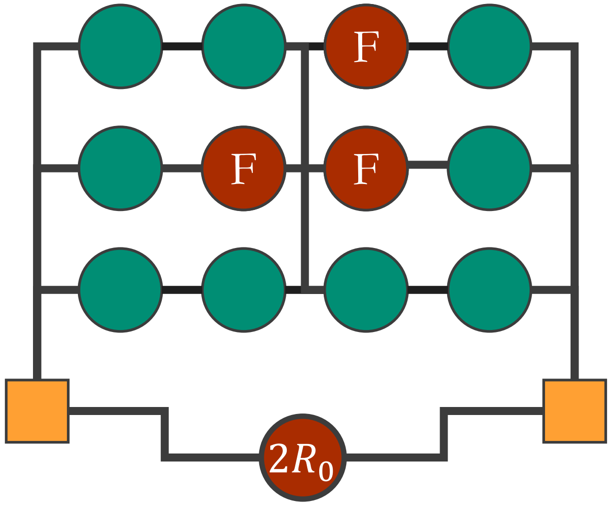
- This Parallelization (λ / l) drastically lowers the minimum yield before which the DC is no longer able to probe the system due to failures.
- The larger w is, the lower this minimum yield limit.
- Minimum yield limit well-described by Weibull Distribution

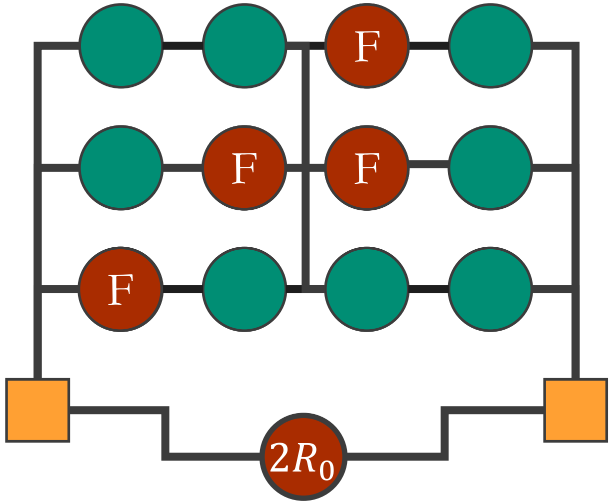


	Weibul α :	Weibull β :	Weibull μ :	Mean
$\lambda = 120$	58.16	0.373	0.608	Y = 0.978
$\lambda = 60$	37.49	0.434	0.627	Y = 0.966
$\lambda = 6$	3.883	0.137	0.714	Y = 0.837
$\lambda = 1$	2.613	0.212	0.293	Y = 0.481

CAVEAT 2: RESOLVING DEGENERACY

$$R = 2\,R_0 \;\rightarrow N_f = \{2,3,4,5,6,7,8\}$$





- By symmetry, reduce $R_{l,w} \rightarrow R_j$ resistance of a column of bumps with $N_f = j$ failed chains of λ elements.

$$R_j = \frac{\lambda \cdot R_0}{w - j}$$

- We introduce degenerate solutions to

$$\Omega = \sum_{j=0}^{wid} N_{f,j} R_j$$

- By adding additional constraints based on transitions between states $s \rightarrow s+1$, we eliminate the majority of degenerate solutions
- SPoF don't fix themselves: N_f cannot decrease

$$\sum_j N_{f,j}(s) \leq \sum_j N_{f,j}(s+1)$$

- Every Failure increases Resistance

$$\sum_j R_{f,j}(s) \leq \sum_j R_{f,j}(s+1)$$

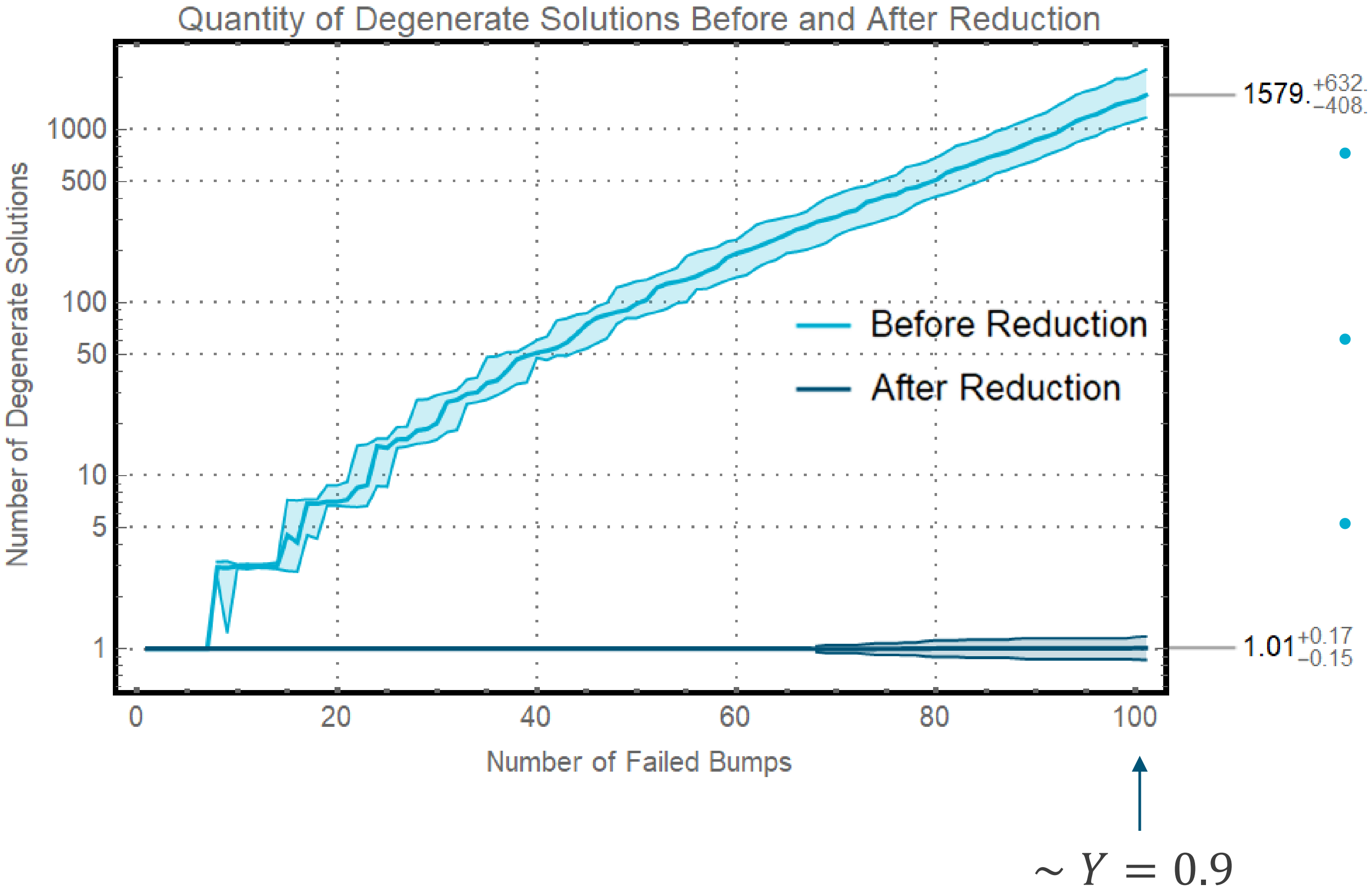
- Self-Consistency:
  - If a bump fails in set $N_{f,j}$, it must now be a member of some set $N_{f,k} : k > j$.

