

Mitigating Parameter Variation with Dynamic Fine-Grain Body Biasing

**Radu Teodorescu, Jun Nakano, Abhishek Tiwari
and Josep Torrellas**

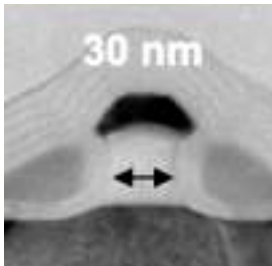
University of Illinois at Urbana-Champaign
<http://iacoma.cs.uiuc.edu>



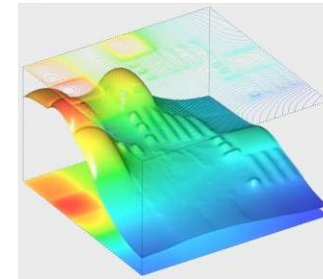


Parameter variation: roadblock to scaling

Process Variation



Temperature Variation

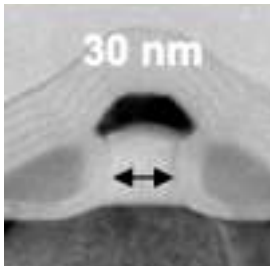


Supply Voltage Variation

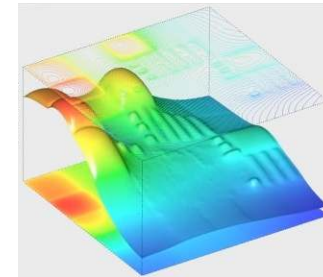


Parameter variation: roadblock to scaling

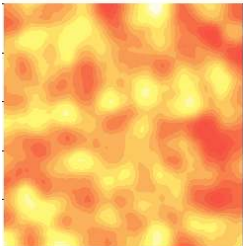
Process Variation



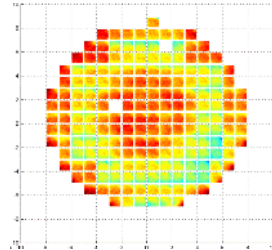
Temperature Variation



Within die (WID)

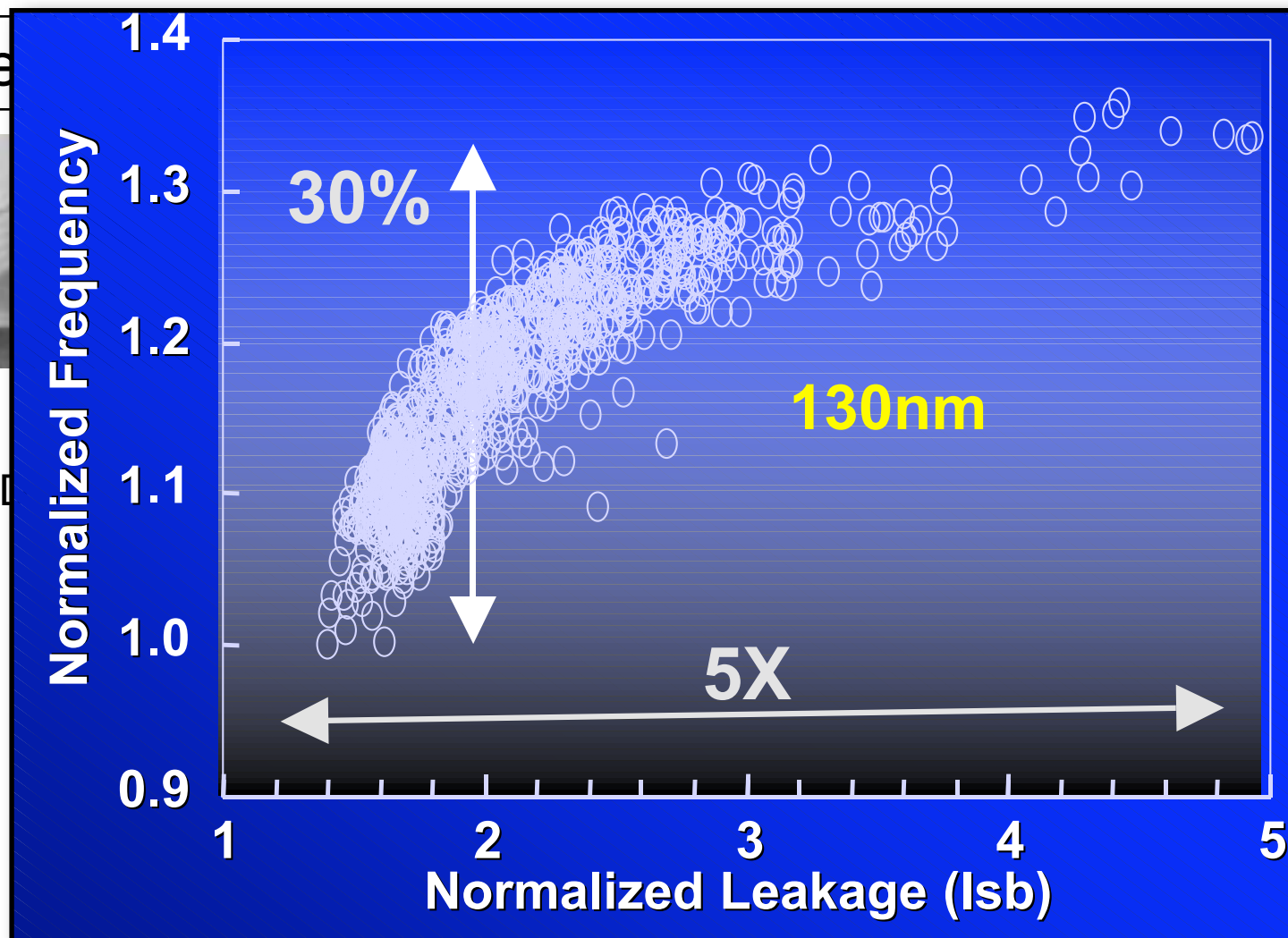


Die-to-die (D2D)





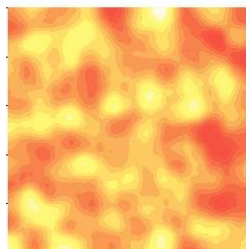
Parameter variation: roadblock to scaling



Process

variation

Within die (WID)

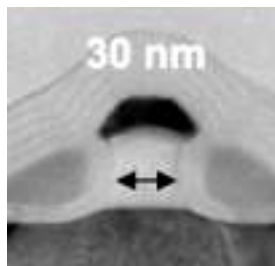


[Shekhar Borkar, Intel Corp.]

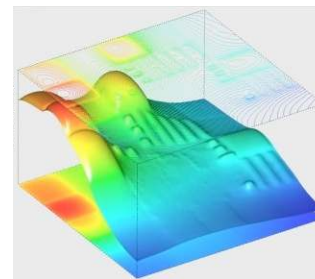


Technology scaling faces a major roadblock

Process Variation



Temperature Variation



Threshold Voltage (V_{th})

Chip frequency



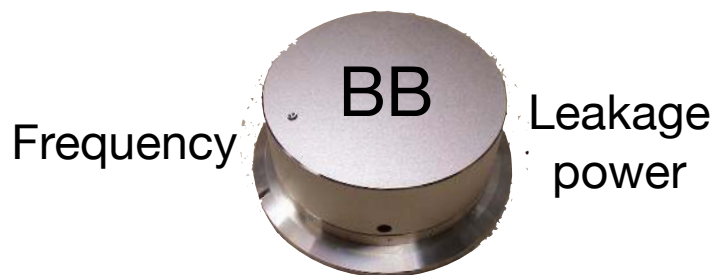
Chip leakage power





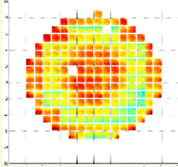
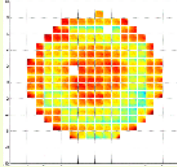
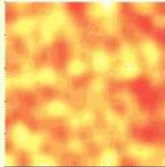
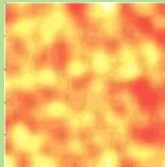
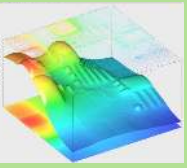
Body biasing

- Well known technique for V_{th} control
- A voltage is applied between source/drain and substrate of a transistor
- Forward body bias (FBB) V_{th} ↓ Freq ↑ Leak ↑
- Reverse body bias (RBB) V_{th} ↑ Freq ↓ Leak ↓
- Key knob to trade off frequency for leakage



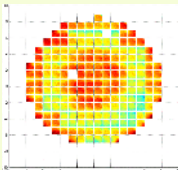
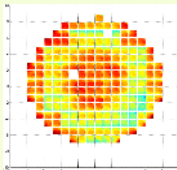
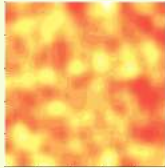
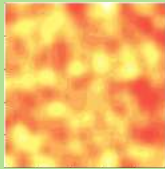
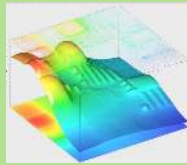


Body bias design space

Time \ Space	Static	Dynamic
	BB fixed for chip lifetime	BB changes with T and workload
Chip-wide	 D2D V_{th} Variation [Intel Xscale] [Intel's 80-core chip]	 D2D V_{th} Variation T Variation
Fine-grain	 WID V_{th} Variation [Tschanz et al]	 WID V_{th} Variation T Variation (space and time) 



Body bias design space

Time \ Space	Static	Dynamic
	BB fixed for chip lifetime	BB changes with T and workload
Chip-wide	 D2D V_{th} Variation [Intel Xscale] [Intel's 80-core chip]	 D2D V_{th} Variation T Variation
Fine-grain	 WID V_{th} Variation S-FGBB [Tschanz et al]	 WID V_{th} Variation D-FGBB (space and time) 



Outline

- Background on S-FGBB
- Dynamic fine-grain body biasing (D-FGBB)
- Environments
- Evaluation
- Conclusions



Outline

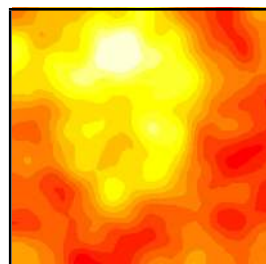
- Background on S-FGBB
- Dynamic fine-grain body biasing (D-FGBB)
- Environments
- Evaluation
- Conclusions



Static fine-grain body biasing

[Tschanz et al, ISSCC 2002]

V_{th} variation



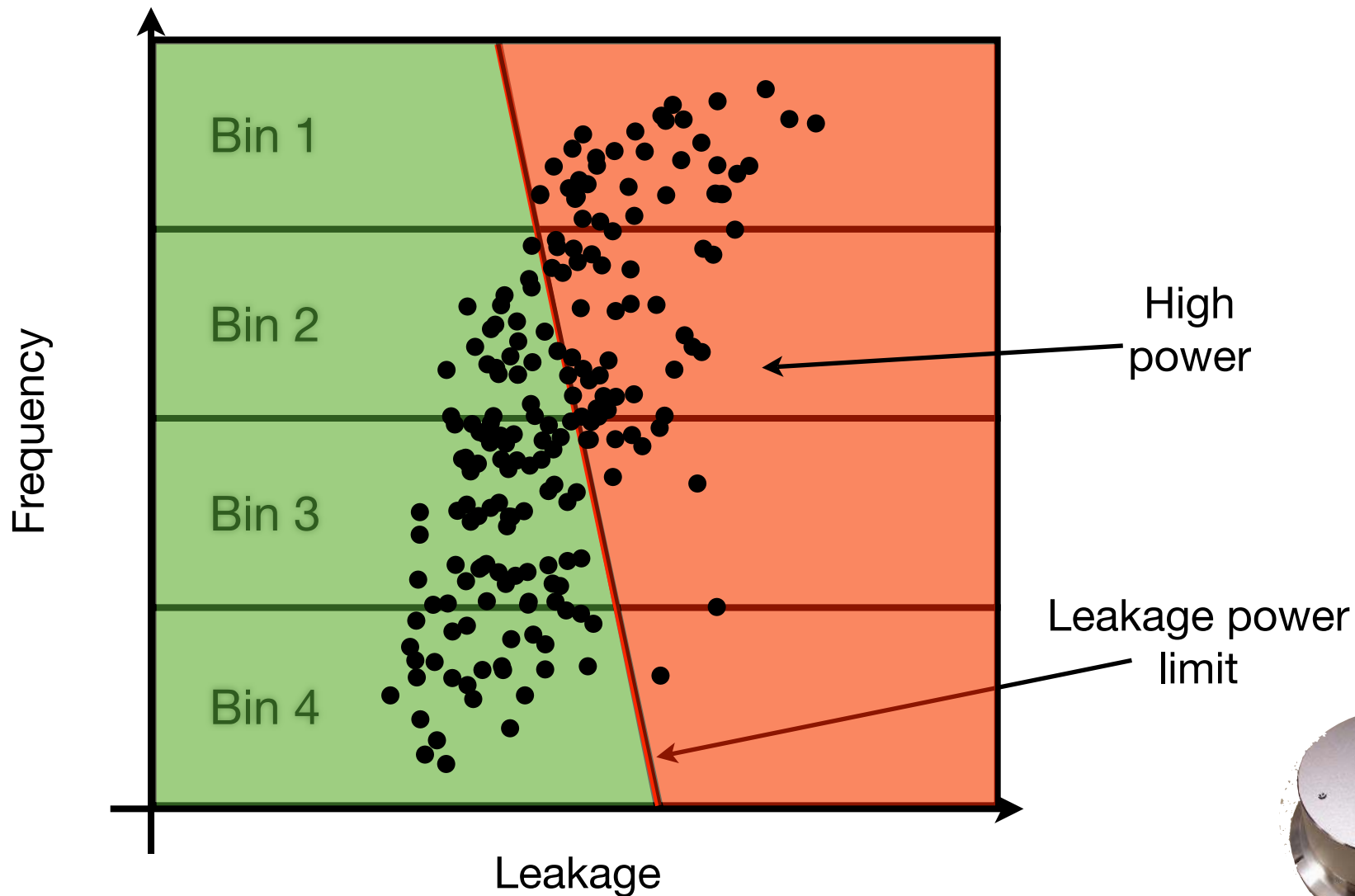
Fine Grain Body Bias

FBB	RBB
RBB	RBB

- The chip is divided in BB cells
 - Slow cells receive FBB - increase speed
 - Leaky cells receive RBB - save leakage
- The result is reduced WID variation (delay, power)
- BB voltages determined at manufacturing
- Fixed for the lifetime of the chip

