Design and Use of Tweakable Devices for Future Logic Implementation

Puneet Gupta

puneet@ee.ucla.edu

http://www.ee.ucla.edu

Agenda

- Introduction
- Example 1: Gate Length Biasing
- Example 2: Shaping Transistor Channels
- Conclusions

Flexibility: The Motivating Observation

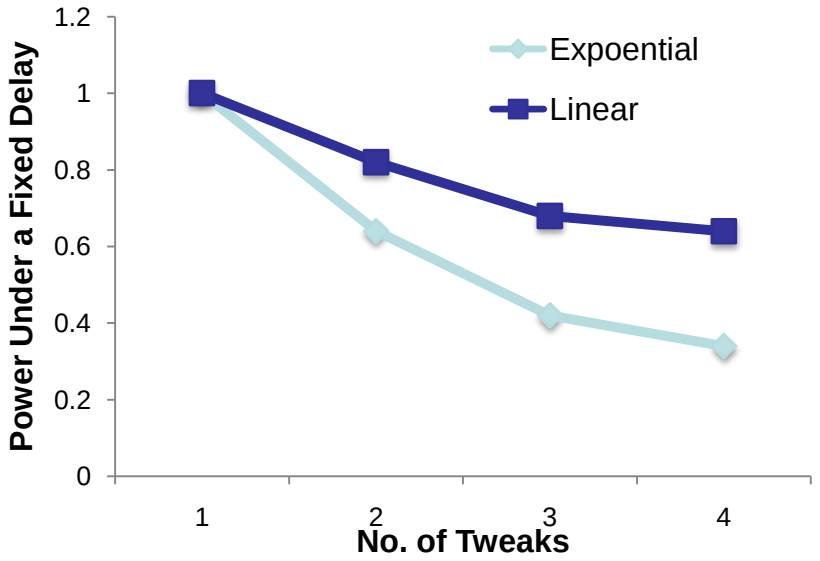
- Circuit designers like flexibility
  - Multiple device options
  - Layout choices
- Circuits implicitly offer flexibility
  - Large number of devices
  - Differing requirements for different parts of the circuit
- *Circuit flexibility can help relax device requirements*
  - *Can tradeoff device "goodness" vs. number of device choices*
  - *Device choices → "tweaks"*
- *This talk: how digital logic implementation leverages tweakable devices*
  - Tweaks to exploit power vs. performance tradeoff

Digital Logic

- Concerns
  - Power
    - Switching
    - Leakage (~Ioff) → focus of this talk
  - Performance (~Ion)
- Characteristics
  - "Random" →no clear structure
  - Huge: O(100M) devices
  - *Huge + Random → Flexibility*
  - Designed with cell-level abstractions → digital designers don't really see transistors
  - Optimization through automated tools
    - Large scale optimization

Digital Logic Optimization

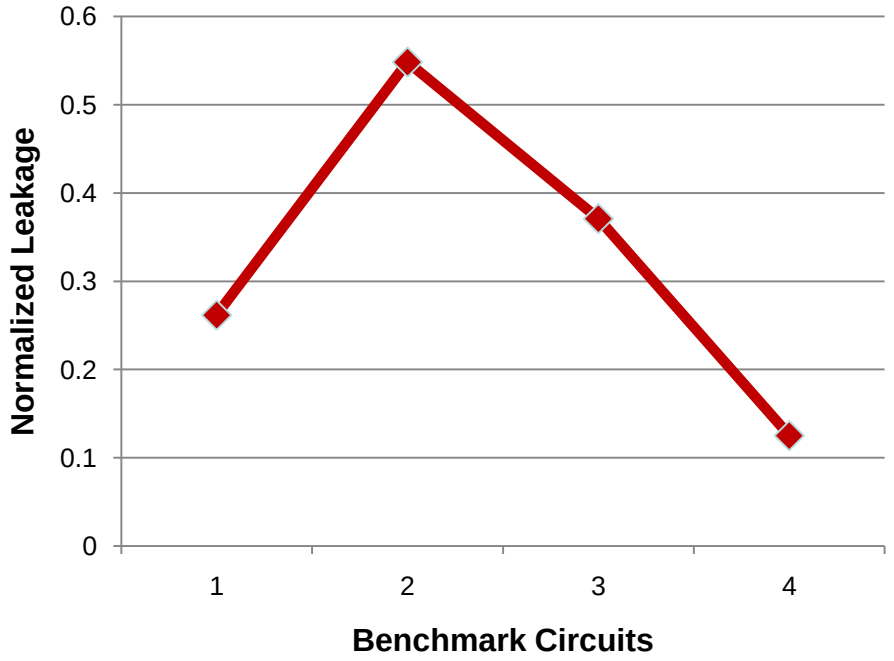
- Only 10%-20% of cells (devices) are timing critical →
  - Performance determined by 10% of devices
  - Power determined by 100% of devices
- Simple experiment
  - 1-4 *tweaks*
  - power-performance tradeoff per tweak
    - Exponential (e.g., Vth)
    - Linear (e.g., width)
  - Usage of device directly proportional to delay
    - Faster devices used only in critical paths!