$$-\left(\Delta N_{f,j}(s, s+1)\right) \leq \sum_{k>j} \Delta N_{f,k}(s, s+1)\ \forall j$$

  - If a bump failure elevated a set into set $N_{f,j}$, it must have been a member of some set $N_{f,k}$: $k < j$ previously.

$$-\left(\Delta N_{f,j}(s, s+1)\right) \leq \sum_{k<j} \Delta N_{f,k}(s, s+1)\ \forall j$$

CAVEAT 2: OVERCOMING DEGENERACY



- By adding additional constraints based on transitions between states $s \to s+1$, we eliminate the majority of degenerate solutions
- SPoF don't fix themselves: N_f cannot decrease

$$\sum_j N_{f,j}(s) \leq \sum_j N_{f,j}(s+1)$$

- Every Failure increases Resistance

$$\sum_j R_{f,j}(s) \leq \sum_j R_{f,j}(s+1)$$

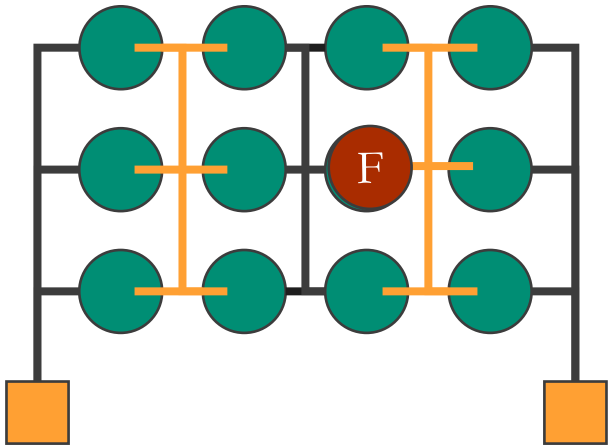
- Self-Consistency:
  - If a bump fails in set $N_{f,j}$, it must now be a member of some set $N_{f,k} : k > j$.

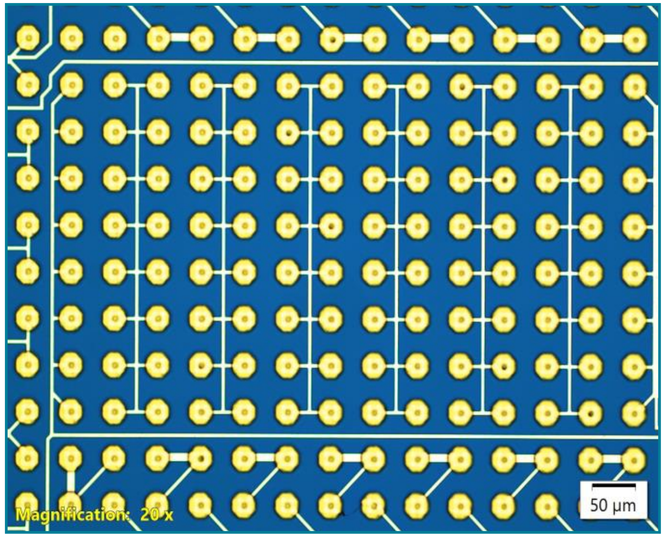
$$-\left(\Delta N_{f,j}(s, s+1)\right) \leq \sum_{k>j} \Delta N_{f,k}(s, s+1) \ \forall j$$