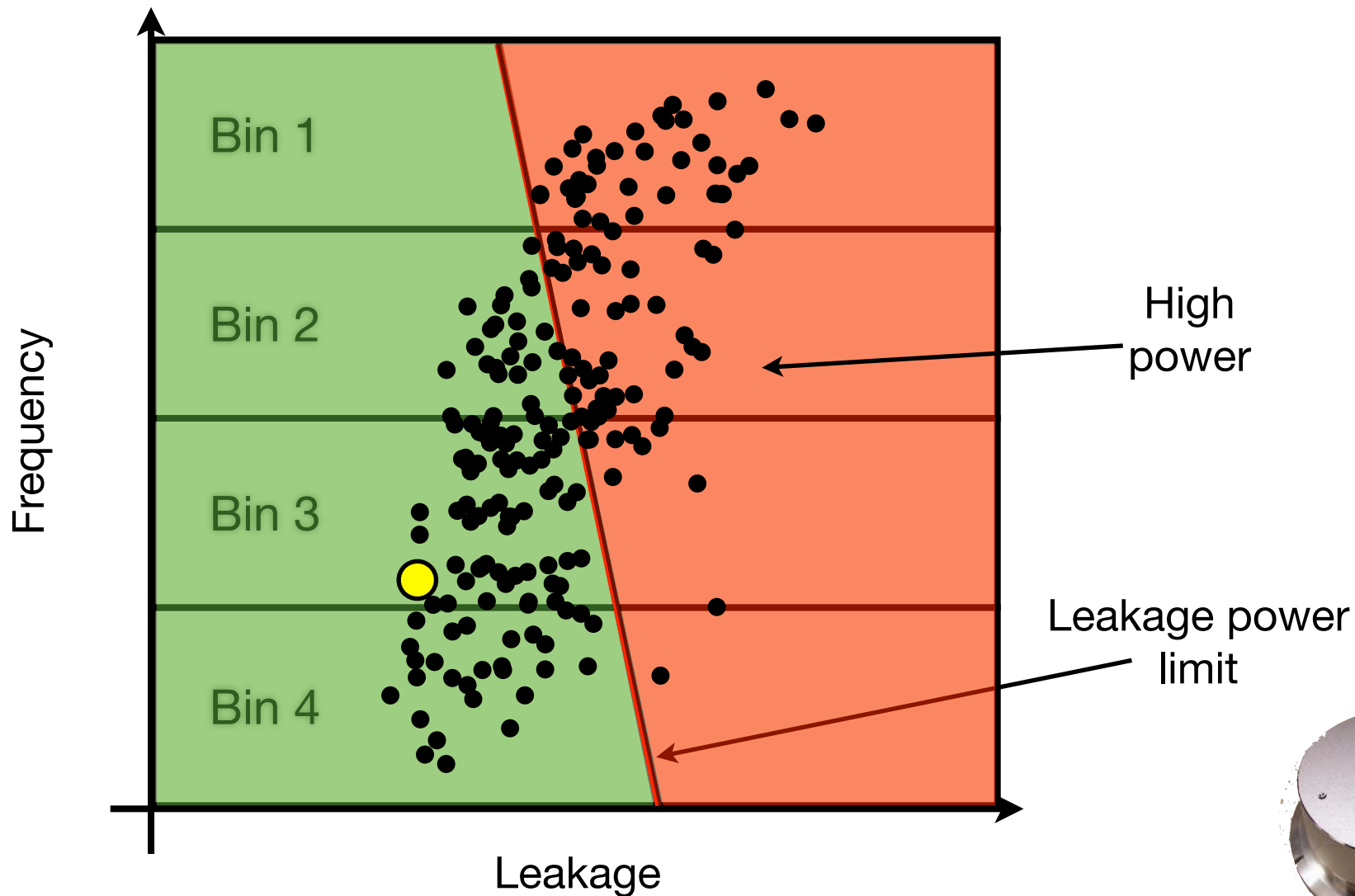


Frequency binning



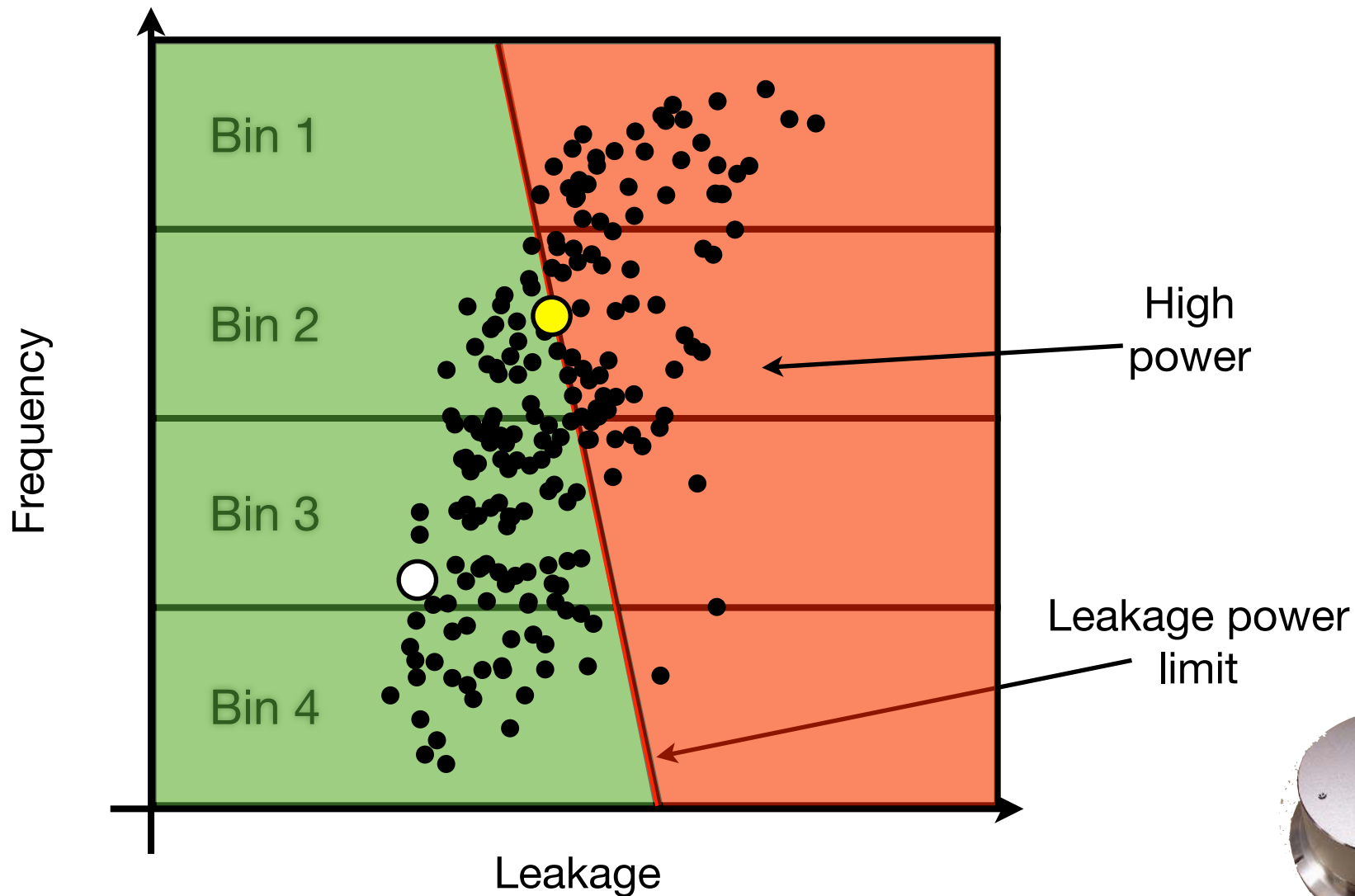


Frequency binning



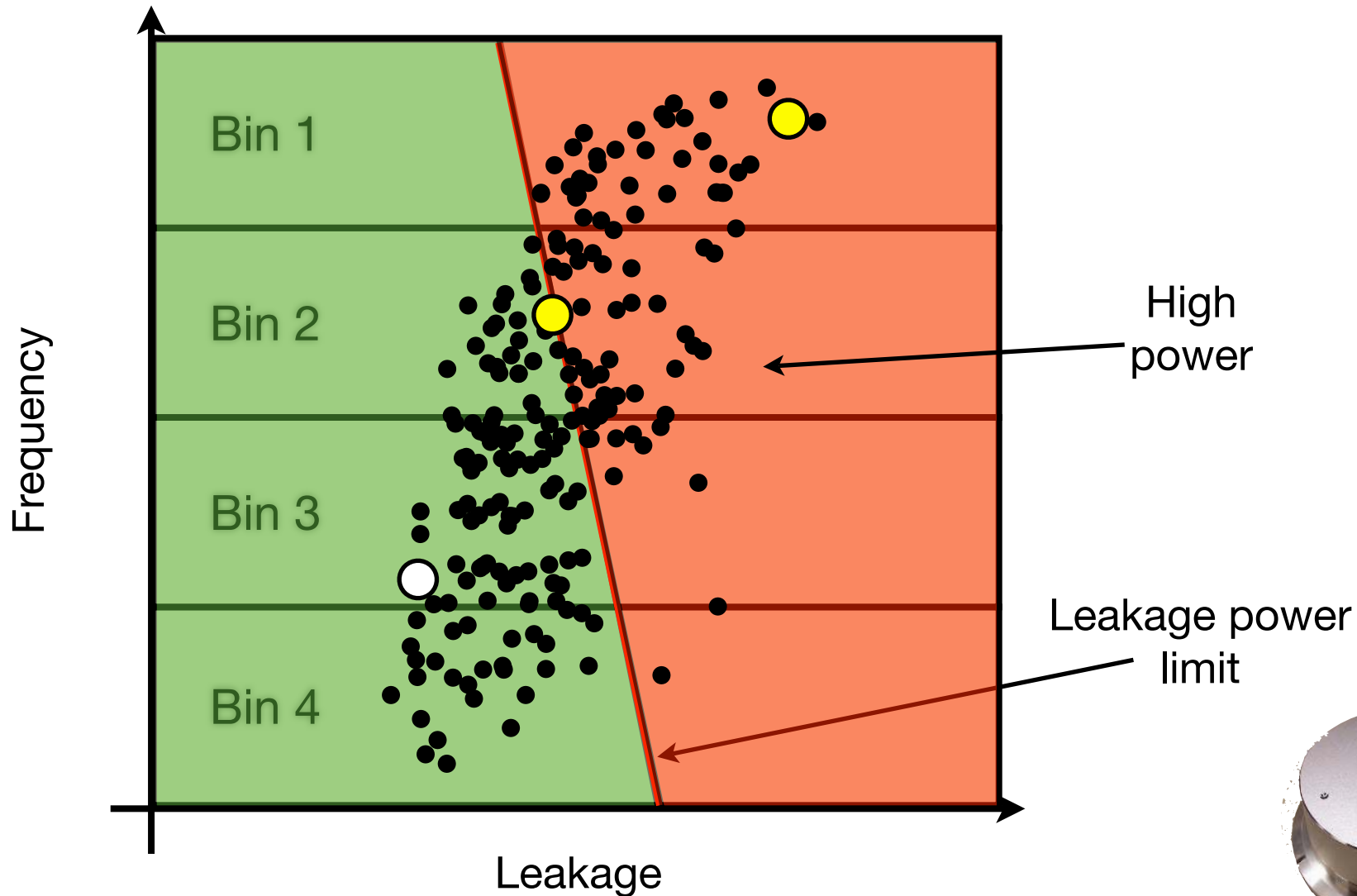


Frequency binning



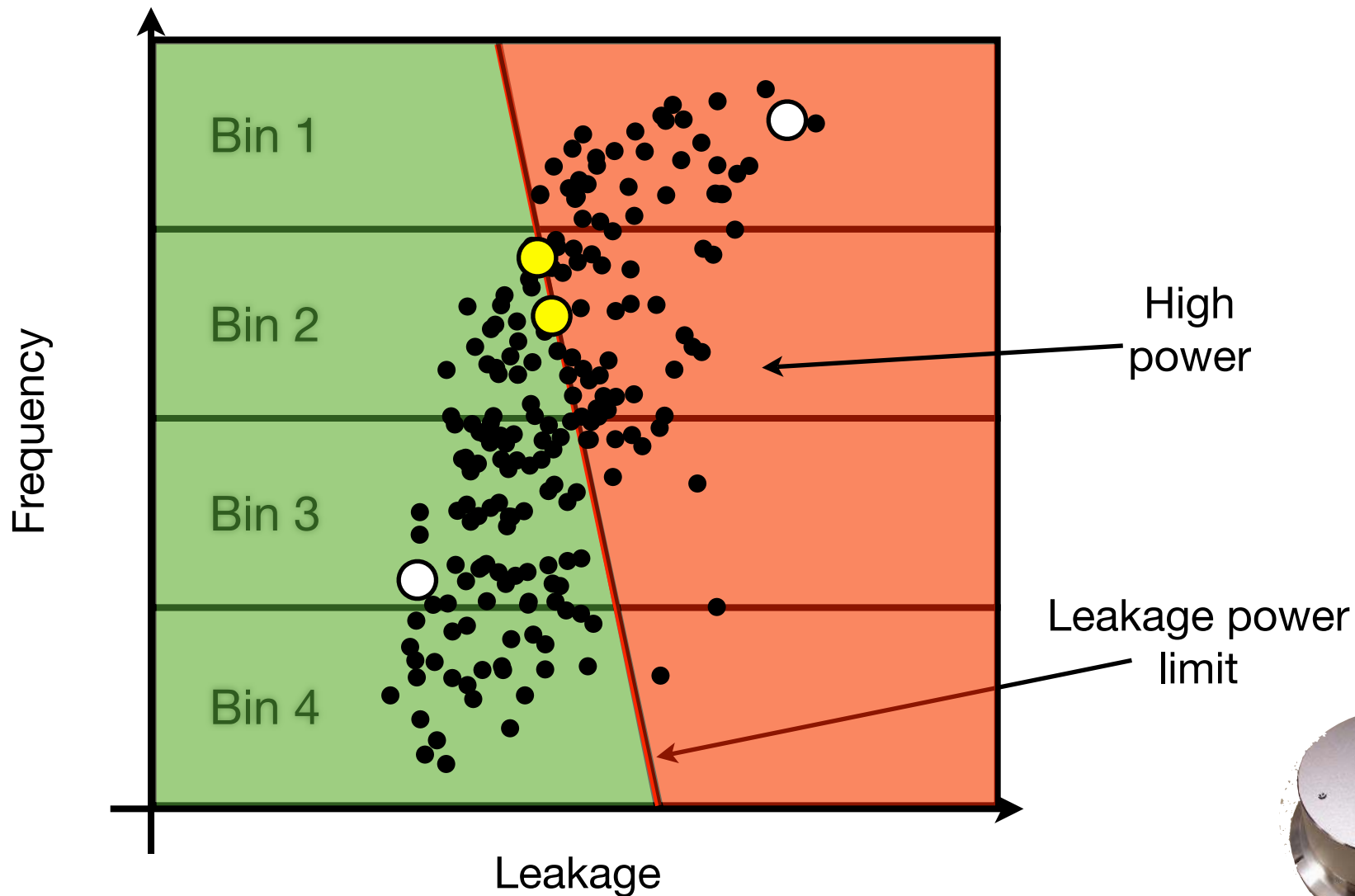


Frequency binning



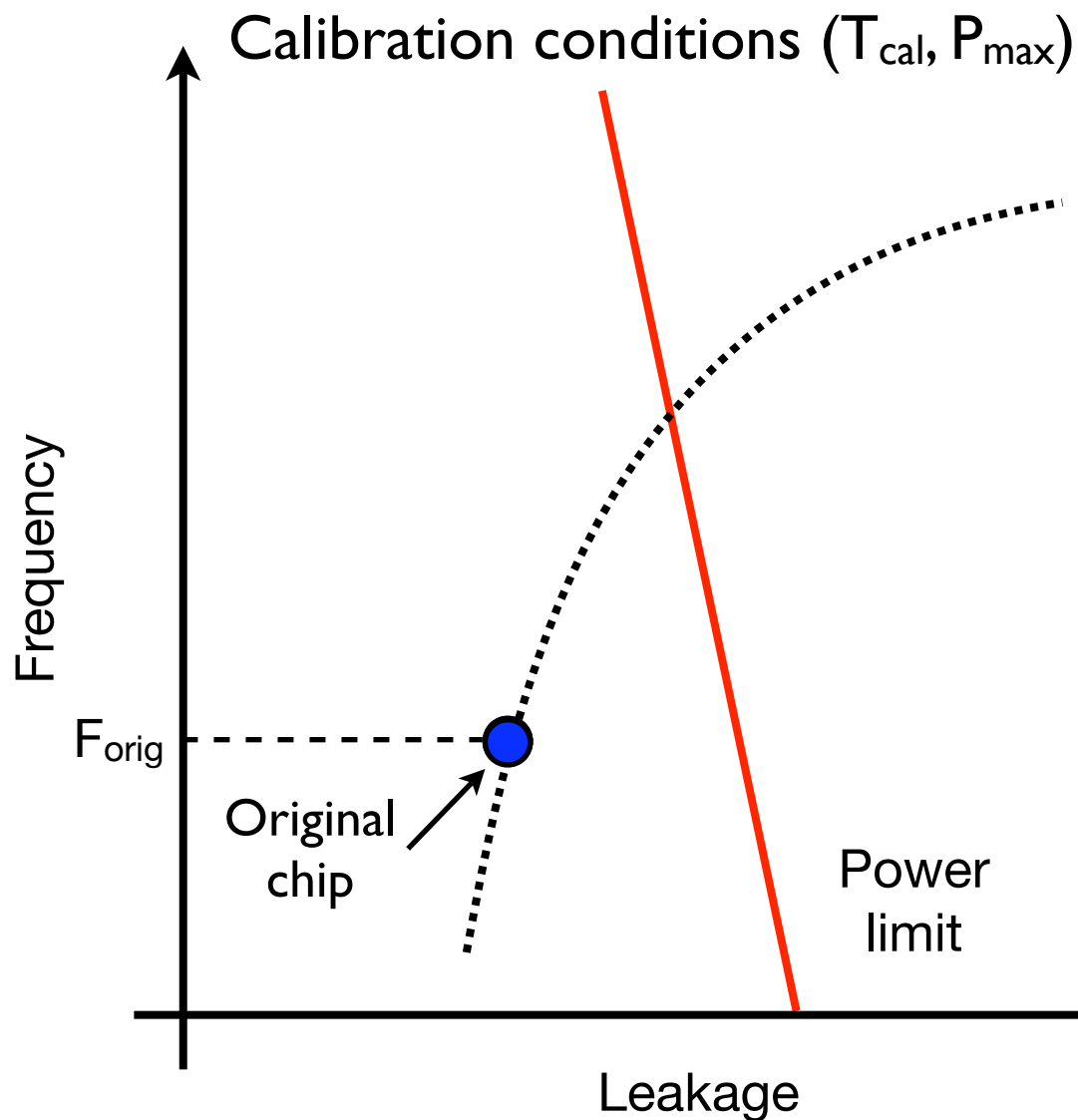


Frequency binning





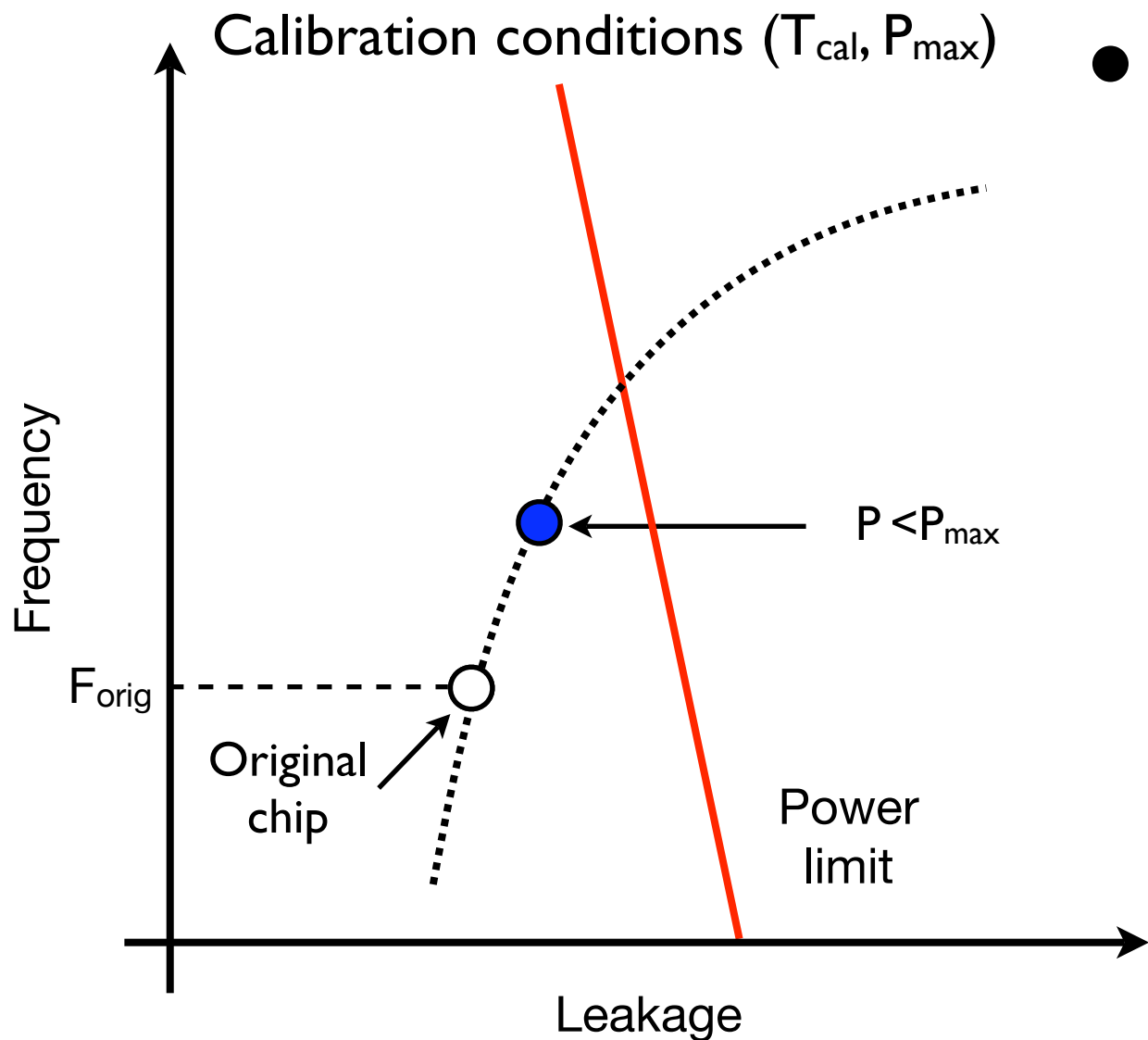
Calibration after manufacturing



- Calibration takes place at maximum temperature T_{cal} (burn-in oven)



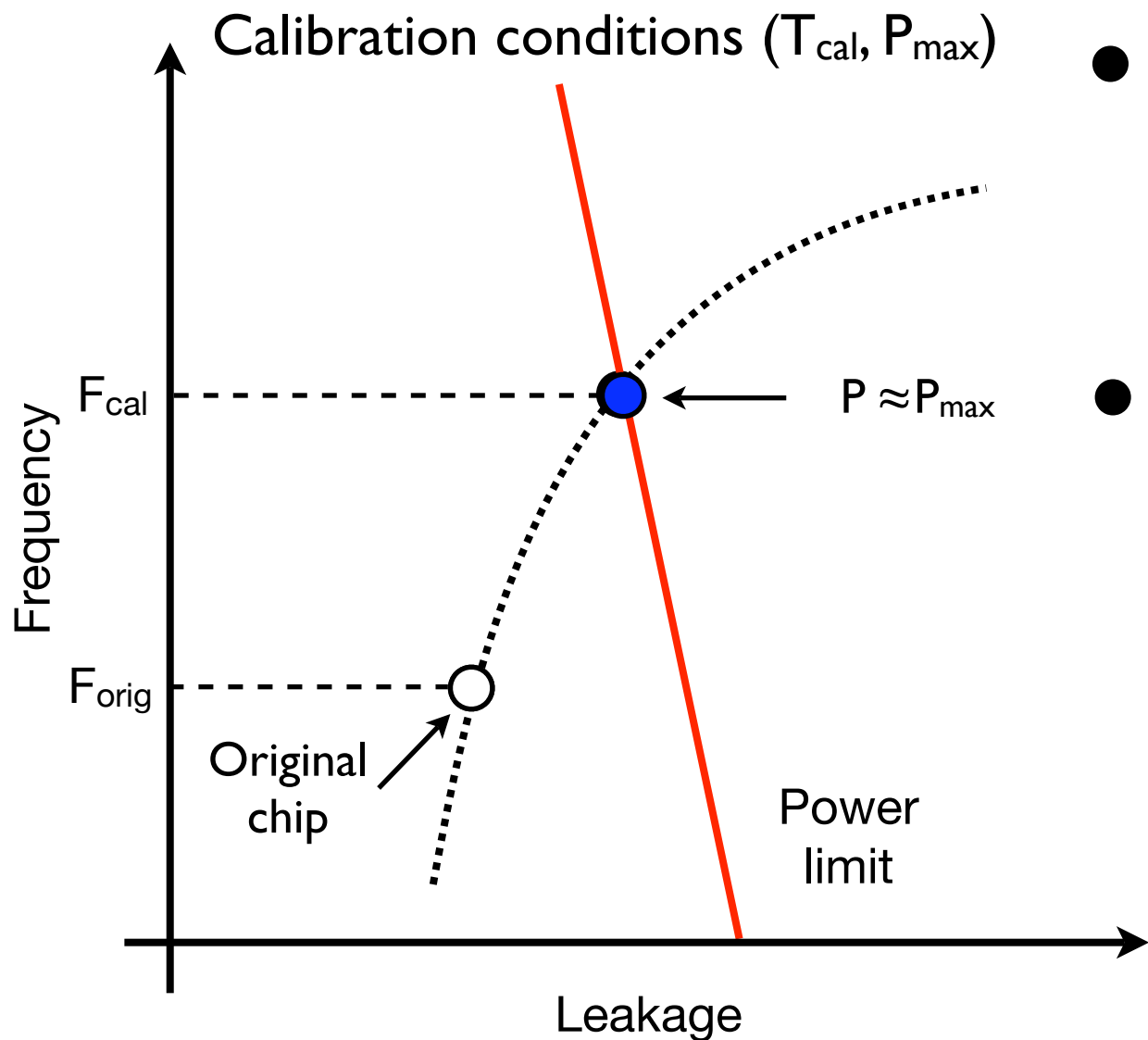
Calibration after manufacturing



- Calibration takes place at maximum temperature T_{cal} (burn-in oven)



Calibration after manufacturing



- Calibration takes place at maximum temperature T_{cal} (burn-in oven)
- F_{cal} becomes the chip's frequency



Outline

- Background on S-FGBB
- Dynamic fine-grain body biasing (D-FGBB)
- Environments
- Evaluation
- Conclusions



Motivation for D-FGBB

- Significant temperature variation:
 - Space: across different functional units, on chip
 - Time: as the activity factor of the workload changes
 - Between average and worst case conditions (T_{cal})
- D-FGBB can exploit this temperature variation
 - Adapt the body bias to changing conditions



Motivation for D-FGGB

- **Optimal** body bias:

The body bias than **minimizes** leakage power at the target frequency

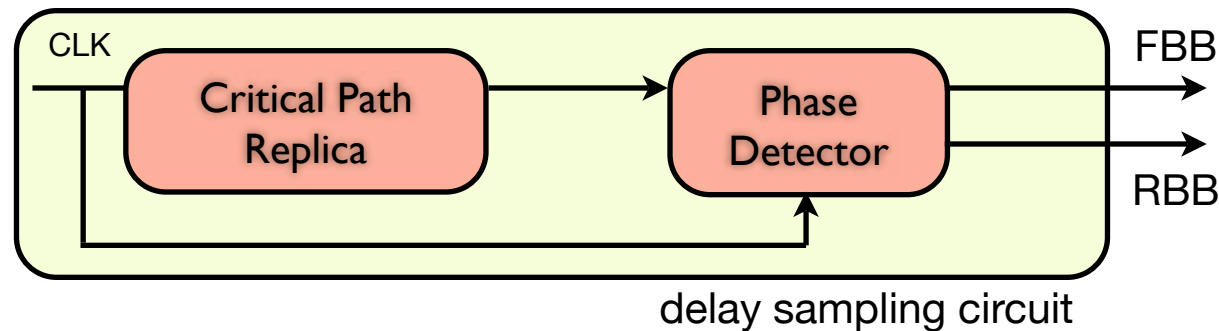
- Circuit delay changes with temperature
- Therefore optimal BB changes with temperature

The goal of D-FGGB is to keep the body bias optimal as T changes



Finding the optimal BB

- Measure the delay of each BB domain (cell)
- Delay sampling circuit:

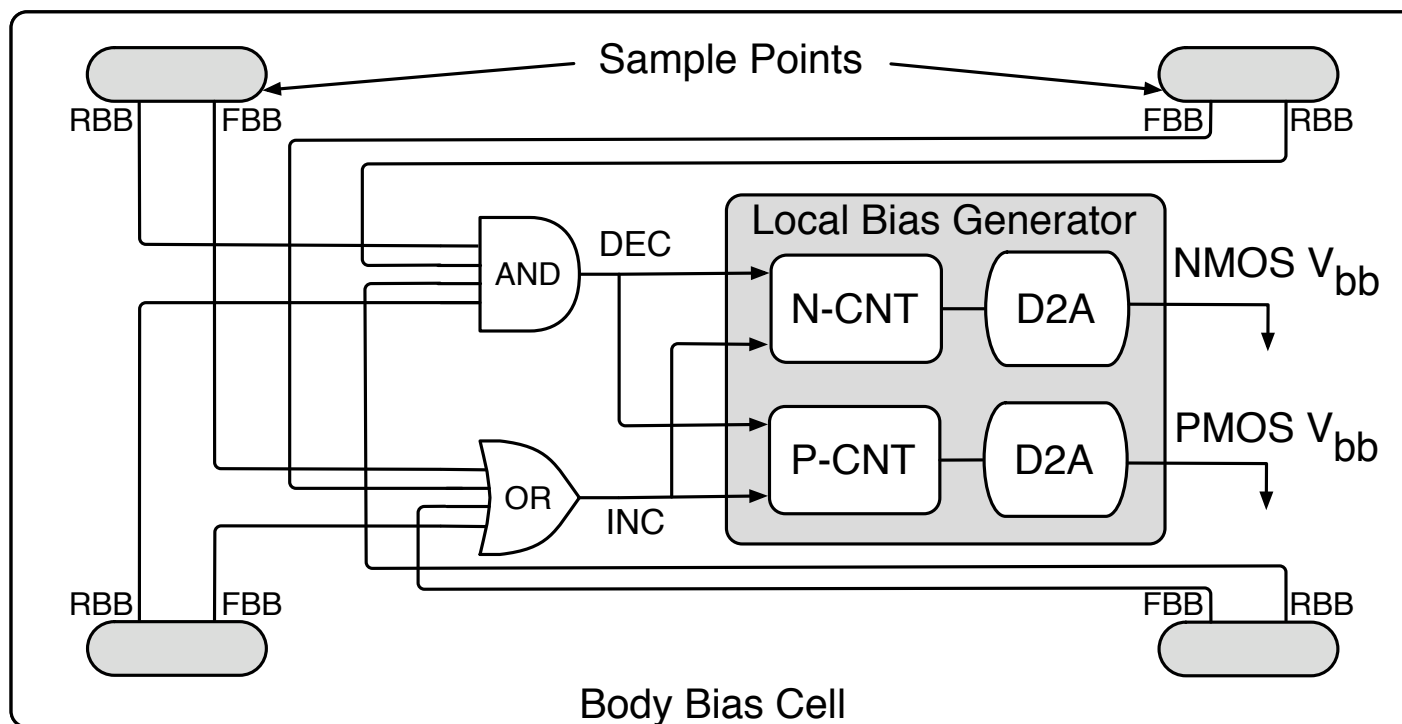


- Phase detector - measures delay of critical path replica
 - If slow - FBB signal raised
 - If fast - RBB signal raised



Applying dynamic fine-grain BB

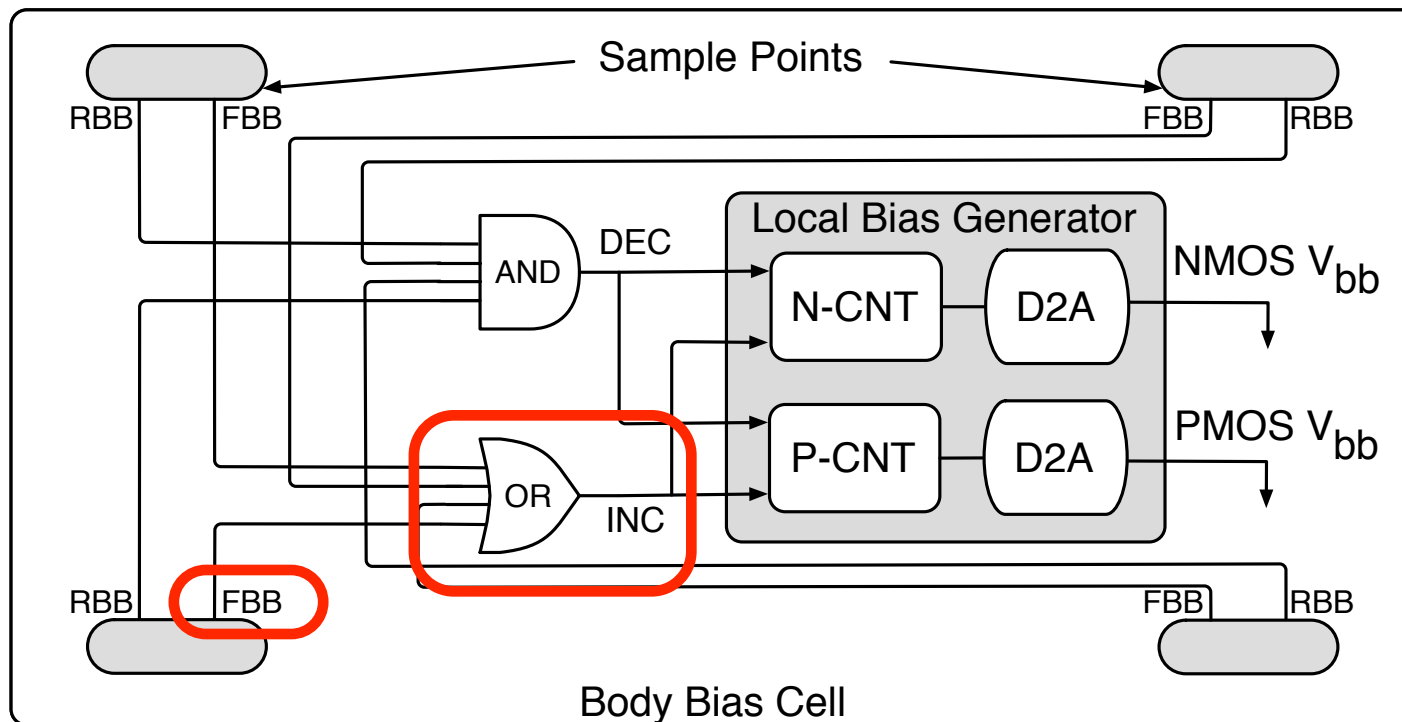
- BB is determined based on feedback from delay samples