More tweaks ≈ Better perceived single device

Example: The Vth Tweak

- 130nm dual-Vth process
  - The tweak: Vth (typically implemented using an implant change)
- 4 benchmark circuits
  - Compare SVT implementation with SVT + HVT with same performance



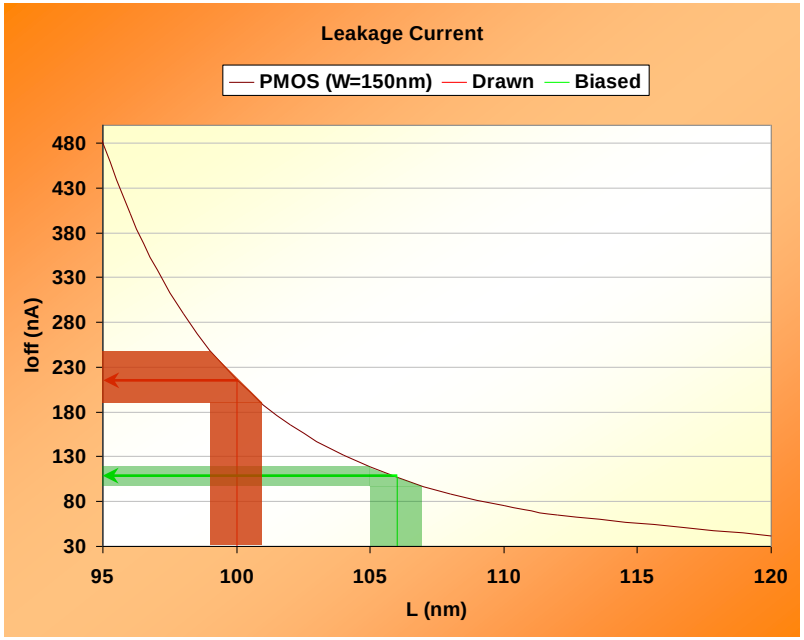
Taxonomy of Tweakable Devices

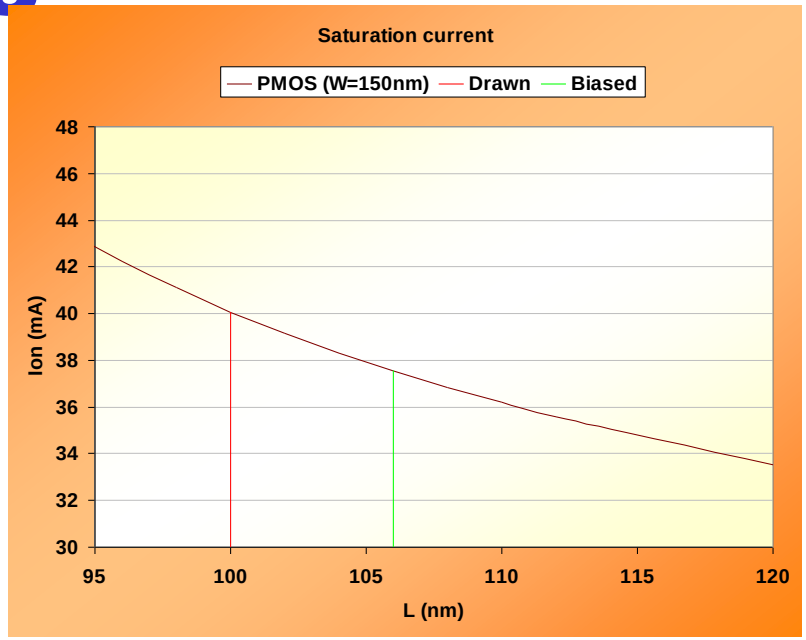
- Runtime tweaks
  - Adjustable essentially through “software”
    - Does not require change in fabrication
    - Will still require change in design/layout
  - E.g., adaptive body biasing, dynamic voltage scaling
- Fabrication-time tweaks
  - Adjusted using manufacturing process
    - May require change in layout
    - Fixed post-fabrication
  - E.g., multiple Vth, strain modulation
- Both are *not* mutually exclusive
  - multiple tweaks → more benefit

