Session 15 – Design Foundations for Advanced Technologies

Analog/Mixed-Signal Design Challenges in 7-nm CMOS and Beyond

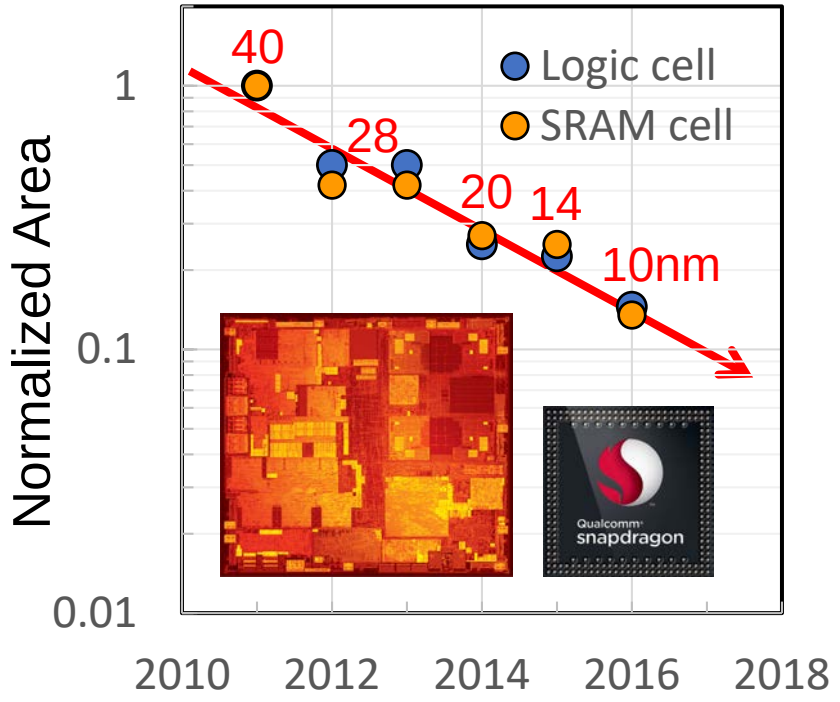
A.L.S. Loke, D. Yang, T.T. Wee, J.L. Holland, P. Isakanian, K. Rim, S. Yang, J.S. Schneider, G. Nallapati, S. Dundigal, H. Lakdawala, B. Amelifard, C.K. Lee, B. McGovern, P. S. Holdaway, X. Kong, and B.M. Leary

Qualcomm

San Diego, CA & Raleigh, NC

Mobile SoC – Main Driver for CMOS Scaling

- 7nm smartphone products imminent
- SoC technology driven by economics of logic & SRAM scaling
- New node feasible with enough PPAC (Power/Performance/Area/Cost) benefit
  - Incremental feature size reduction
  - Extensive logic & SRAM DTCO
- AMS (Analog/Mixed-Signal) ubiquitous
  - PLLs, wireline I/Os, ESD, regulators, data converters, bandgap references
  - AMS device palette slaved to logic