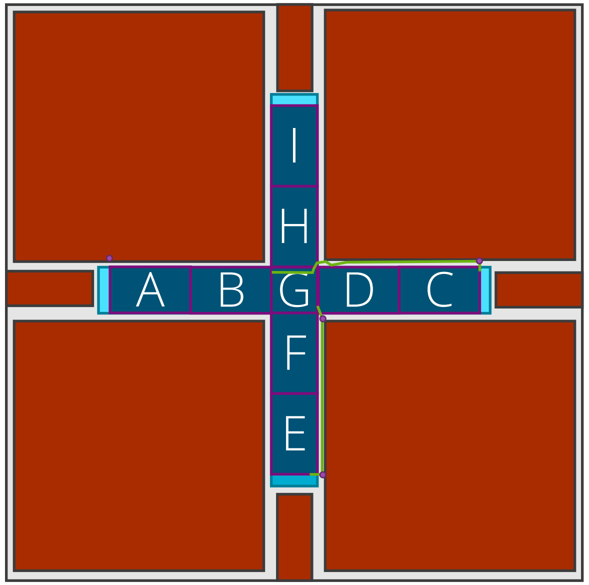
  - If a bump failure elevated a set into set $N_{f,j}$, it must have been a member of some set $N_{f,k}$: $k < j$ previously.

$$-\left(\Delta N_{f,j}(s, s+1)\right) \leq \sum_{k<j} \Delta N_{f,k}(s, s+1) \ \forall j$$

FULL PARALLELIZATION: $\lambda \rightarrow 1$



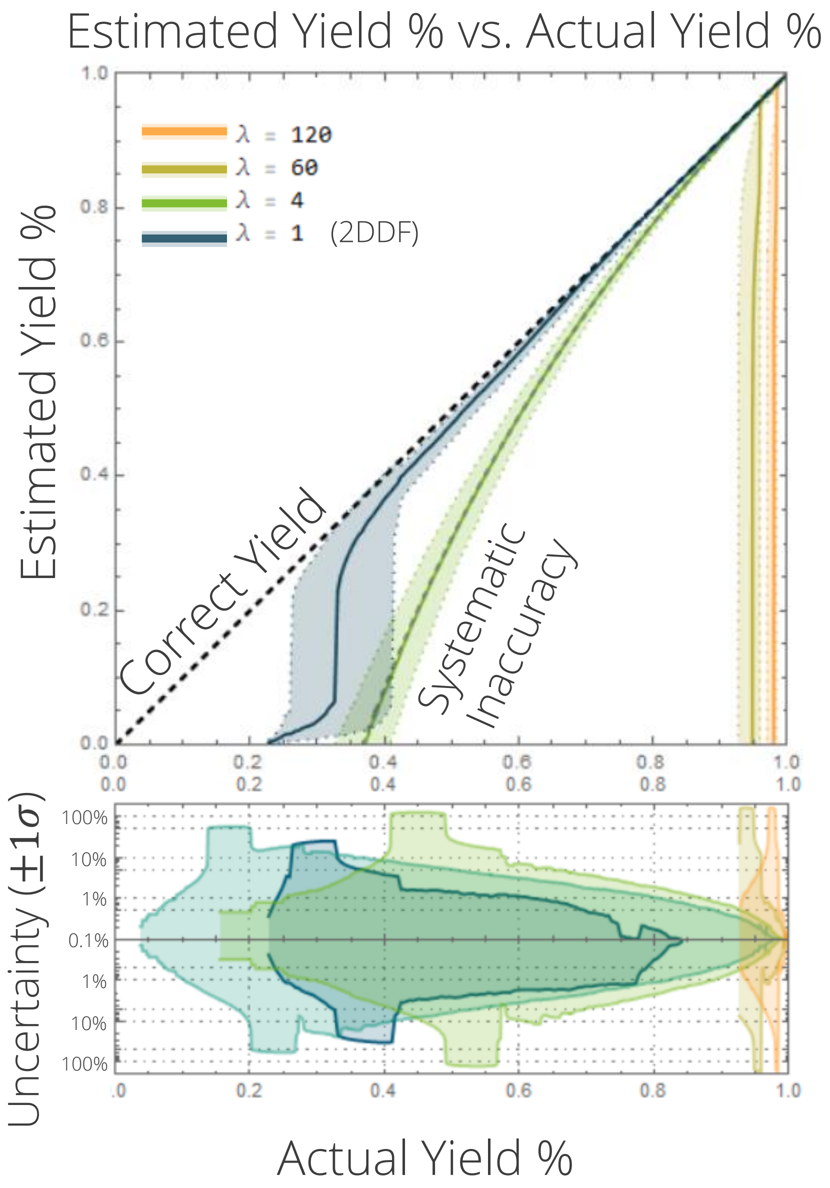




- 2D Daisy 'Fields' (2DDFs) = Can make direct estimates of yield.
- Full Parallelization = Similar quality analysis in any configuration – fit onto any apparatus.
- Sensitivity Requirement: $\Delta R < \frac{R_0}{l*w}$

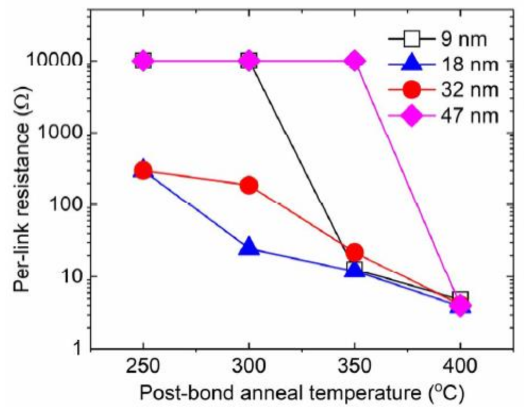
PARALLELIZATION: IMPROVED YIELD ANALYSIS

- 2DDFs provide better accuracy and precision for any circuit footprint.
- Yield Estimate is valid for all yield processes
  - Ideal for processes with low and unknown yields.