Applying dynamic fine-grain BB

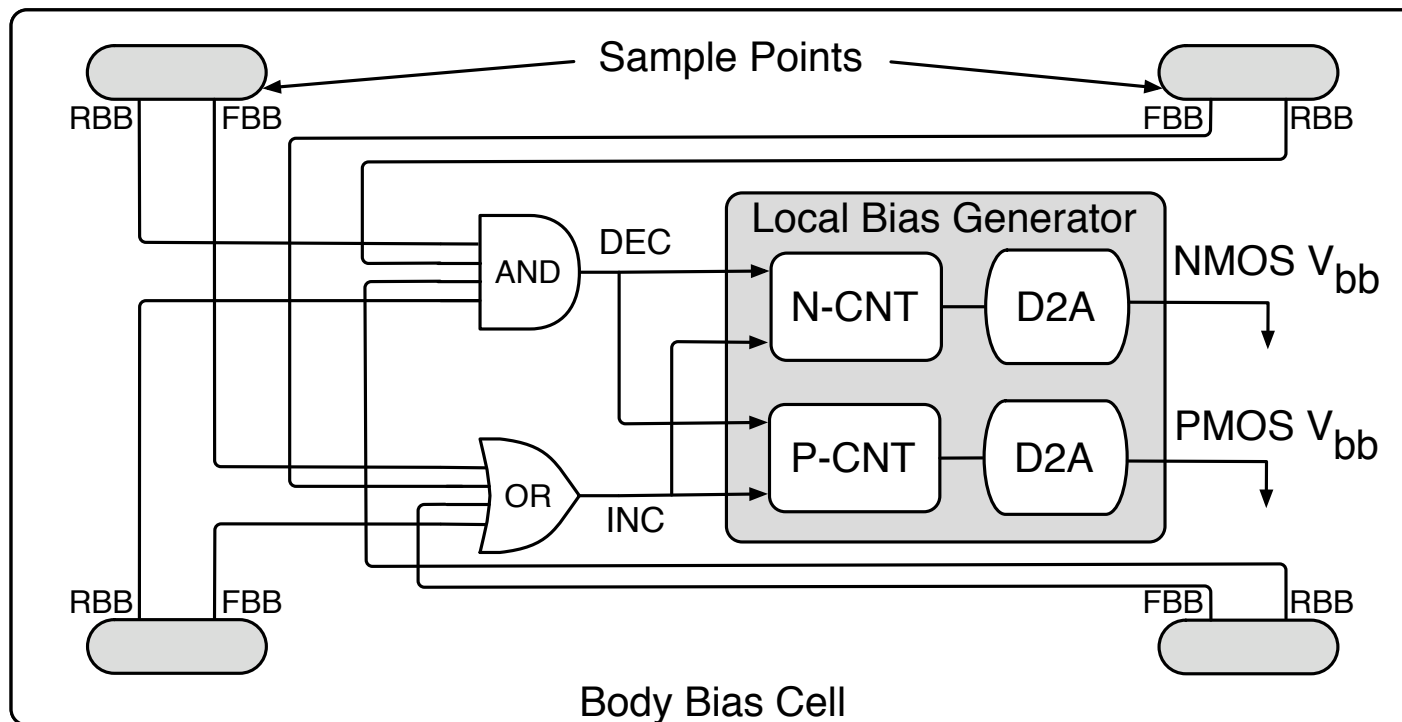
- BB is determined based on feedback from delay samples





Applying dynamic fine-grain BB

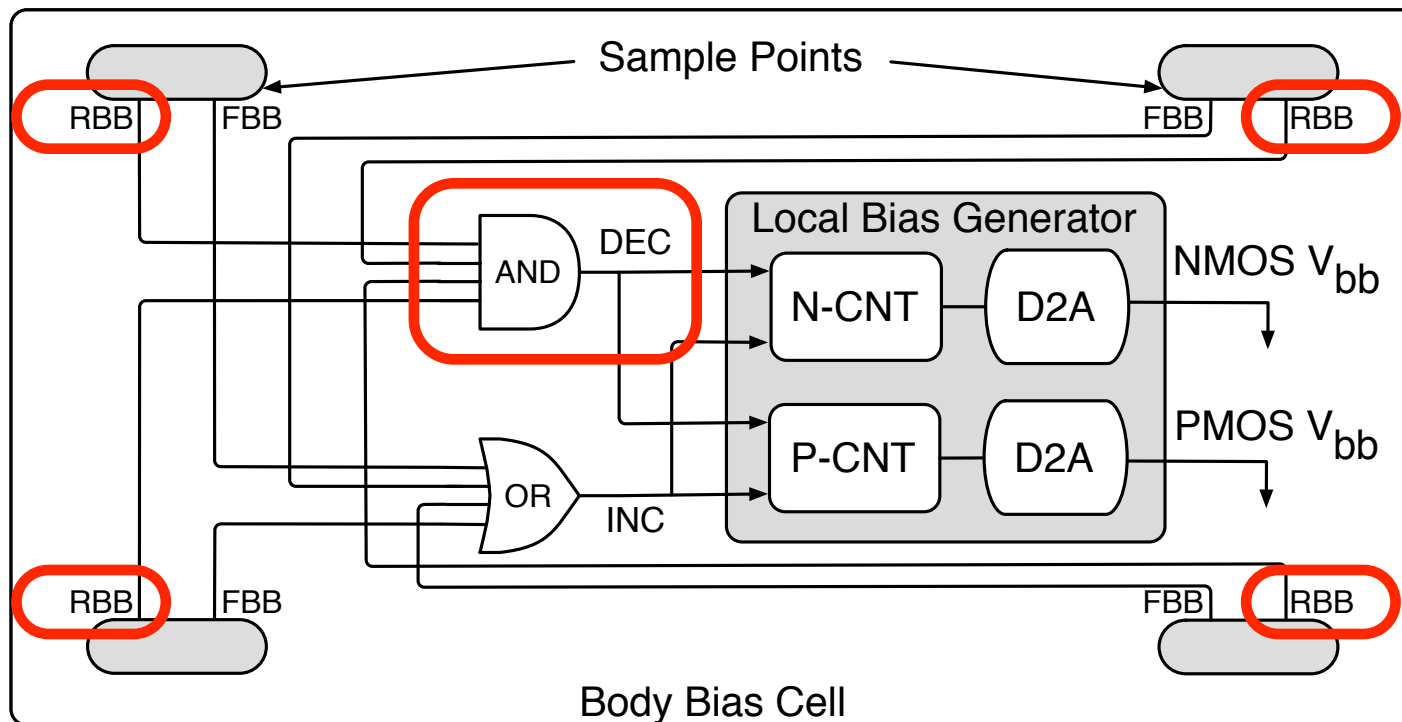
- BB is determined based on feedback from delay samples





Applying dynamic fine-grain BB

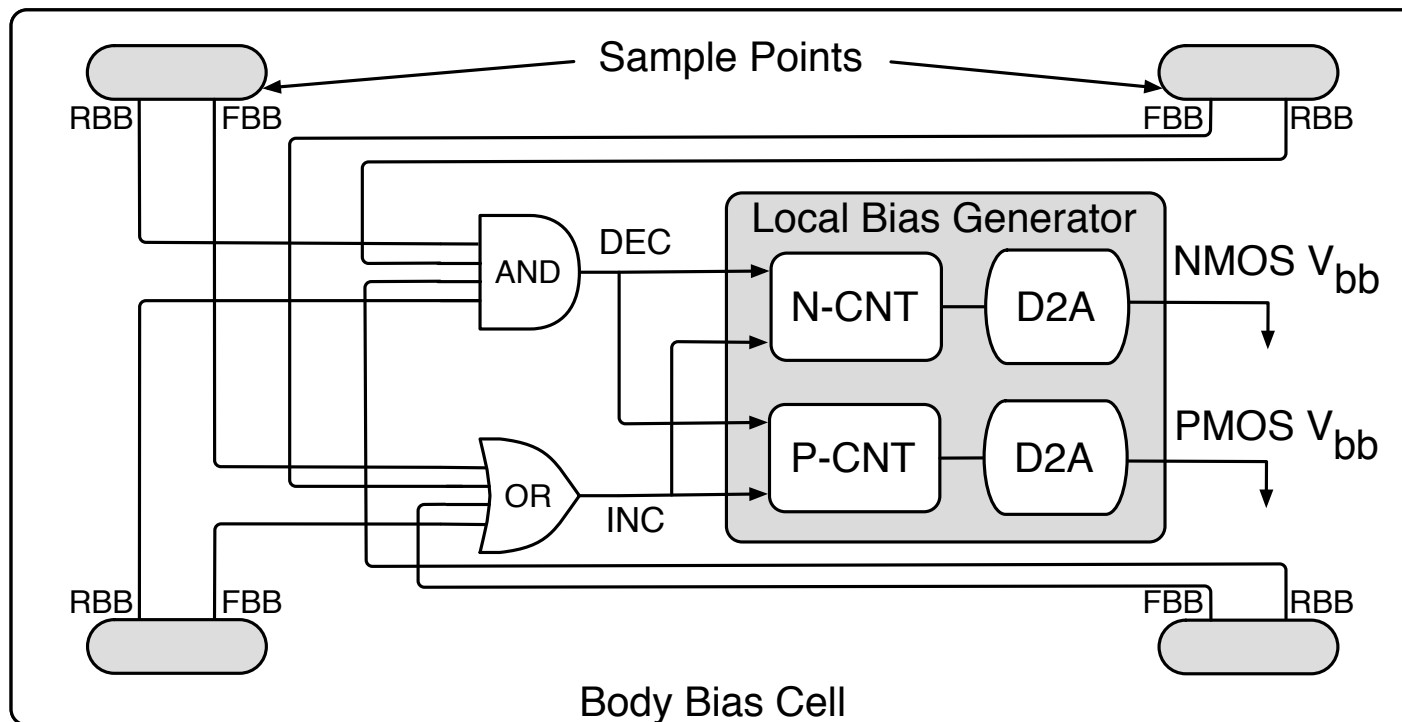
- BB is determined based on feedback from delay samples





Applying dynamic fine-grain BB

- BB is determined based on feedback from delay samples



- The BB changes until optimal delay is reached
- BB stays constant, until T conditions change again






Outline

- Background on S-FGBB
- Dynamic fine-grain body biasing (D-FGBB)
- **Environments**
- Evaluation
- Conclusions






D-FGGB environments

Operating environments	D-FGGB
Standard 	Minimize leakage power at F_{cal}
High performance 	Maximize average frequency
Low Power 	Minimize leakage power at F_{orig}

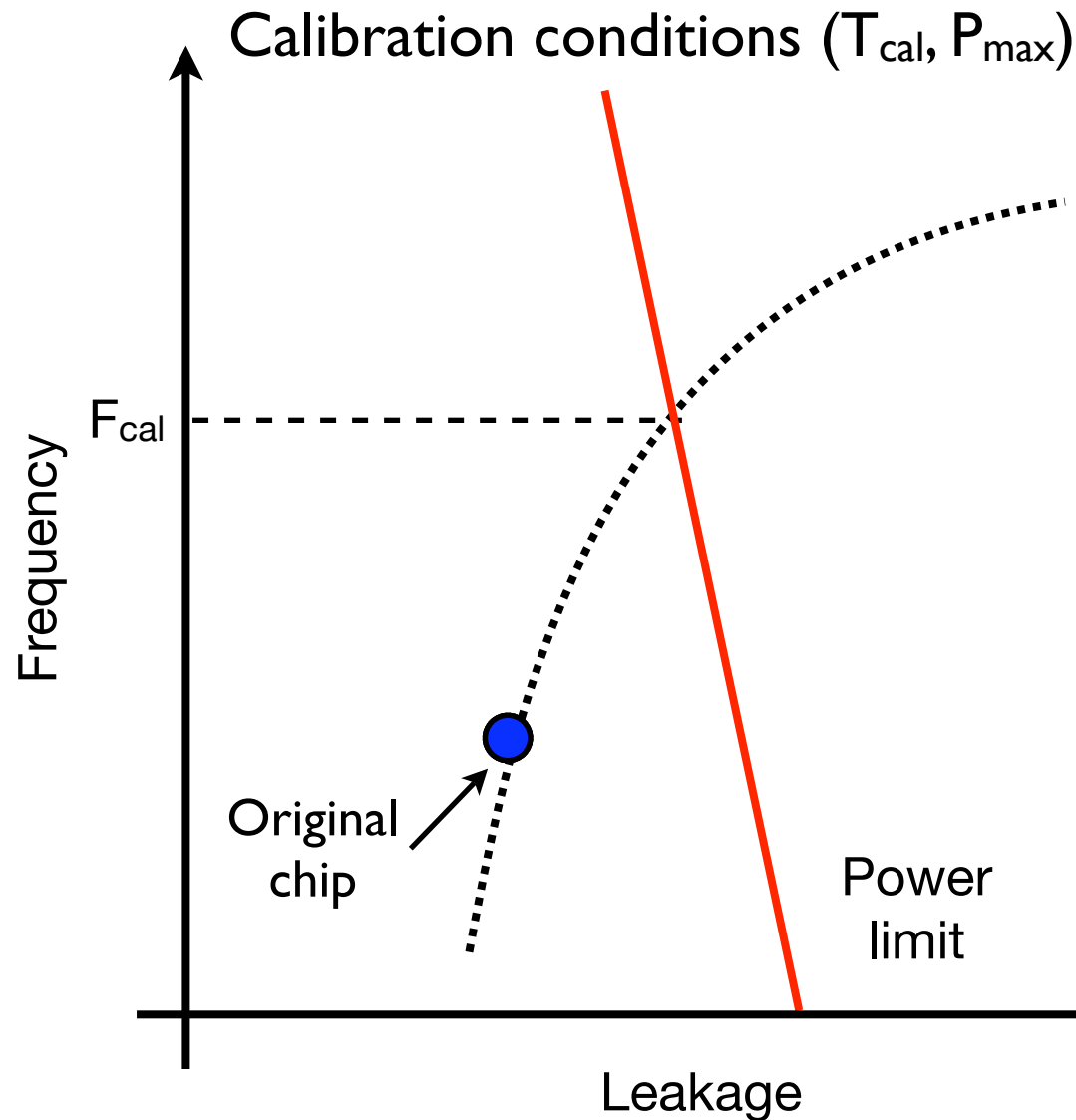


D-FGGB environments

Operating environments	D-FGGB
Standard 	Minimize leakage power at F_{cal}
High performance 	Maximize average frequency
Low Power 	Minimize leakage power at F_{orig}