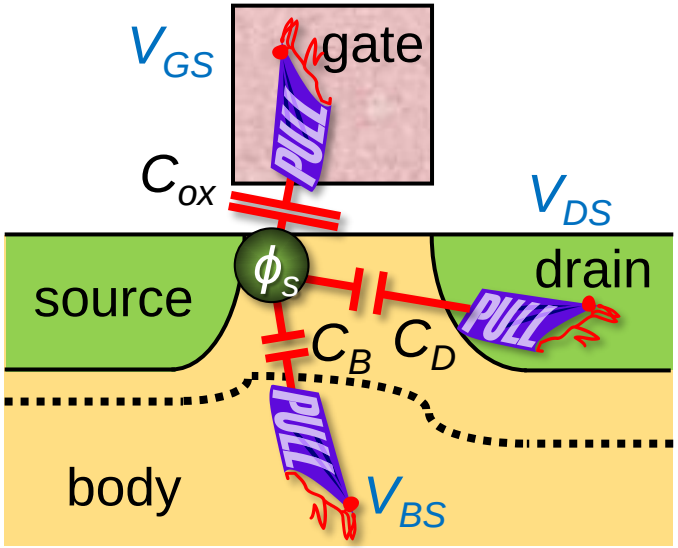
Outline

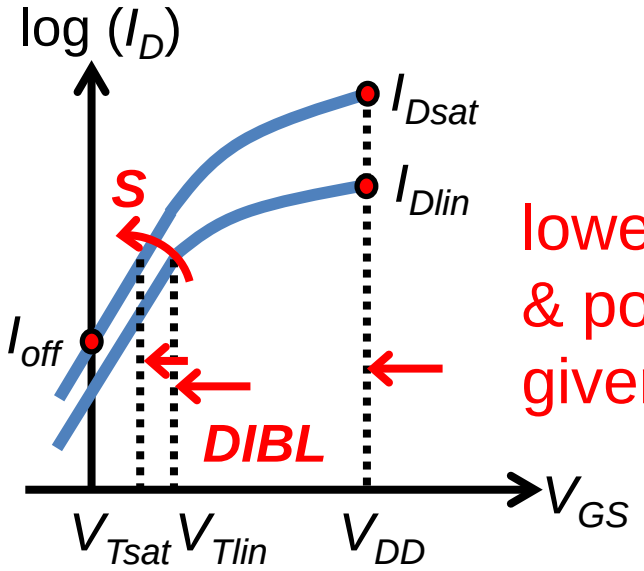
- Introduction
- Technology Scaling Enablers
- AMS Device Palette
- AMS Design Impact
- Concurrent Technology/Design Development
- Conclusion

Outline

- Introduction
- **Technology Scaling Enablers**
  - **FinFET**
  - **Lithography & Self-Aligned Patterning**
  - **High-*K* Gate Dielectric & Metal Gate (HKMG)**
  - **Mechanical Stressors**
  - **Middle-End-Of-Line (MEOL)**
- AMS Device Palette
- AMS Design Impact
- Concurrent Technology/Design Development
- Conclusion