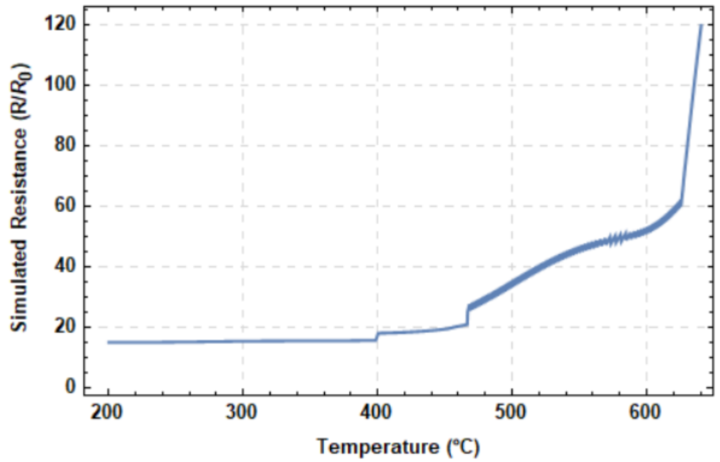


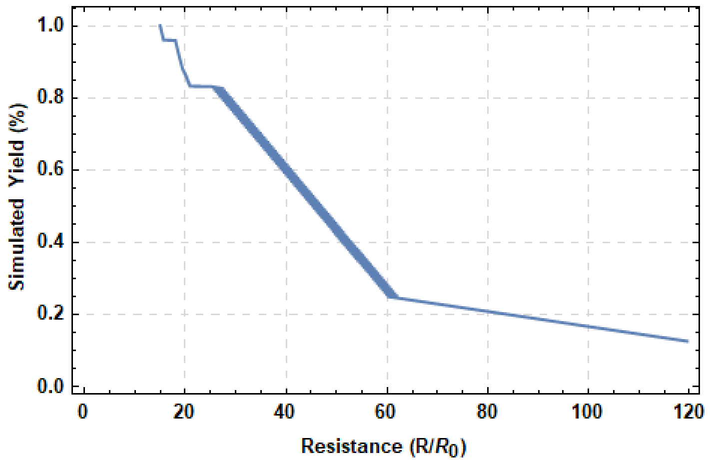
CONTINUOUS MONITORING

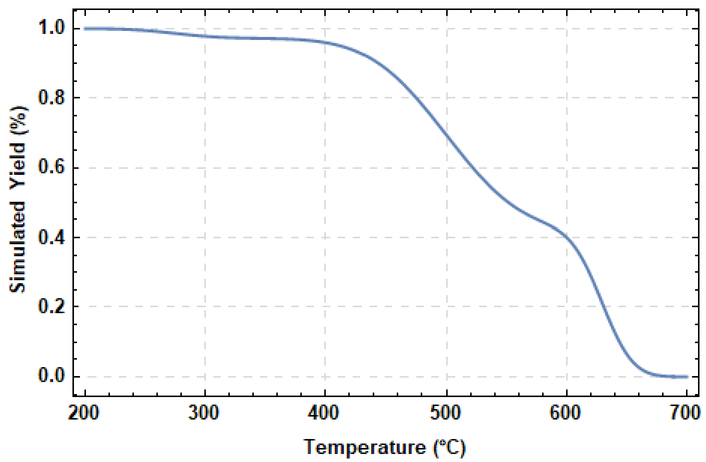
- Continuous Yield Analysis = Can monitor *when* failures happen.
- Can identify source of failures and thus determine process limits
- All in a single experiment, no more guess and check repeatedly (→).
- Current iteration of program is not suited for real-time application
  - Workaround: "pre-bake" a confidence interval for "Resistance → Yield" to get yield estimates in "real time" as a lookup table function.



* J.P.Mudrick, et al (2019), used with permission.