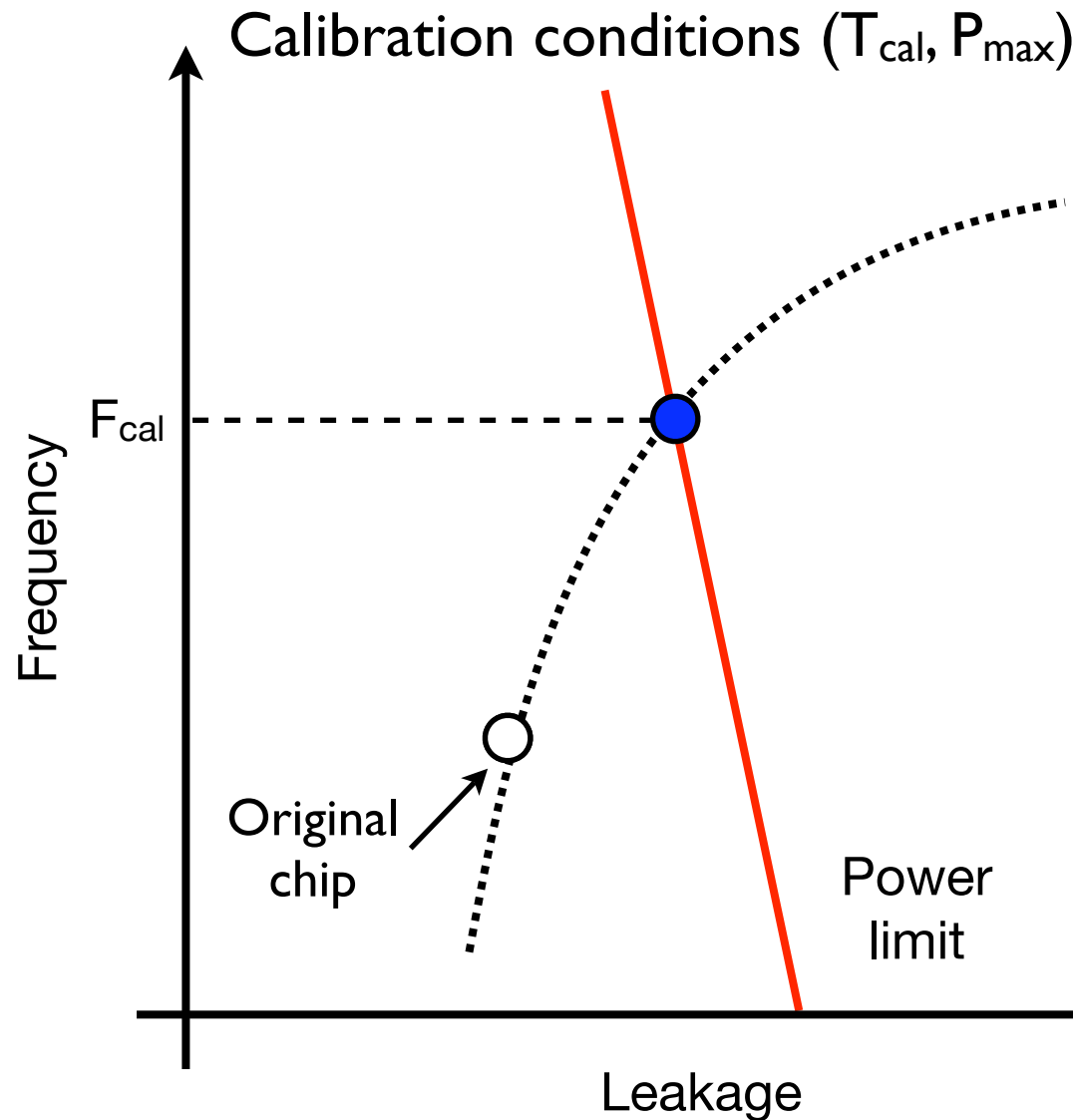
Standard environment



- S-FGGB finds and sets F_{cal}



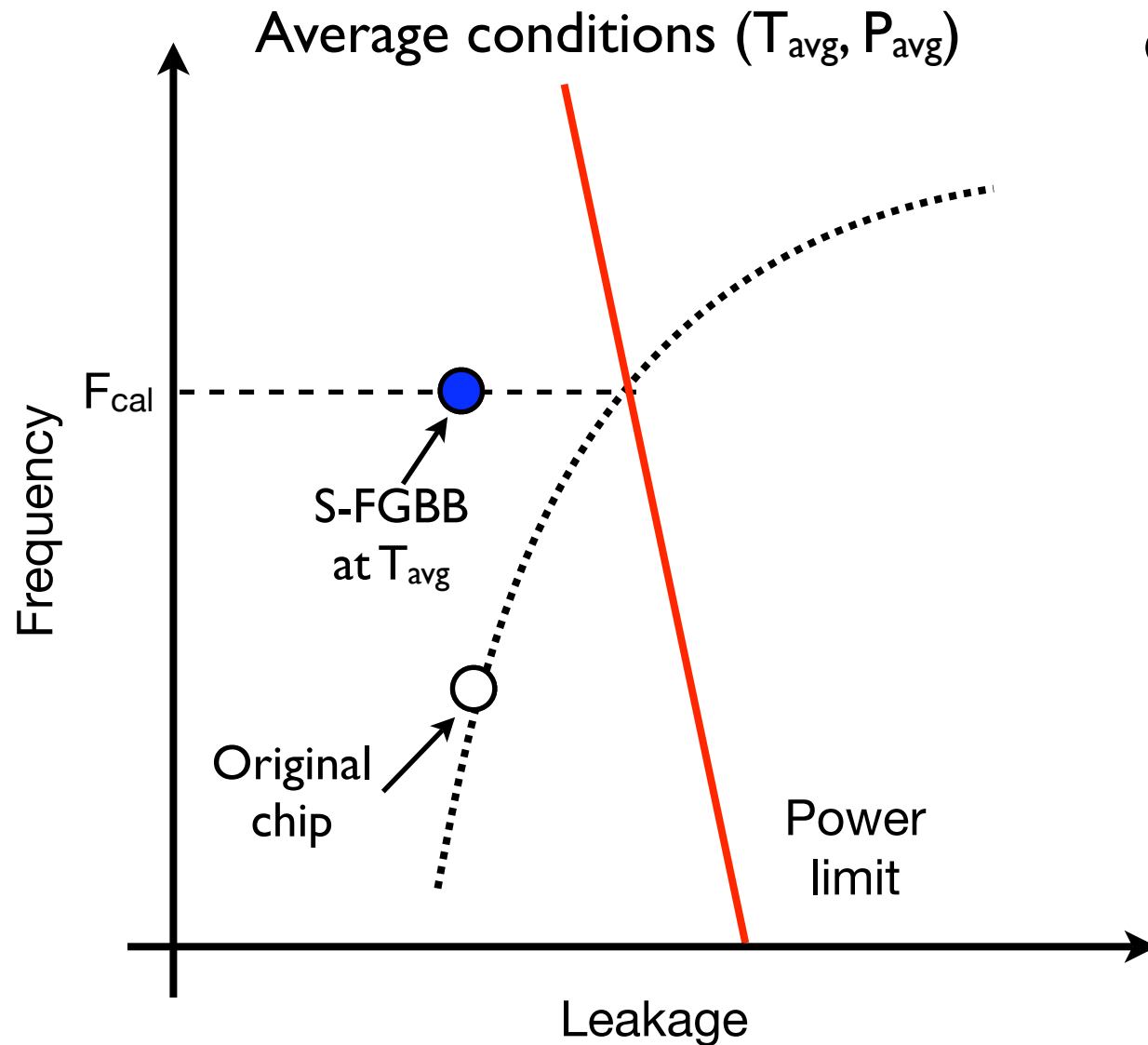
Standard environment



- S-FGGB finds and sets F_{cal}



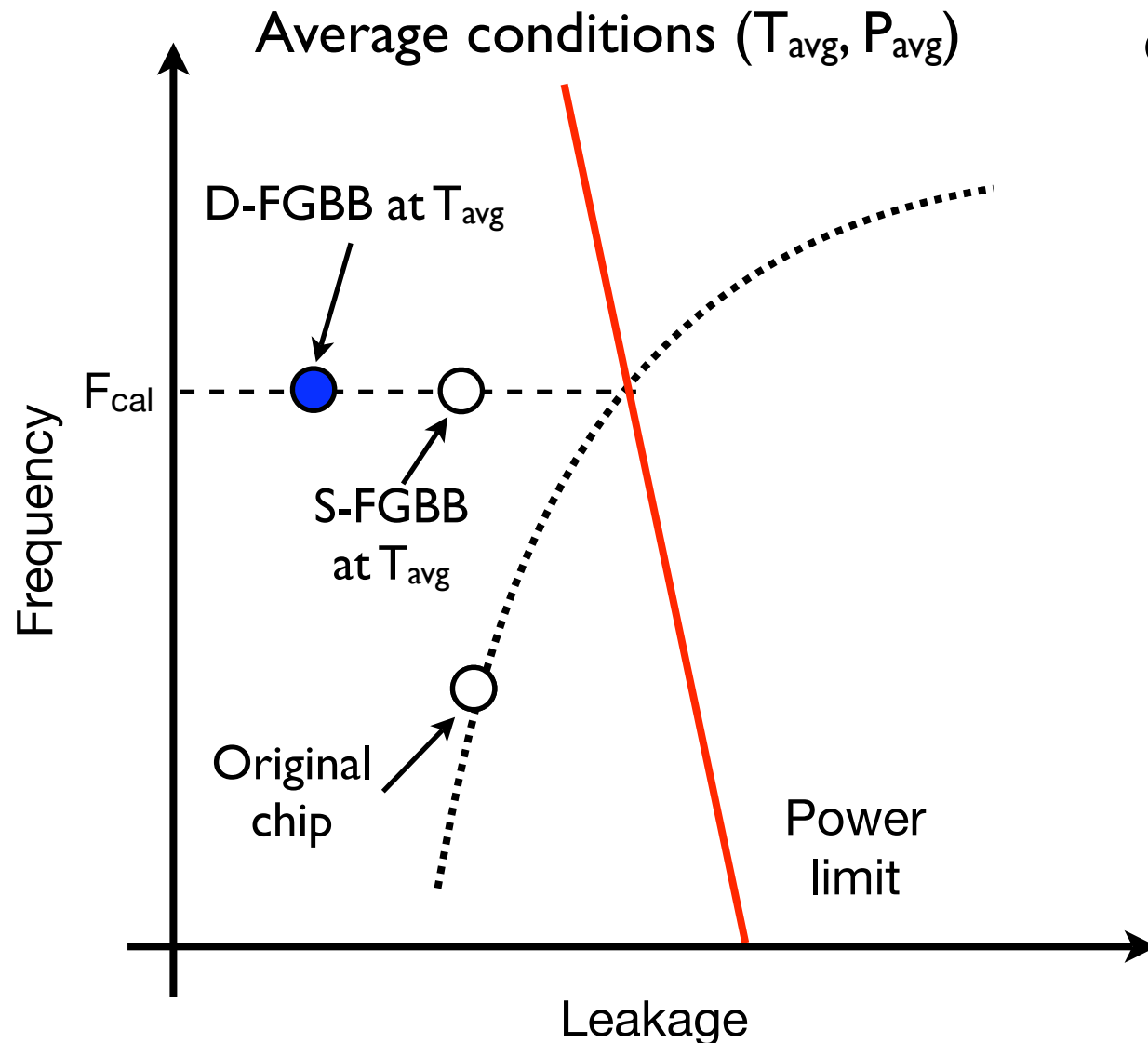
Standard environment



- S-FGGB finds and sets F_{cal}



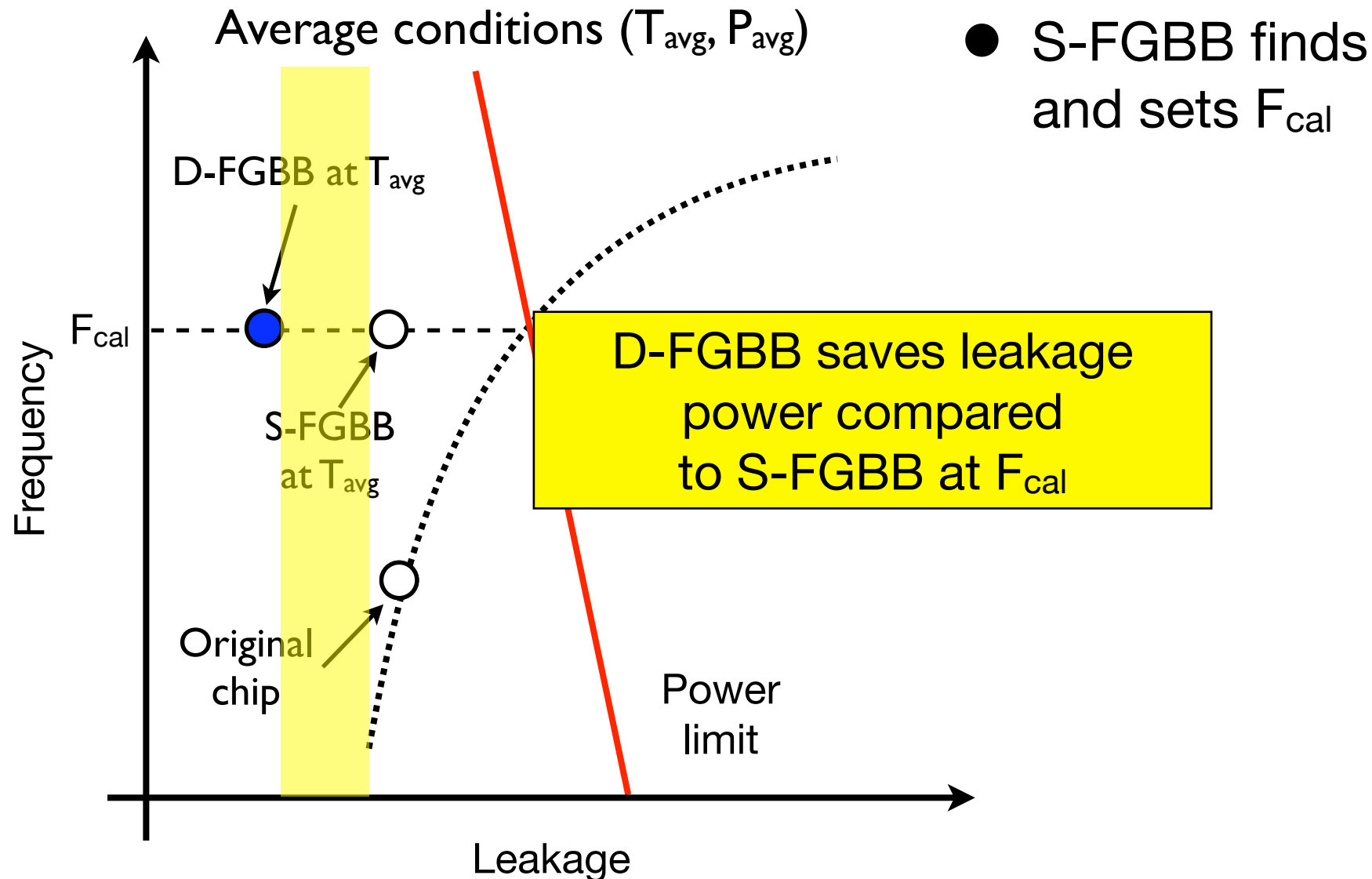
Standard environment



- S-FGGB finds and sets F_{cal}






Standard environment



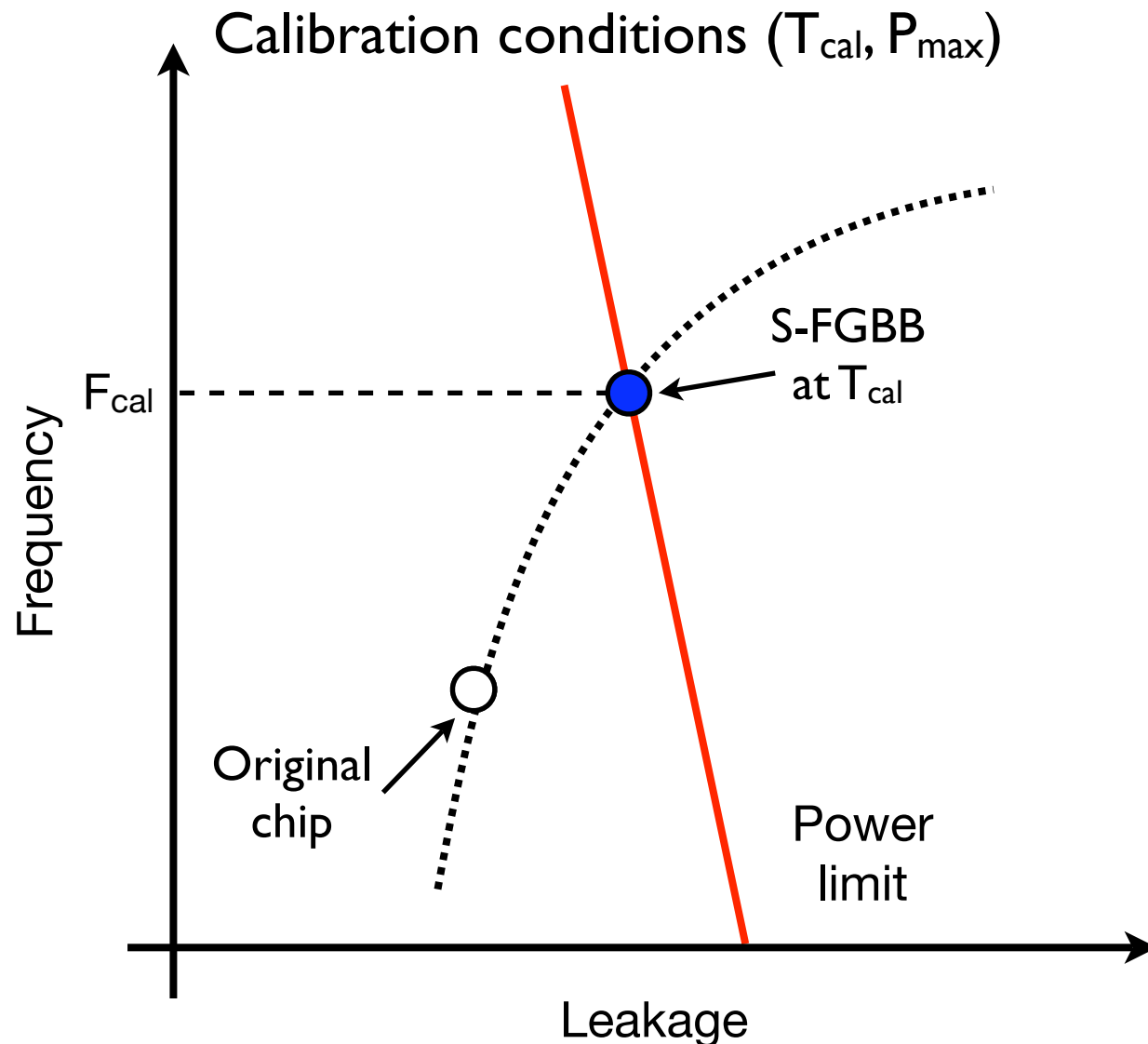


D-FGGB environments

Operating environments		D-FGGB
Standard		Minimize leakage power at F_{cal}
High performance		Maximize average frequency
Low Power		Minimize leakage power at F_{orig}