Tweak Requirements

- Absolutely need SPICE-simulatable model for device behavior
  - Hopefully *n* tweaks does not mean *nX* model extraction time
- Small layout/design overhead
  - Least intrusive tweaks get adopted most
- Manageable process overhead
  - Good process control
  - No major process changes
- Nominal and tweaked device have similar behavior across process-voltage-temperature
  - Important for design verification
- Nice-to-have: small mismatch between nominal and tweaked device
  - to allow usage in “skew” sensitive structures (e.g., clock)

Example 1: Gate-Length Biasing





Impact on leakage:

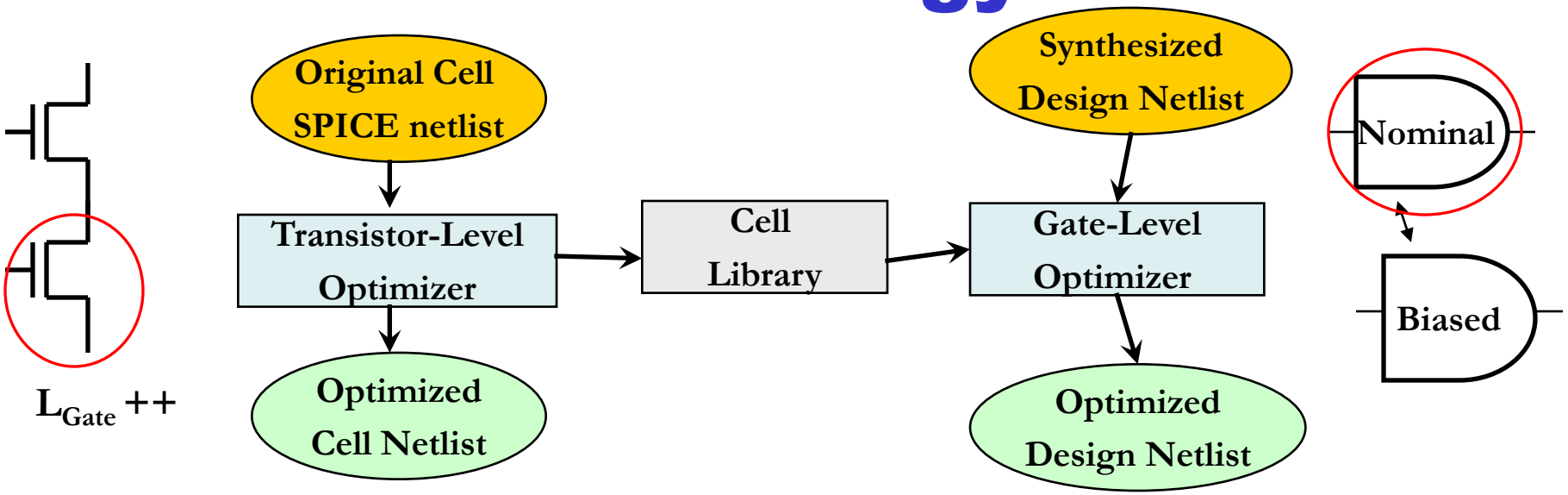
- Exponential leakage reduction
- Exponential leakage variation reduction