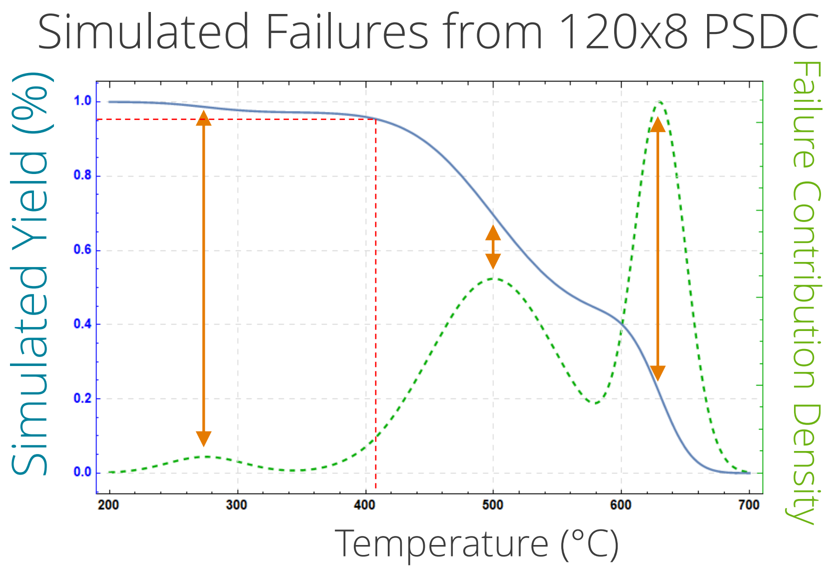


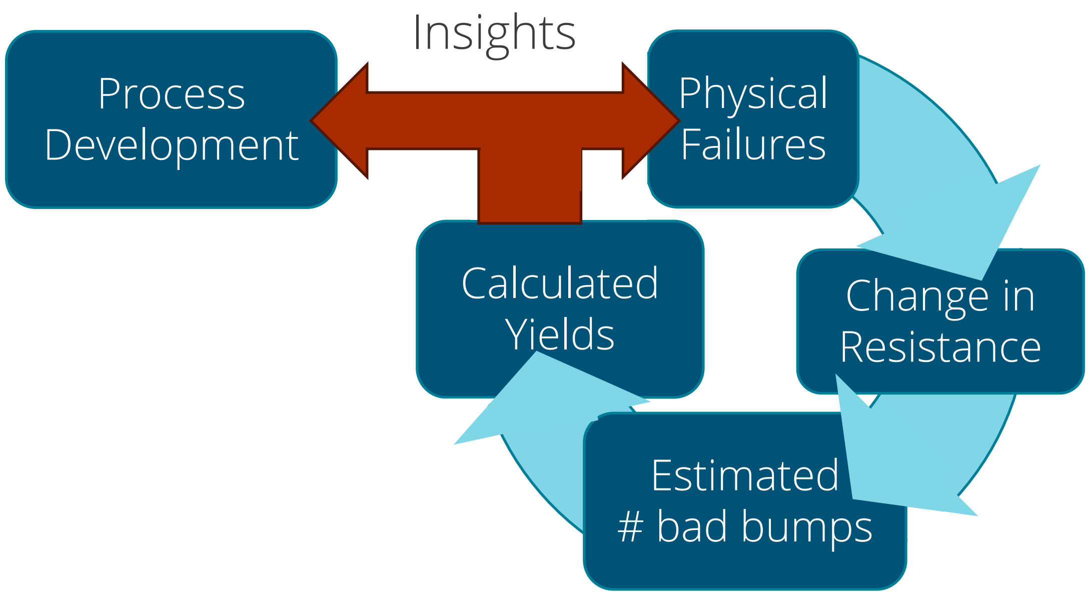




CONTINUOUS MONITORING: IDENTIFYING MODES OF FAILURE

- Identify modes of failure from features in failure contribution density
- Use to determine process limits, reliability budgets, or opportunities for optimization.
- Can also be used with HAST for component lifetime quantification.