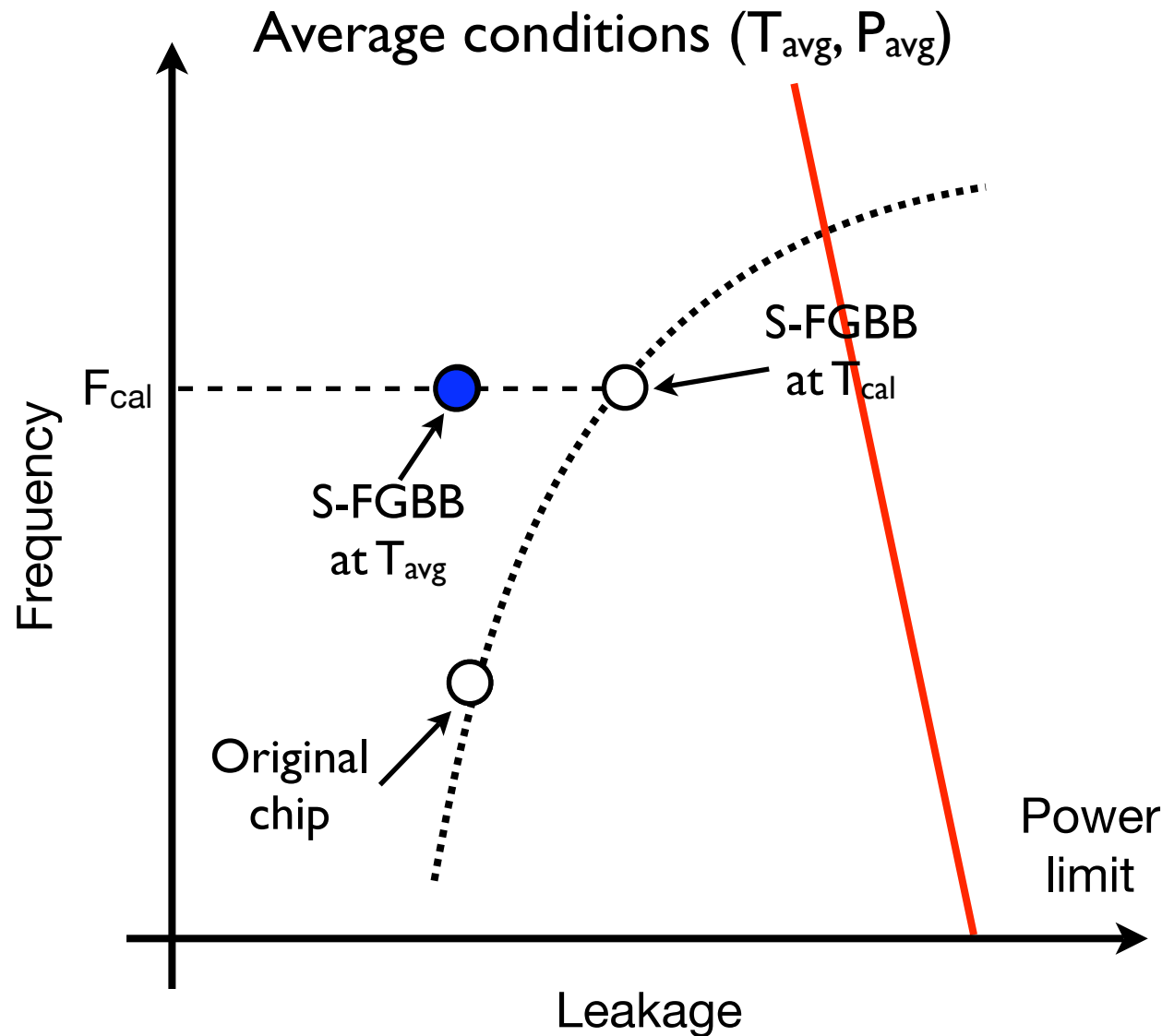
High performance



- Average power
 $P_{avg} \ll P_{max}$



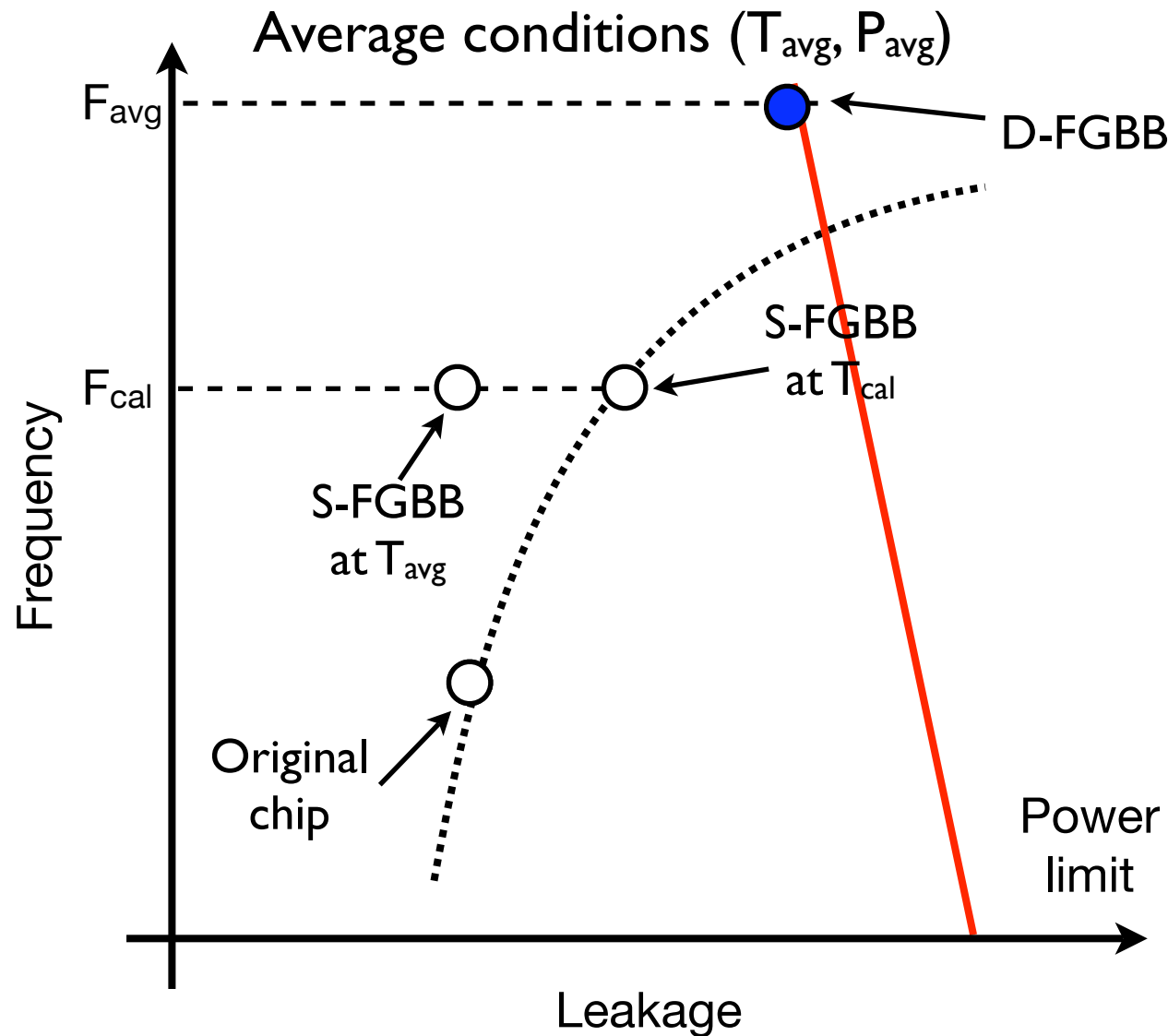
High performance



- Average power
 $P_{avg} \ll P_{max}$



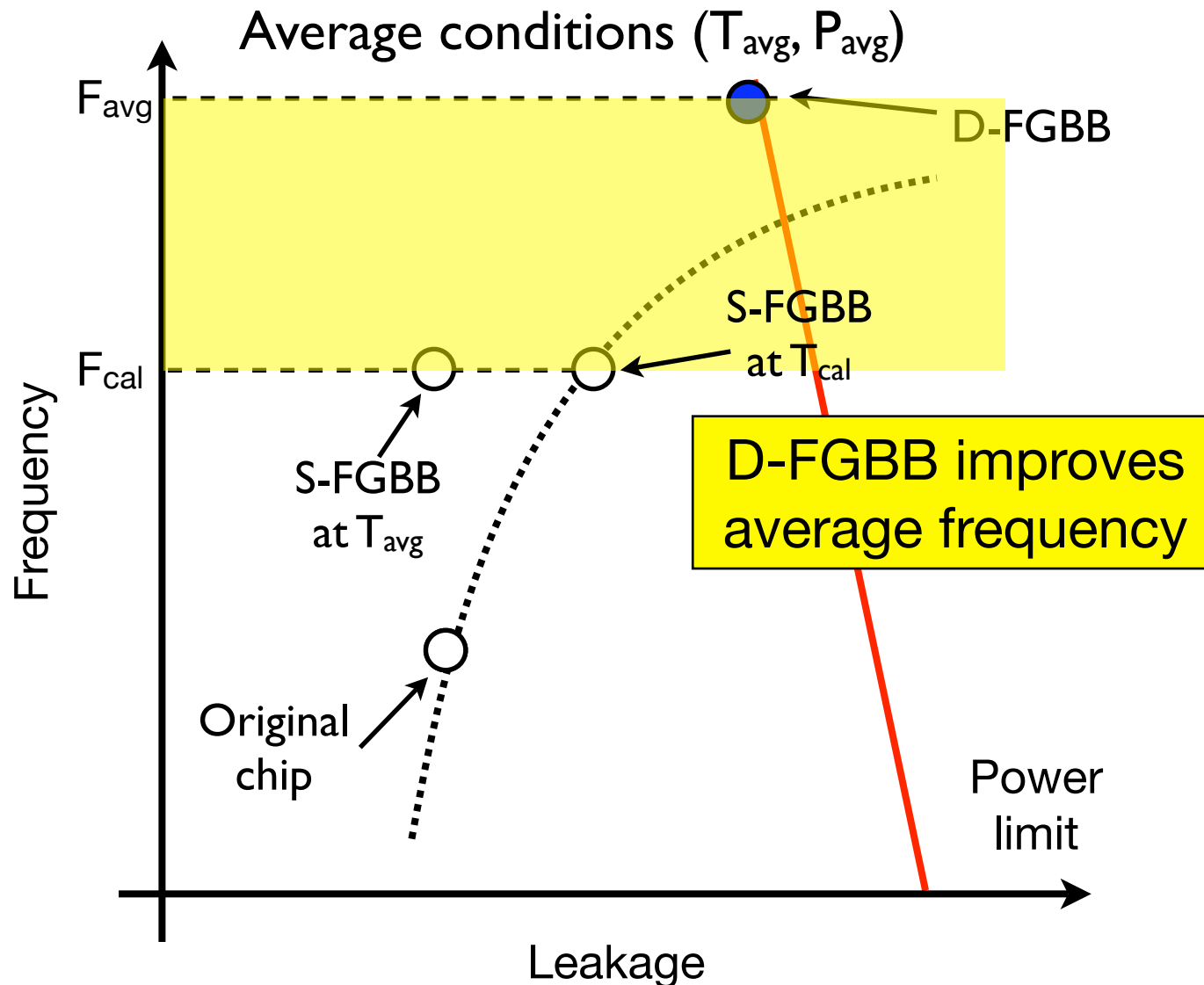
High performance



- Average power
 $P_{avg} \ll P_{max}$






High performance



- Average power $P_{avg} \ll P_{max}$

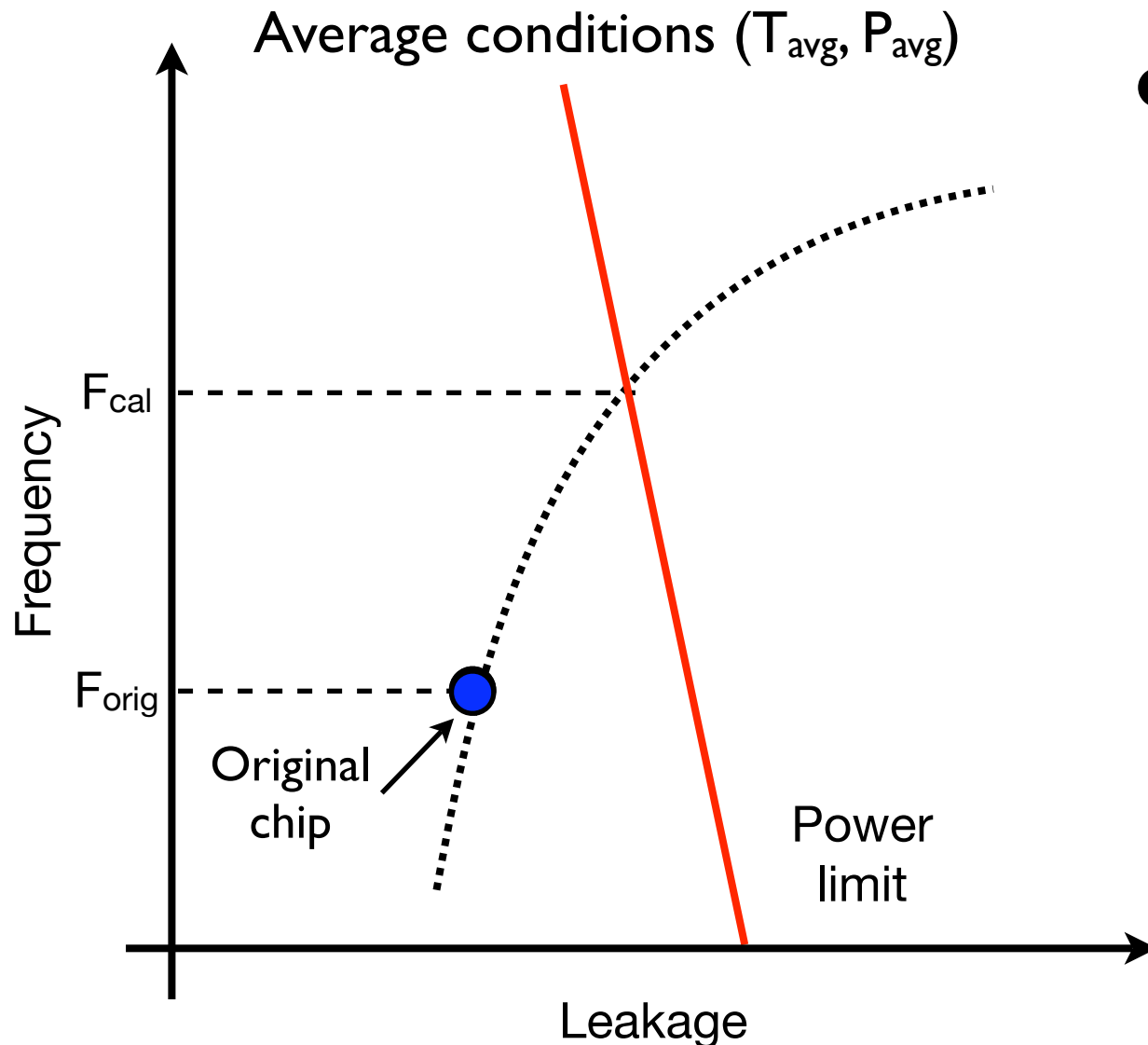


D-FGGB environments

Operating environments		D-FGGB
Standard		Minimize leakage power at F_{cal}
High performance		Maximize average frequency
Low Power		Minimize leakage power at F_{orig}