Impact on drive strength:

- Linear drive strength reduction

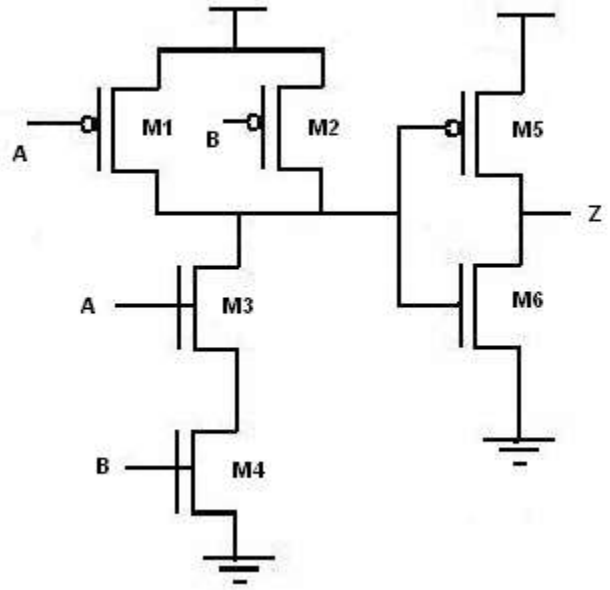
- Idea: Use small biases with a fine granularity (e.g. 2nm, 4nm)
  - Small leakage reduction beyond 10% biasing
  - Preserve pin-compatibility → Technique can be applied post-routing

Gate Length Biasing Methodology



- Extend a standard cell library with biased L_{Gate} versions of all cells
  - Cells optimized at transistor-by-transistor to achieve best leakage-delay tradeoff
- Optimize circuit for leakage by using biased L_{Gate} versions for non-critical cells
  - A static-timing driven heuristic sizing algorithm used

Transistor-Level Biasing



Input State		Device					
A	B	M1	M2	M3	M4	M5	M6
0	0	D	D	N	N	L	D
0	1	D	N	L	N	L	D
1	0	N	D	N	L	L	D
1	1	L	L	D	D	D	L

M3,M4,M5

Only 1 delay dominant state, ’11’

M3,M4

Only 1 leakage dominant state

M5

3 leakage dominant states

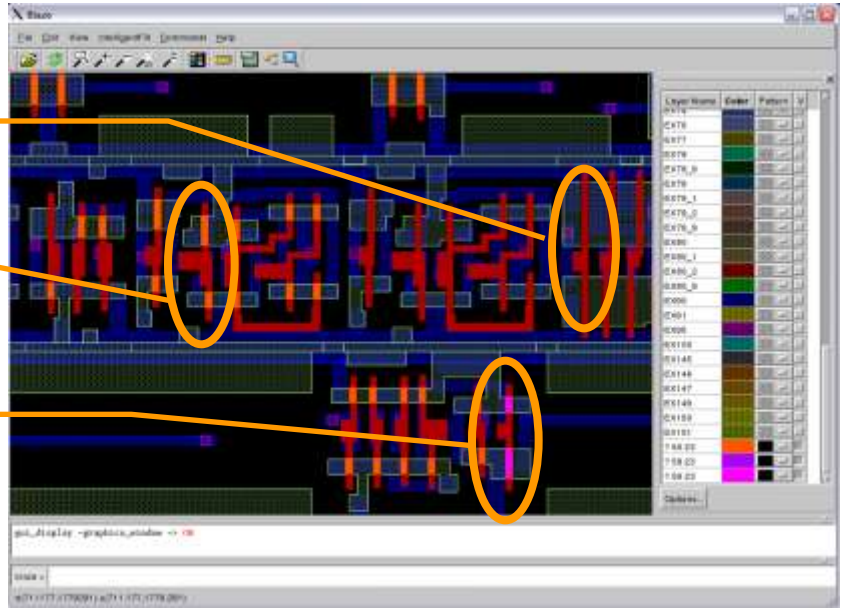
Reduce bias of M3,M4; increase bias of M5, to maintain delay of ’11’ input state