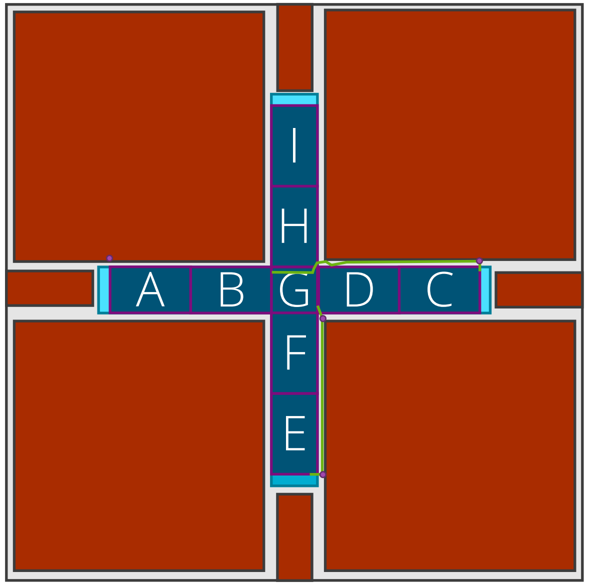




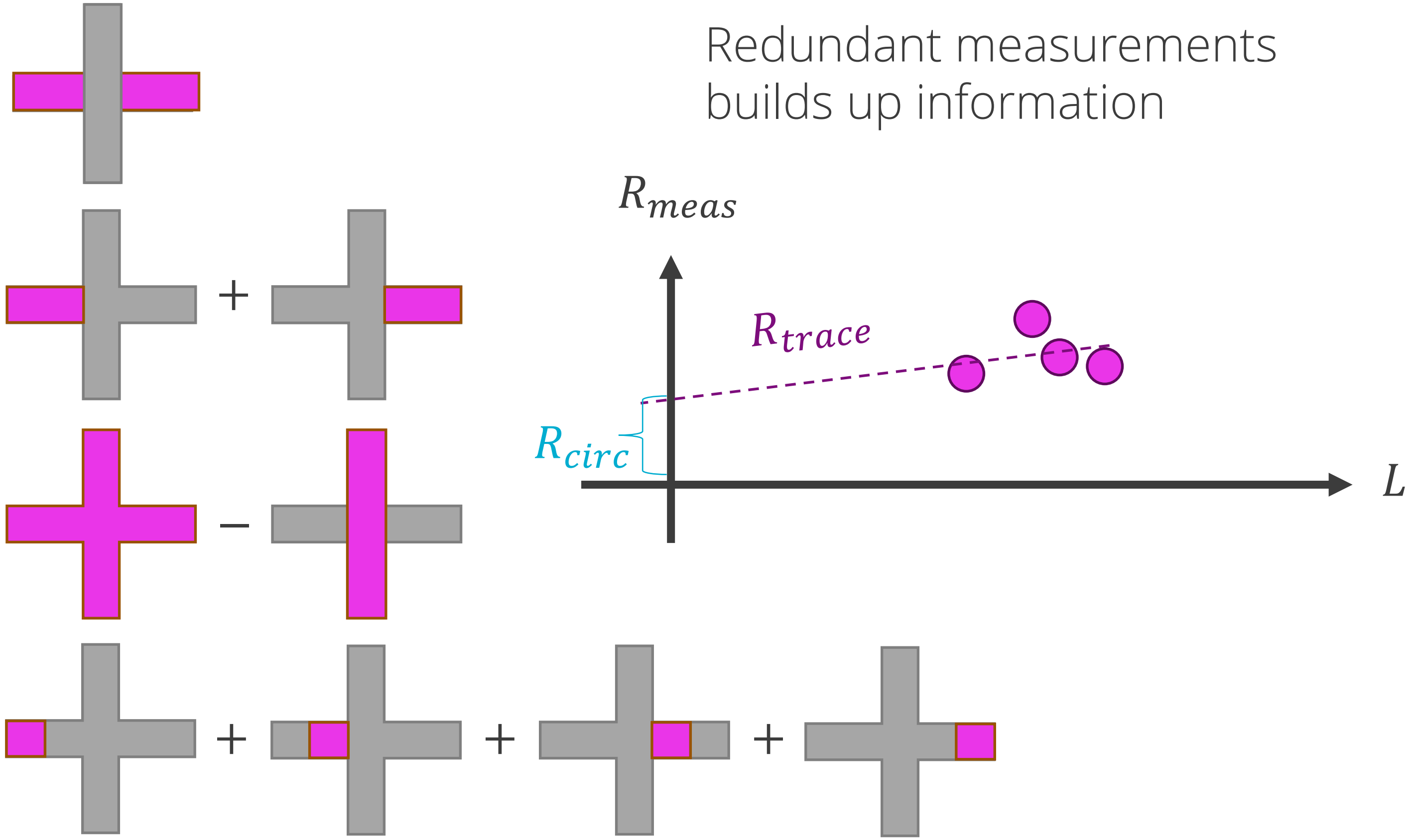
PHYSICAL CALIBRATION: R_0, R_{trace}

Redundant measurements builds up information

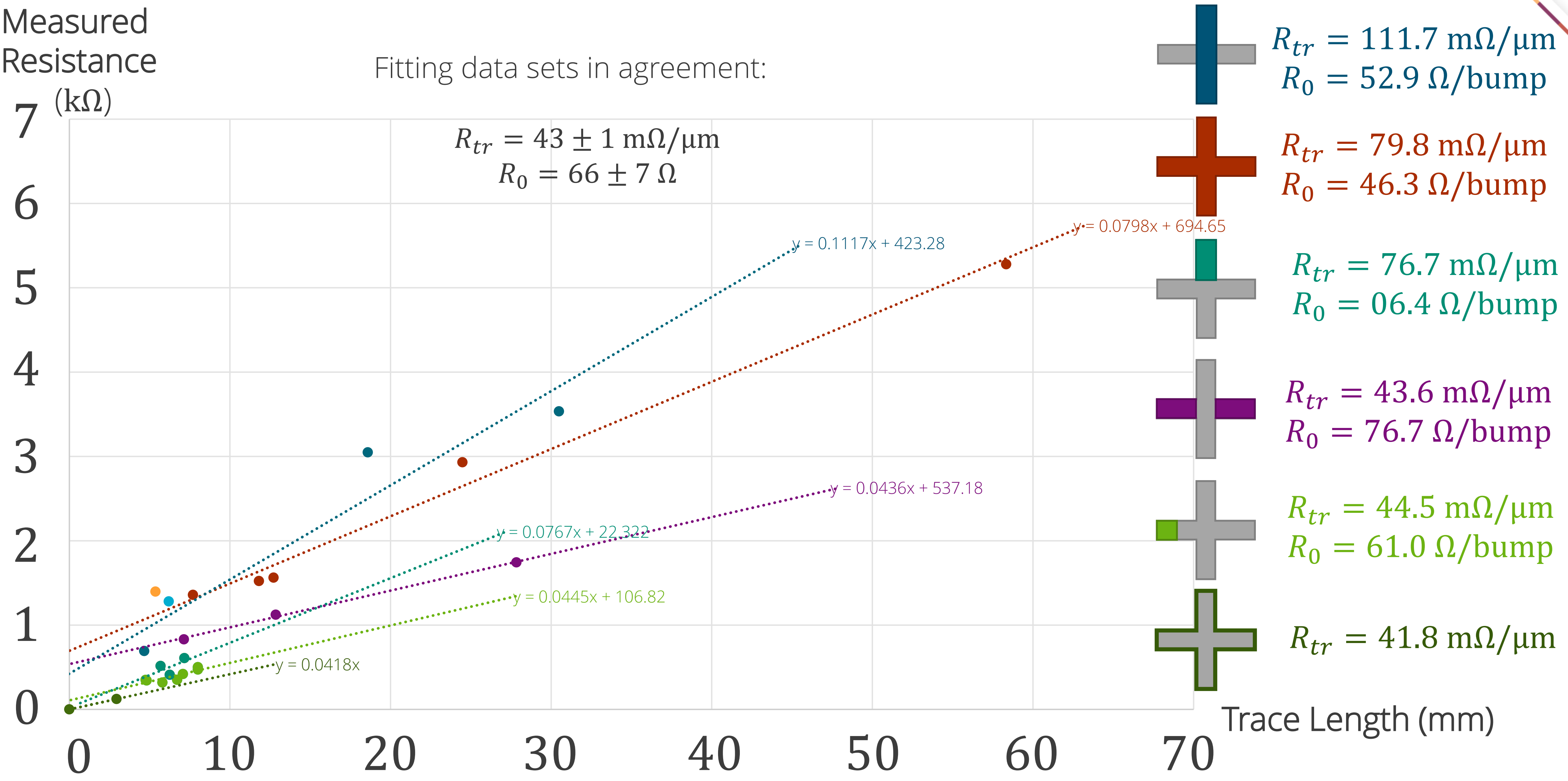
R_0, R_{trace} may not be known, so calibration is needed to find N_F