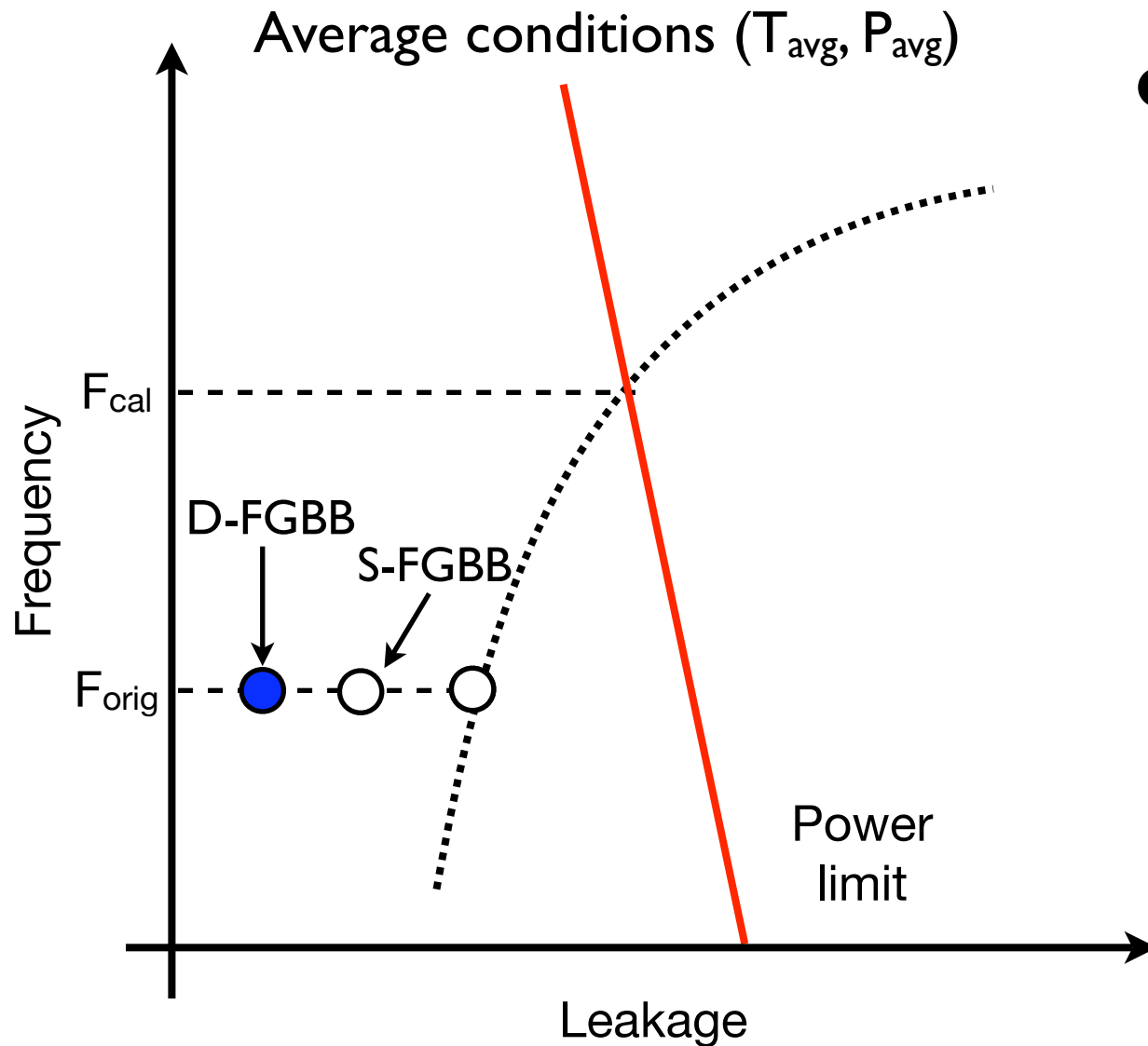
Low power



- The chip runs at its original frequency (F_{orig})



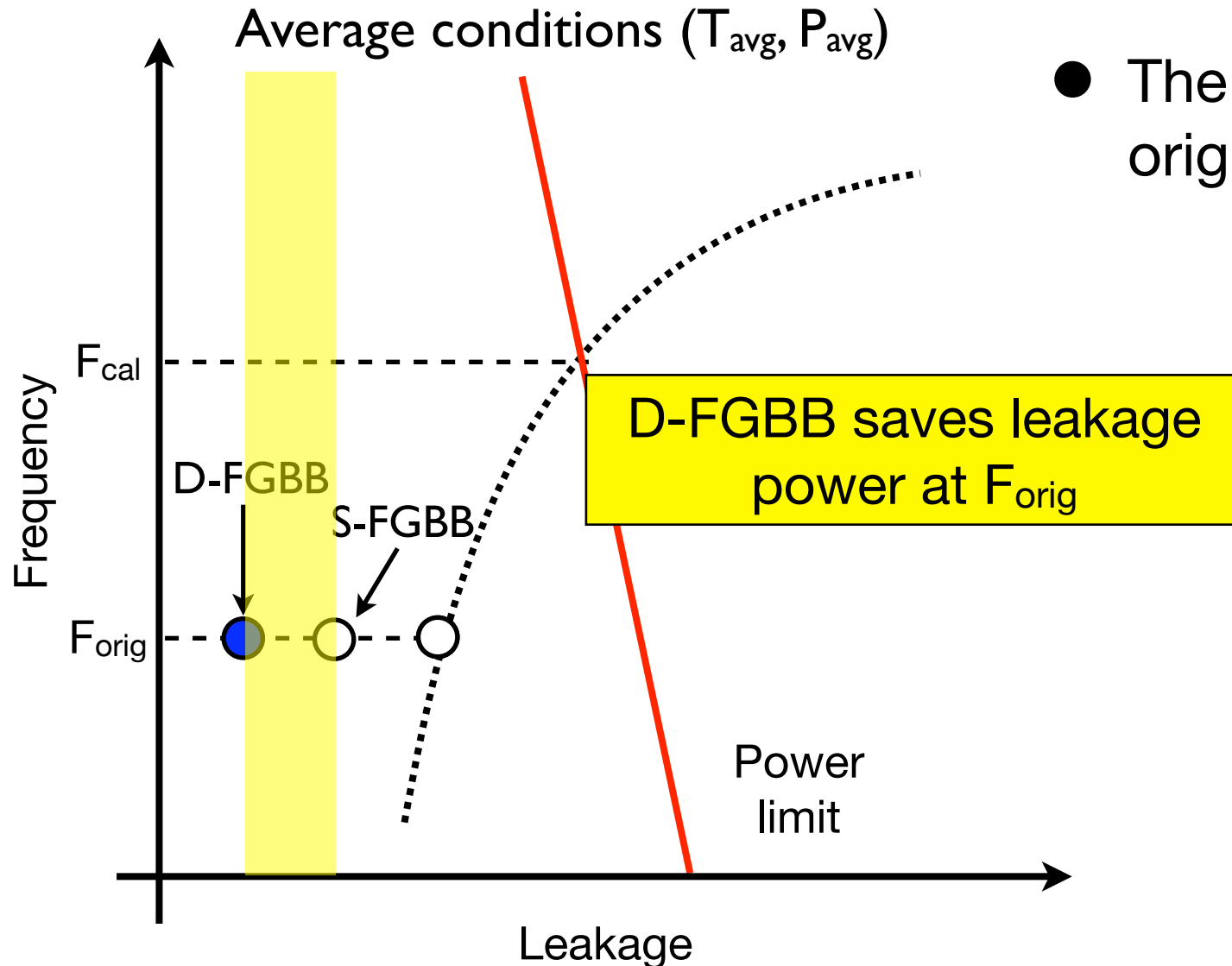
Low power



- The chip runs at its original frequency (F_{orig})



Low power



- The chip runs at its original frequency (F_{orig})



Outline

- Background on S-FGBB
- Dynamic fine-grain body biasing (D-FGBB)
- Environments
- Evaluation
- Conclusions



Evaluation infrastructure

- Process variation model - *VARIUS* [ASGI'07]
 - Generate V_{th} and L_{eff} variation maps for 200 chips
- SESC - cycle accurate microarchitectural simulator - execution time, dynamic power
 - Mix of SPECint and SPECfp benchmarks
- HotLeakage, SPICE model - leakage power
- Hotspot - temperature estimation

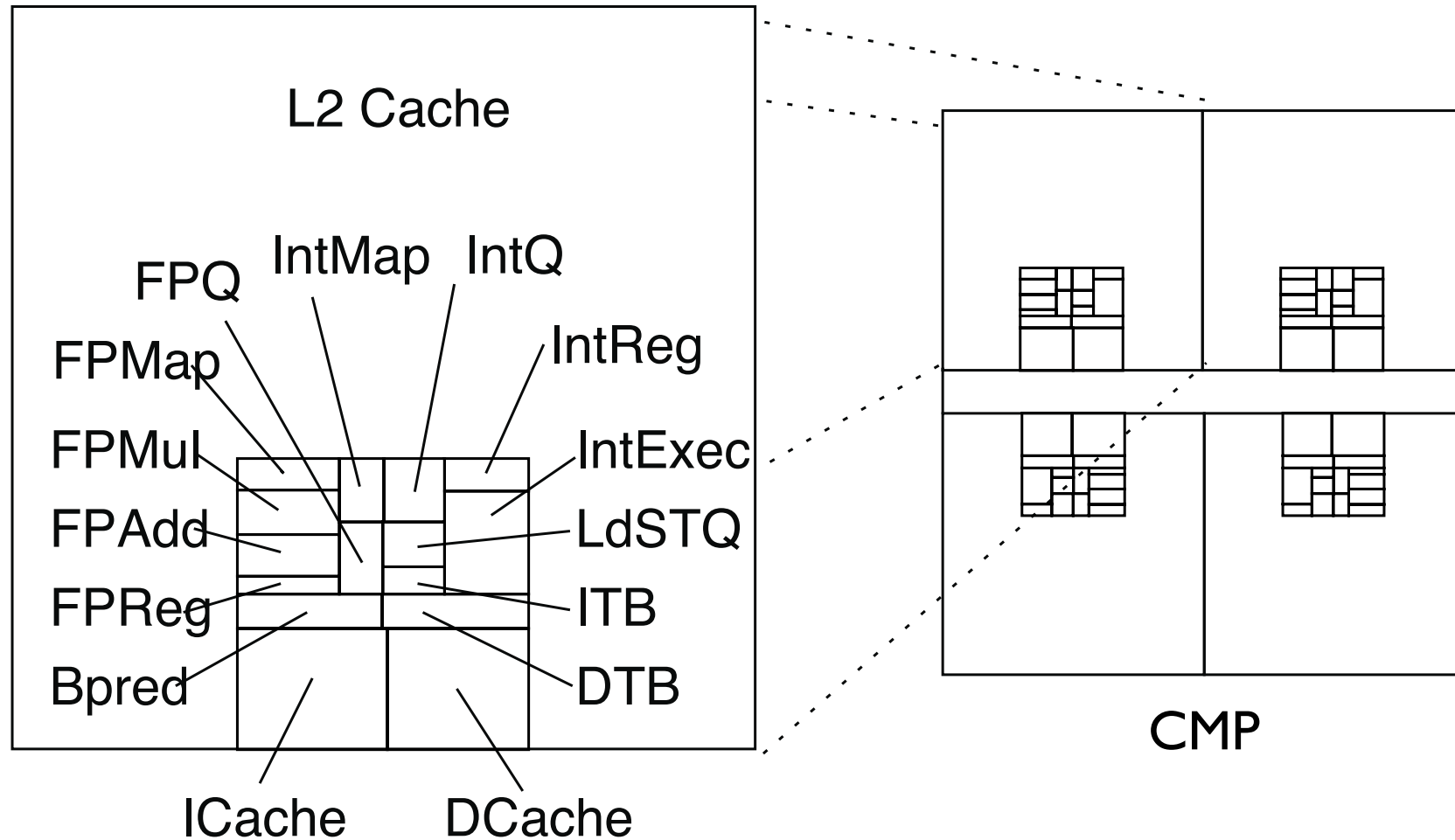


Evaluation parameters

- 4-core CMP, based on Alpha 21364
- 45nm technology, 4GHz
- V_{th} variation: $\sigma_{Vth}/\mu_{Vth}=3-12\%$, $\sigma_{sys}=\sigma_{rand}$
- L_{eff} variation $\sigma_{Leff} = \sigma_{Vth}/2$
- $V_{dd}=1V$, $V_{th0}=250mV$, $V_{bb} = \pm 500mV$



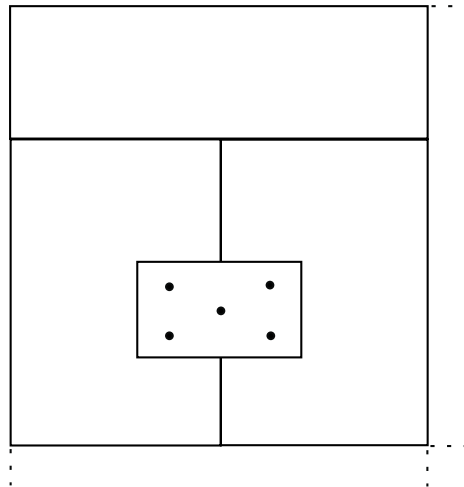
CMP architecture



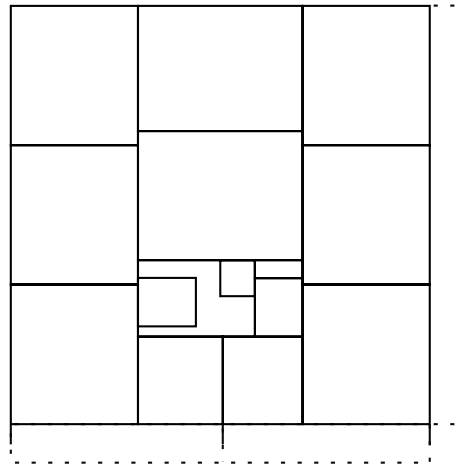


Body bias granularity

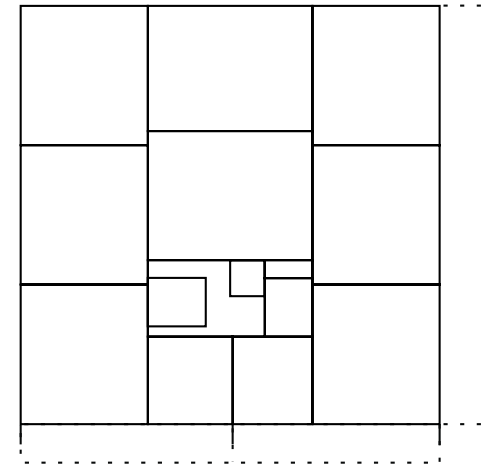
- We evaluate FGBB at different granularities
 - 1 - 144 BB cells per chip
 - Shapes and sizes follow functional units