LGate Bias Through OPC

Transistor on critical path
⇨ use no bias (0nm)

Transistor on nearly critical path
⇨ use minimum bias (+2nm)

Transistor on non-critical path
⇨ use maximum bias (+6nm)



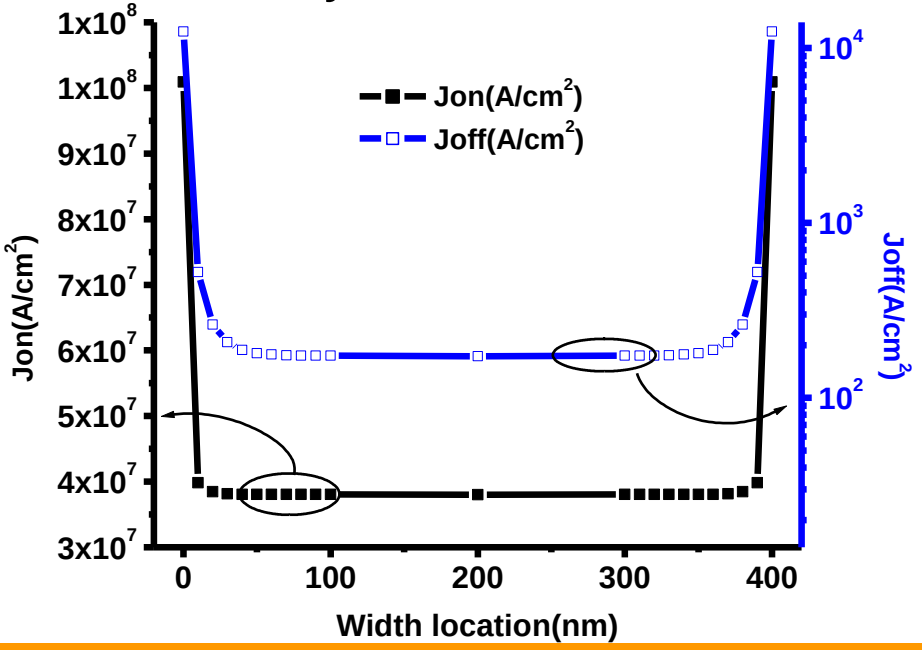
- Tiny gate-length changes (within foundry-supported bounds) implemented during OPC
  - No design changes, no methodology changes
  - Marker shapes passed in extra GDS layer to foundry → shift target edge placements for OPC