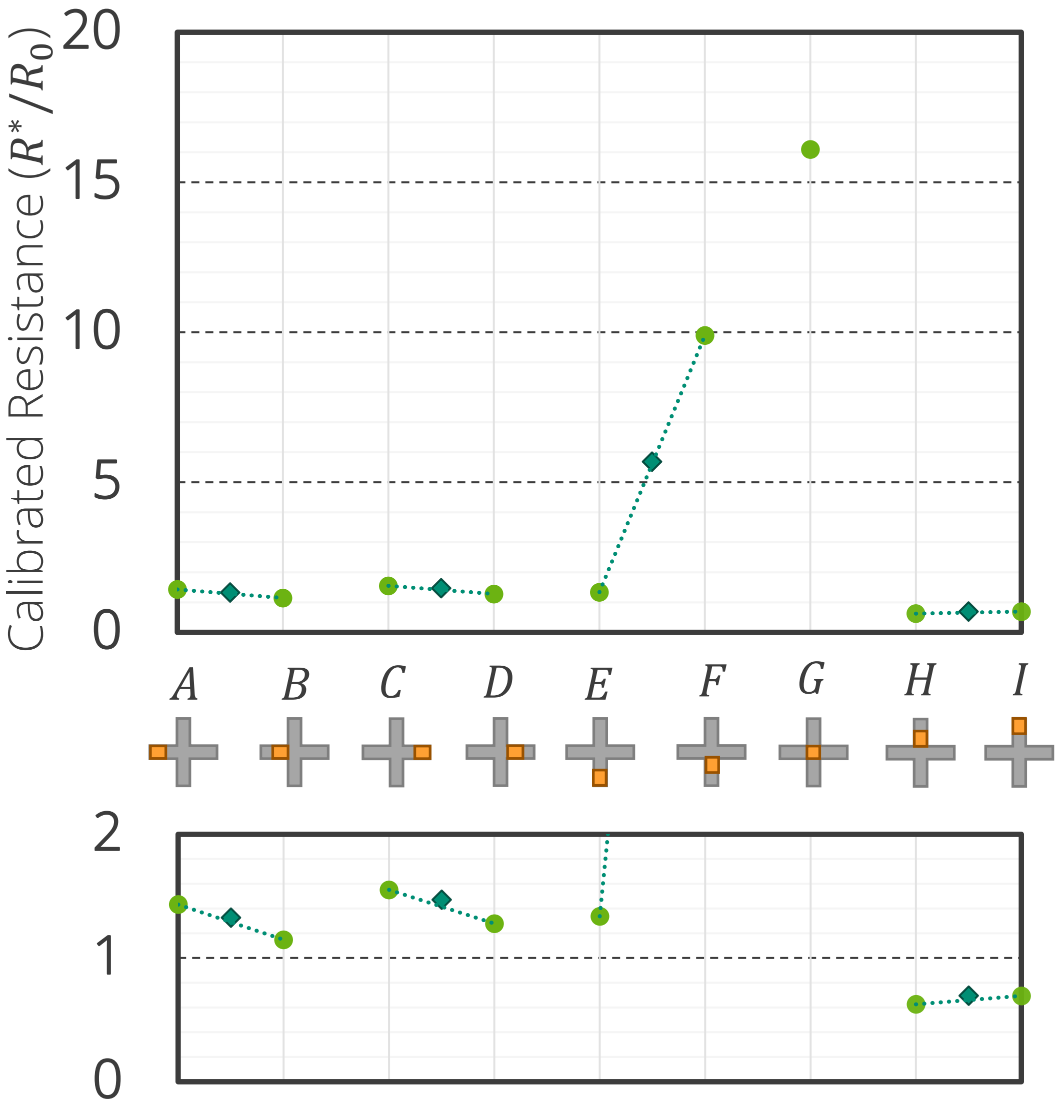
Connected Like:



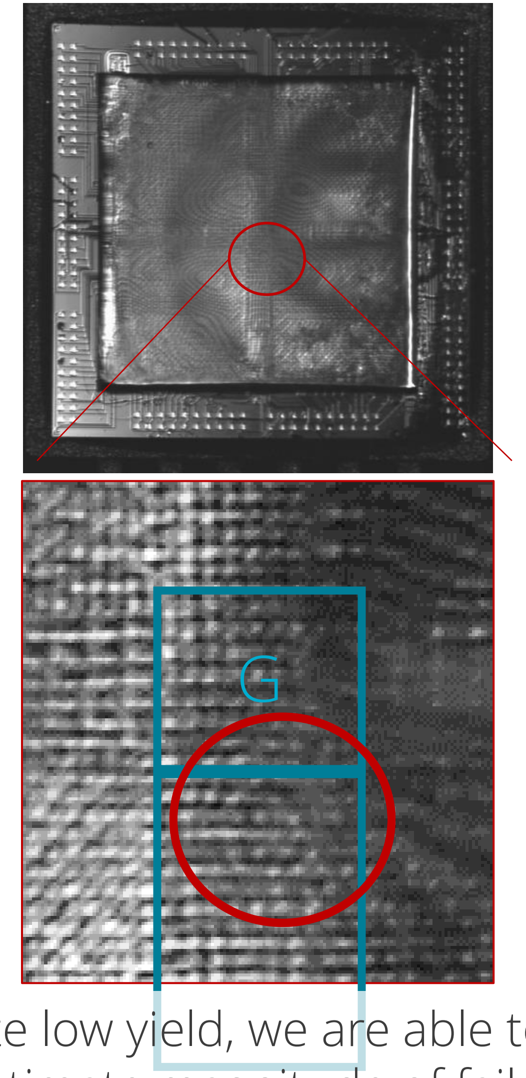
COMPARING METRICS FOR CALIBRATION



COMPARING CALIBRATED RESULTS



Defect Present near the boundary of subsections "F" and "G"