FGBB16






FGBB64



FGBB144

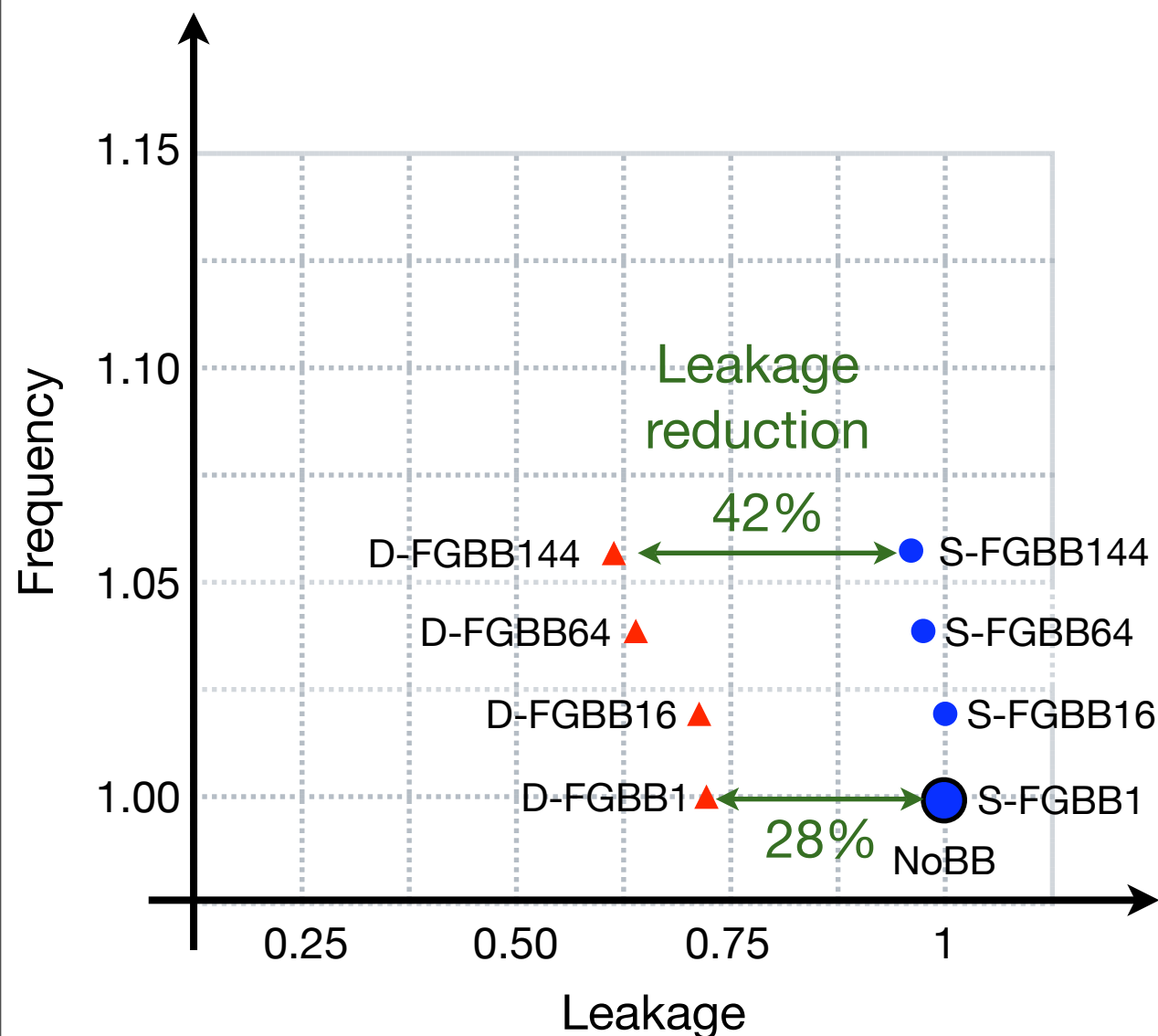


D-FGGB environments

Operating environments	D-FGGB	
Standard 	Minimize leakage power at F_{cal}	
High performance 	Maximize average frequency	
Low Power 	Minimize leakage power at F_{orig}	






D-FGBB reduces leakage



- D-FGBB reduces leakage significantly
- More BB cells result in higher frequency and lower leakage

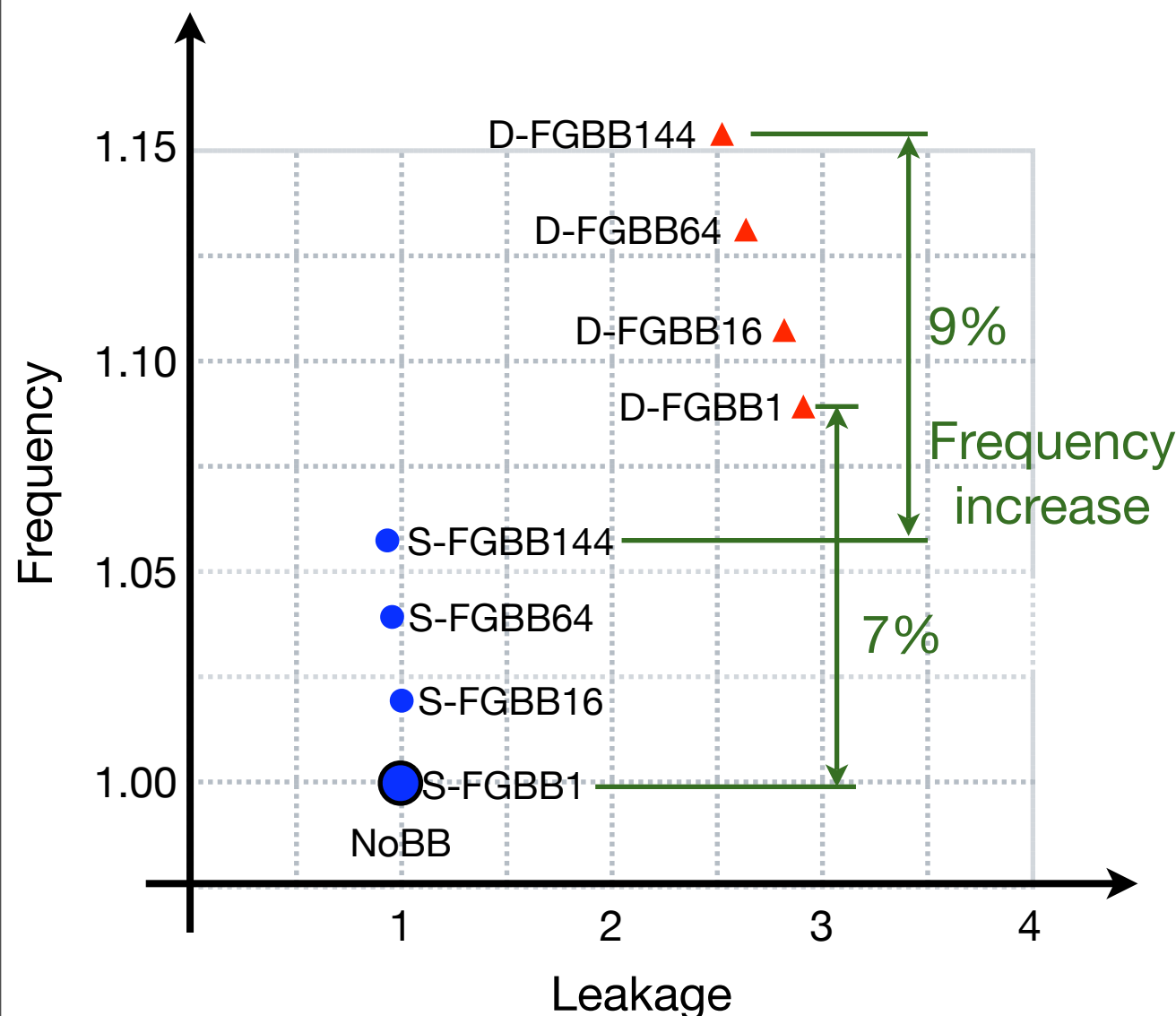


D-FGGB environments

Operating environments		D-FGGB
Standard		Minimize leakage power at F_{cal}
High performance		Maximize average frequency
Low Power		Minimize leakage power at F_{orig}



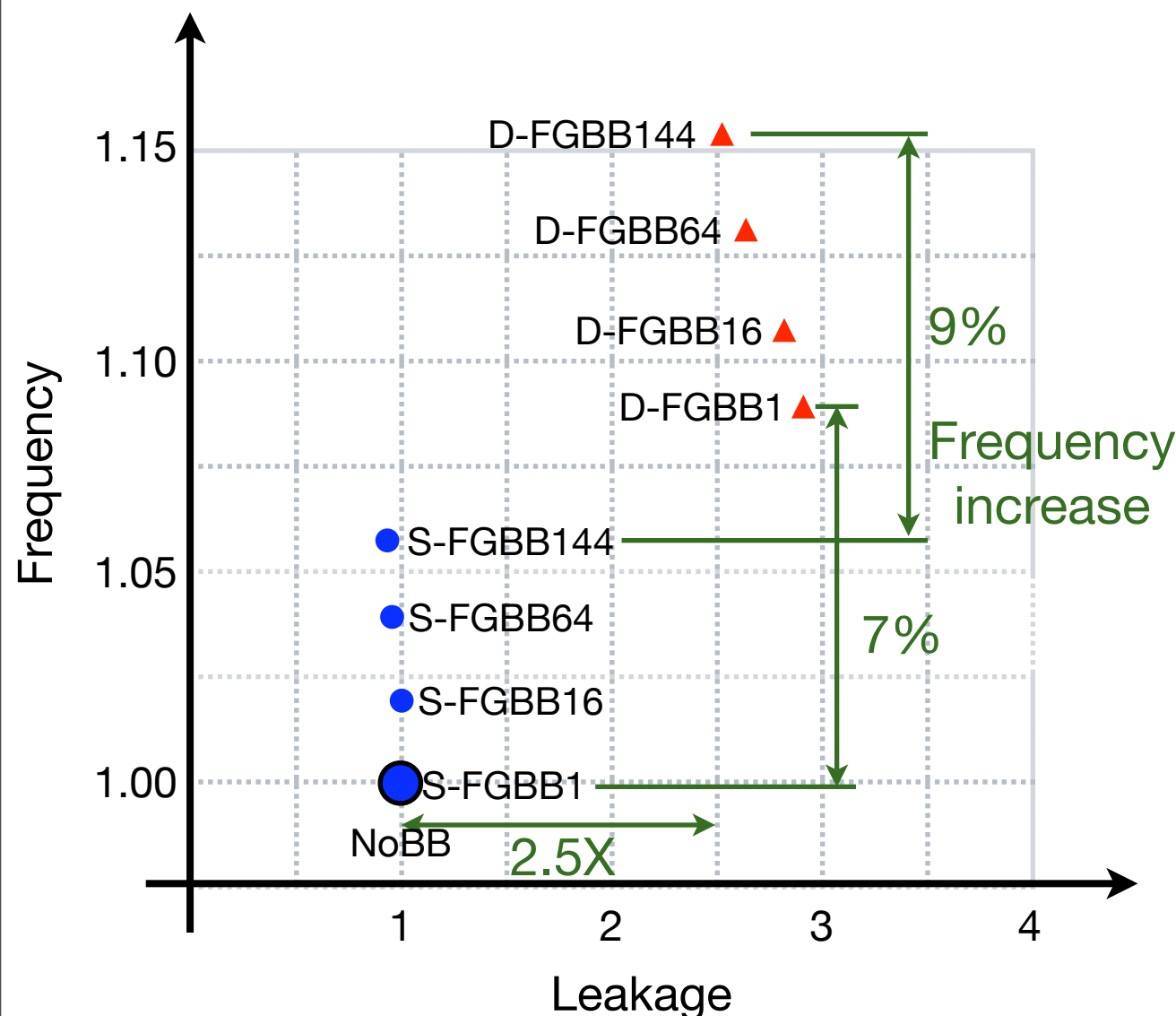
D-FGGB improves frequency



- More BB cells result in a higher increase






D-FGBB improves frequency



- More BB cells result in a higher increase
- Significant power cost, but still within the power budget

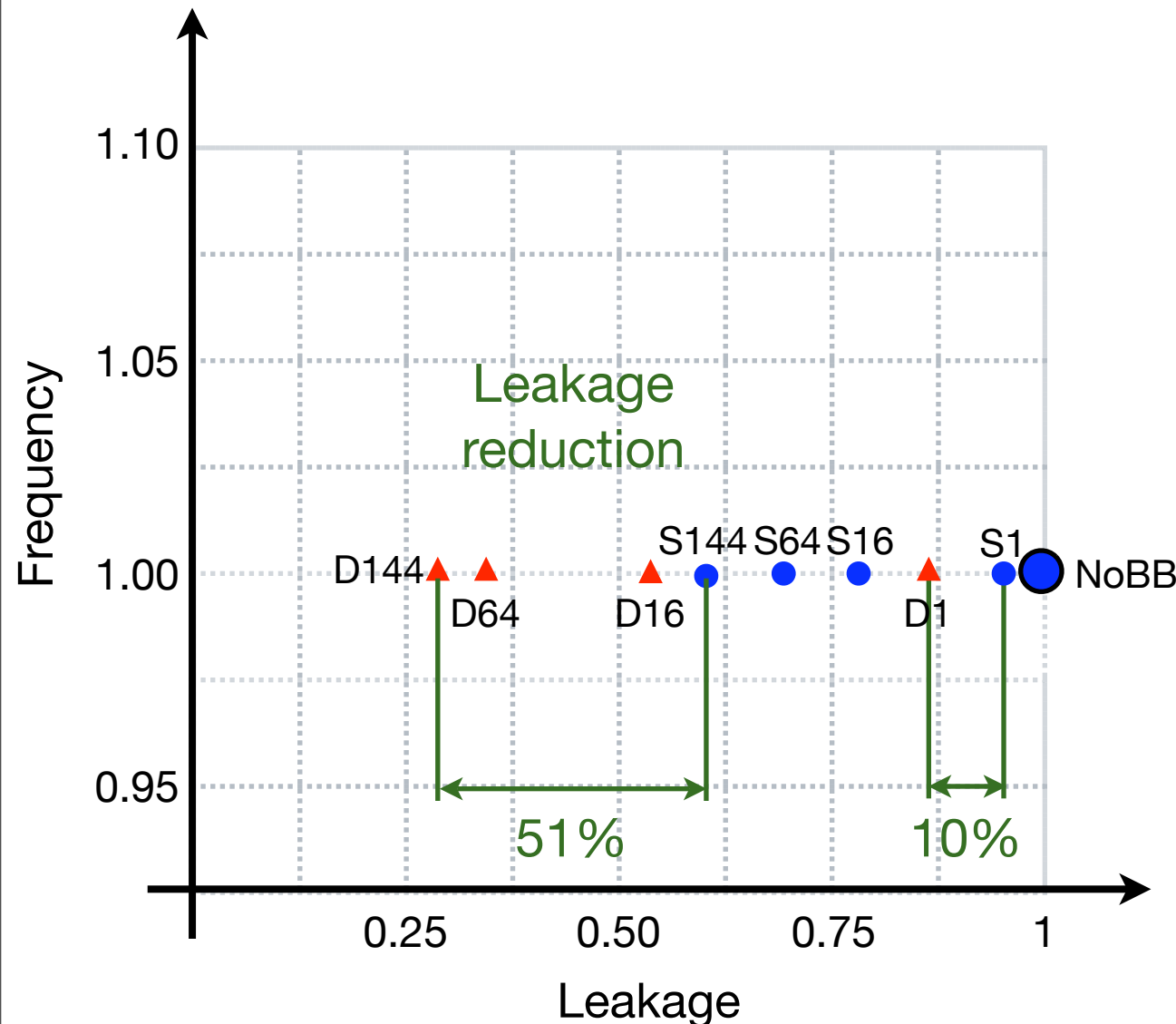


D-FGGB environments

Operating environments	D-FGGB
Standard 	Minimize leakage power at F_{cal}
High performance 	Maximize average frequency
Low Power 	Minimize leakage power at F_{orig}



D-FGBB reduces leakage

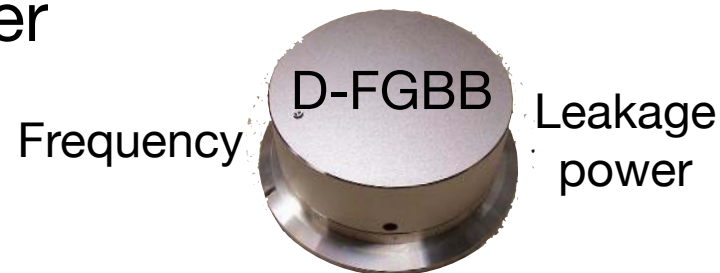


- More BB cells result in higher savings



Conclusions

- D-FGGB is more effective than S-FGGB at reducing WID variation:
 - 50% lower leakage
 - 10% higher frequency
 - because D-FGGB adapts to T variation
- D-FGGB can give architects an additional knob to tradeoff frequency/power





More in the paper...

- Details about our variation model
- A solution for combining D-FGGB with DVFS
- Estimated overheads of D-FGGB
- More implementation details

Mitigating Parameter Variation with Dynamic Fine-Grain Body Biasing

**Radu Teodorescu, Jun Nakano, Abhishek Tiwari
and Josep Torrellas**

University of Illinois at Urbana-Champaign
<http://iacoma.cs.uiuc.edu>