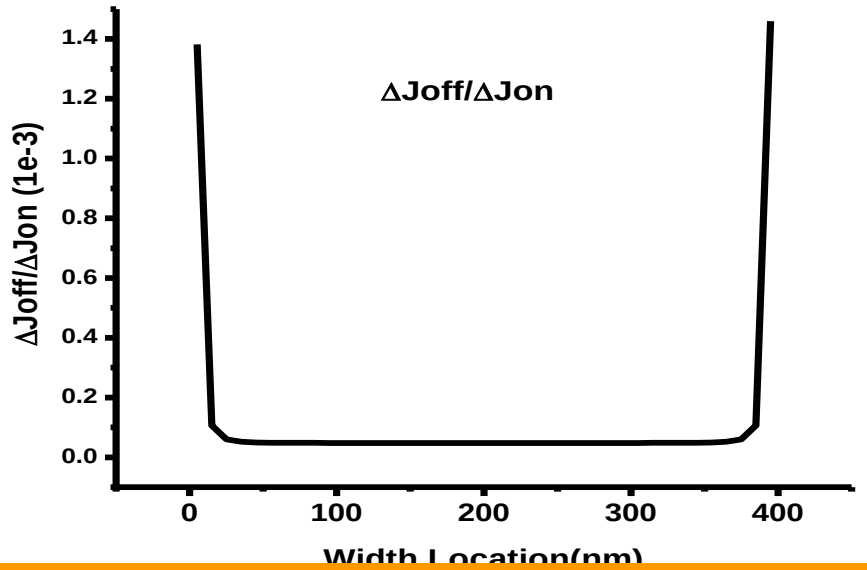
Sample Results

Process	Library Type	Cell Count	Leakage Reduction at Block/Full-Chip Level (inc. Mem)	Leakage Reduction at Std-Cell Level
55nm GP	Multi Vt	>10M	**19%**	**24%**
65nm LP	Multi Vt	>100K	**30%**	**50%**
65nm LP	Multi Vt	>100K	**20%**	**35%**
65nm G	Multi Vt	>100K	**30%**	**45%**
65nm G	Multi Vt	>1M	**25%**	**30%**
90nm G	Multi Vt	>100K	**30%**	**38%**
90nm LP	Multi Vt	>1M	**15%**	**30%**
90nm LP	Multi Vt	~100K	**20%**	**40%**
90nm LP	Multi Vt	>1M	**20%**	**27%**
90nm G	Single Vt	~100K	**30%**	**48%**
90nm G	Single Vt	>100K	**30%**	**52%**

Example 2: Non-Rectangular Gate Channels

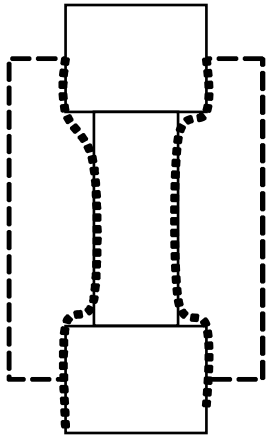
- Ion/Ioff densities depend on distance from device (STI) edge
  - Line-end capacitance, dopant concentrations → lowered Vth near edges → Better delay-leakage tradeoff at the center than at edges → Uniform channel length suboptimal
  - Longer L at edges and shorter L in center → lower leakage for same delay