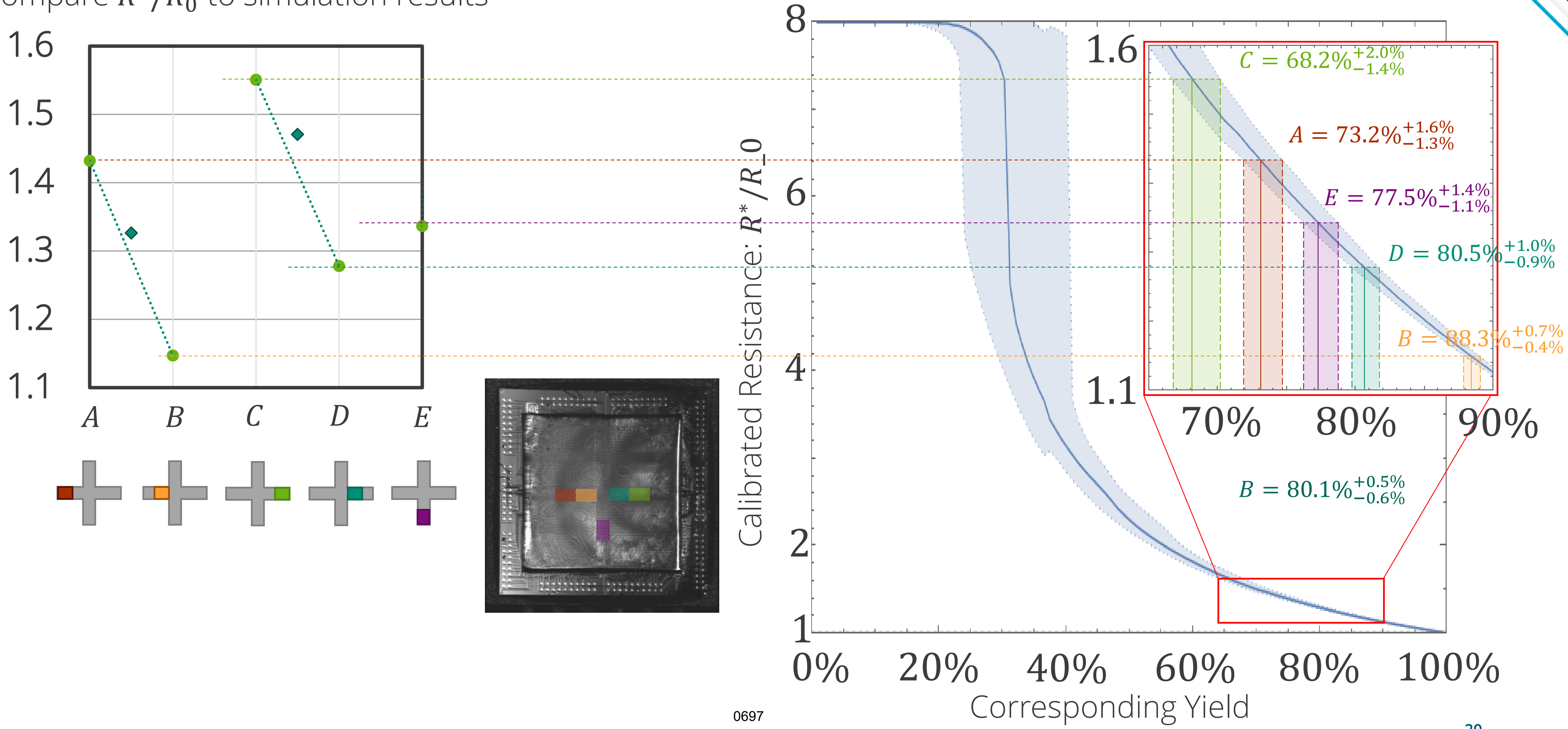


Despite low yield, we are able to detect and estimate magnitude of failures

DETERMINING YIELDS

Compare R^*/R_0 to simulation results

Simulated R^*/R_0 -- Yield Relation



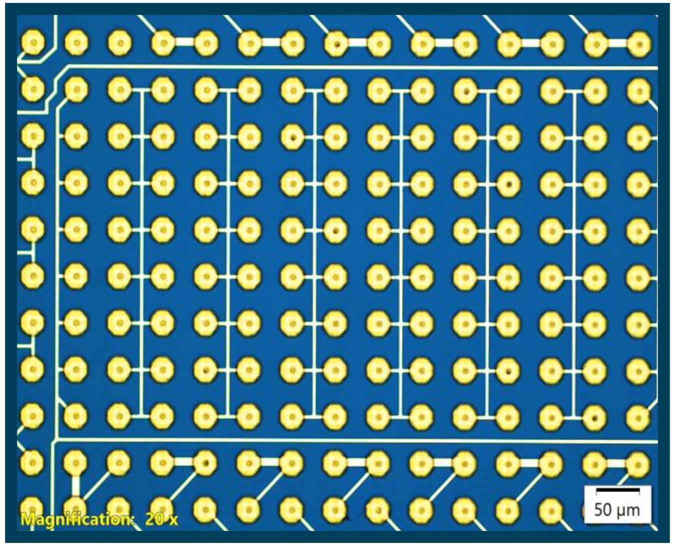
IN CONCLUSION:

Compared to DCs, 2D Daisy 'Fields' are:

- Faster
  - (single measurement)
- Continuous Data
- More Precise
  - (> x10 improvement)
- More Accurate
  - Use earlier in development
  - Over a wider range of processes
- Lower measurement Complexity

Additionally,

- Easy to implement
- Enables new *in situ* yield analysis capability



THANK YOU TO:

Elijah Smythurst

Matthew B. Jordan

Jeremy A. Walraven

Funded or Sponsored By:

This work was supported by the Defense Advanced Research Projects Agency under contract number DENA0003525

Sandia National Laboratories