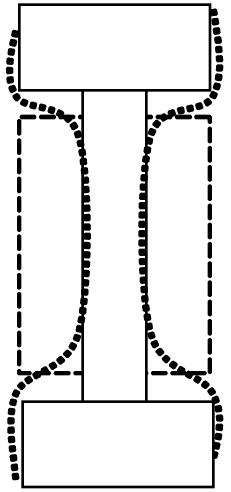




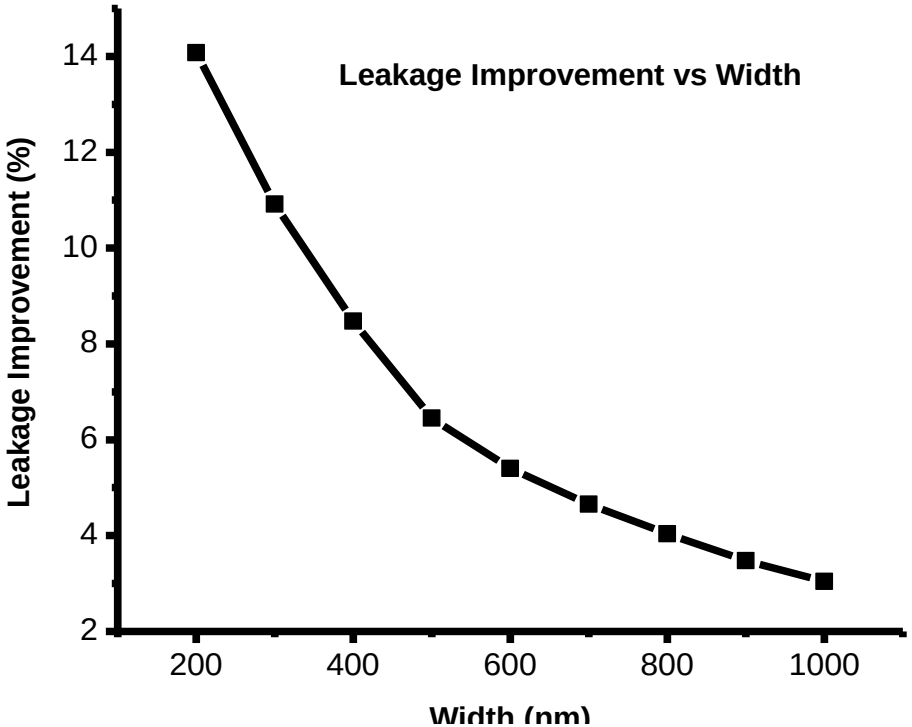
Laying out Non-Rectangular Transistors

- Active shape perturbation
  - Explicitly draw poly as intended
  - May require RET and DR waivers
- Passive shape perturbation
  - Non-gate poly is changed
  - Does not require waivers
  - Much weaker a knob
- Active perturbation only in this work
  - Only dumbbell shape considered → maximum 4 extra jogs per device.
- Electrical Constraints
  - Same or less Ion
  - Same or less area → reduced capacitance
- Goal: minimum Ioff





Results: Device-Level



- A commercial 90nm technology
- Results shown are for NMOS (PMOS similar)
  - Ioff reduction with constant Ion

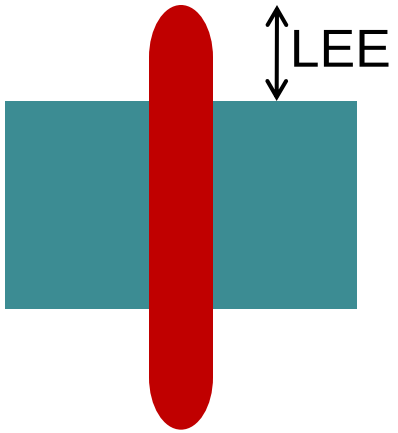
Results: Design-Level

Circuit Name	Orig. Delay (ns)	Opt. Delay (ns)	Orig. Leakage (uW)	Opt. Leakage (uW)	% Imp.
C432	1.87	1.87	9.60	9.11	5.1
C1908	2.24	2.24	11.98	11.38	5.0
C2670	1.55	1.55	18.62	17.68	5.0
C3540	2.84	2.84	4.44	4.22	4.9
C5315	1.96	1.95	31.93	30.46	4.6
C6288	5.62	5.61	39.66	38.38	3.2
C7552	3.19	3.19	36.78	35.08	4.6
i2	0.86	0.86	13.55	12.80	5.5
i3	0.45	0.45	6.07	5.74	5.4
i4	0.58	0.58	5.46	5.21	4.6
i5	0.52	0.52	9.22	8.77	4.9
i6	0.59	0.59	10.72	10.23	4.8
i7	0.72	0.72	14.22	13.50	5.1
i8	1.01	1.01	25.19	23.98	4.8
i9	1.37	1.37	16.28	15.40	5.4
i10	2.27	2.27	55.82	53.06	4.9

- ISCAS85 and MCNC Benchmarks
- Average 4.9% reduction

Other Possible Tweaks

- Axes of interest: delay, leakage power, switched capacitance, area, variability
  - Tweaks trade off one with other
- Strain tweakable per device ?
  - E.g., [Kahng et al.'07] modulate STI-induced strain by inserting active dummies
- Area-variability tradeoffs ?
  - Which spacing/extension/enclosure design rules can we shrink at cost of increased device variation ?
    - E.g., line-end extension rule [Gupta et al.'08]
- Others ?



Tweaks are Not Free

- Probably cheaper than engineering and manufacturing a new device though..
- Overheads
  - Layout: almost always require new cell library layout (sometimes the changes are trivial)
  - Characterization: lots of circuit simulation runs to get power/performance models for tweaked cells
  - Physical Design: more complex circuit optimization
    - Sometimes methodology changes as well. E.g., separate routing for body bias lines
  - Modeling: additional extraction
  - Process: hopefully just a different parameter value

Conclusions

- Circuit optimization based use of multiple device tweaks gives an illusion of a better device
  - Too many tweaks not necessarily good → not enough bang for buck for increased circuit optimization complexity
- When engineering a device
  - Allow for few controllable perturbations to its Ioff/Ion characteristics
    - Tweaked device may be worse → just need to offer a tradeoff
- Side benefit
  - Models for tweaks may allow for relaxed process control
    - E.g., non-rectangular gate models may allow less aggressive OPC
- *Give designers easily tweakable devices!*