Stronger Short-Channel Gate Control

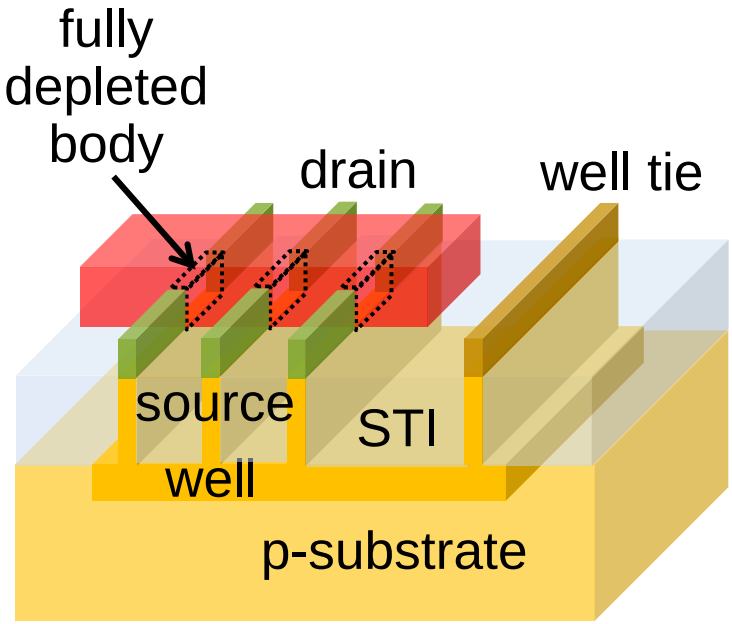




lower supply & power for given I_{on} & I_{off}

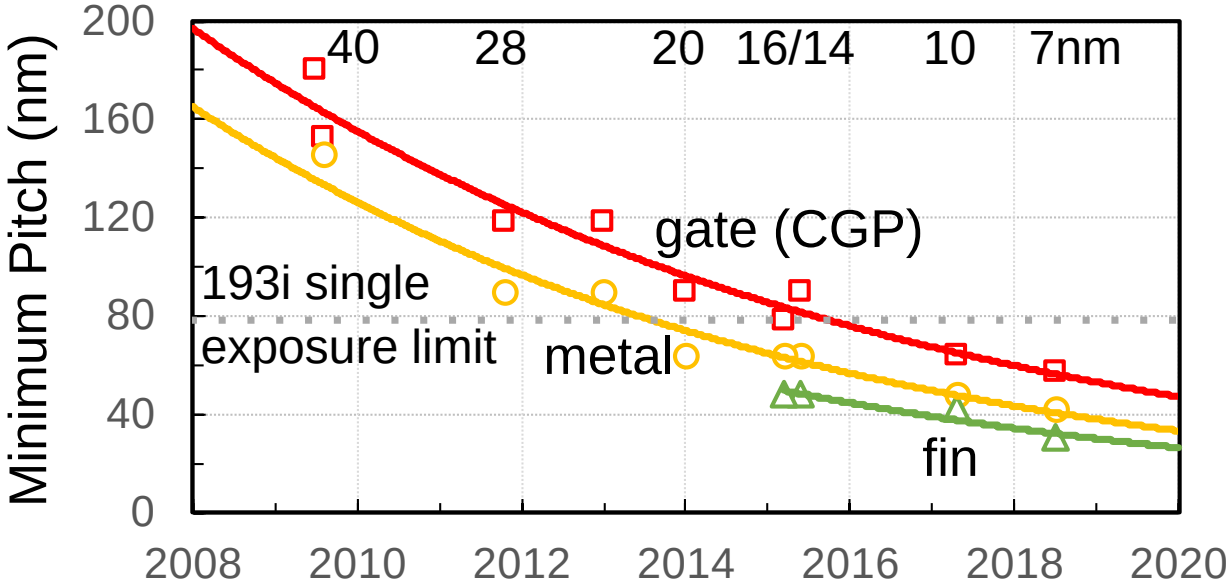
- Subthreshold controlled by C_{ox} vs. (C_B+C_D)
- Fully-depleted finFET weakens C_B & C_D → less *S*, *DIBL* & body effect

Fully Depleted Tri-Gate FinFET



- More I_{on} & g_m per area
- Quantized channel width
- Less DIBL → higher r_{out}, 3× intrinsic gain
- Negligible body effect ($\Delta V_T < 10mV$)
- Less RDF mismatch
- Parasitics
  - High S/D resistance → big deal
  - High S/D coupling to gate
  - Fin width << fin pitch → low C_j, high R_{well}

Foundry Pitch Scaling

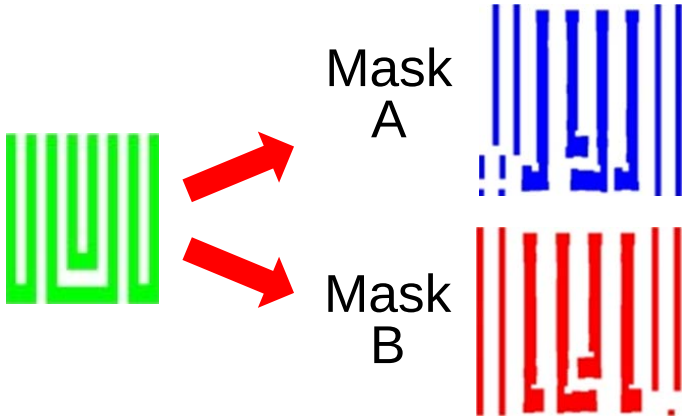


- Scaling rate slower than 0.7x per node → node name = PPAC marketing
- EUV late, only started at 7nm → process complexity for sub-80nm pitch

Lithography Innovations

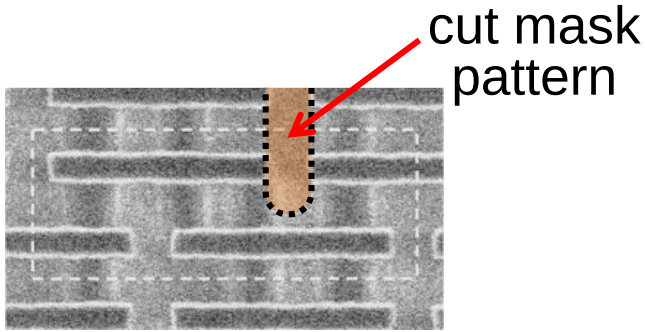
Pitch splitting (LELE)

- Interleave single exposures
- Mask color decomposition & balance
- Extendible to LELELE
- Limited by overlay between masks



Orthogonal cutting

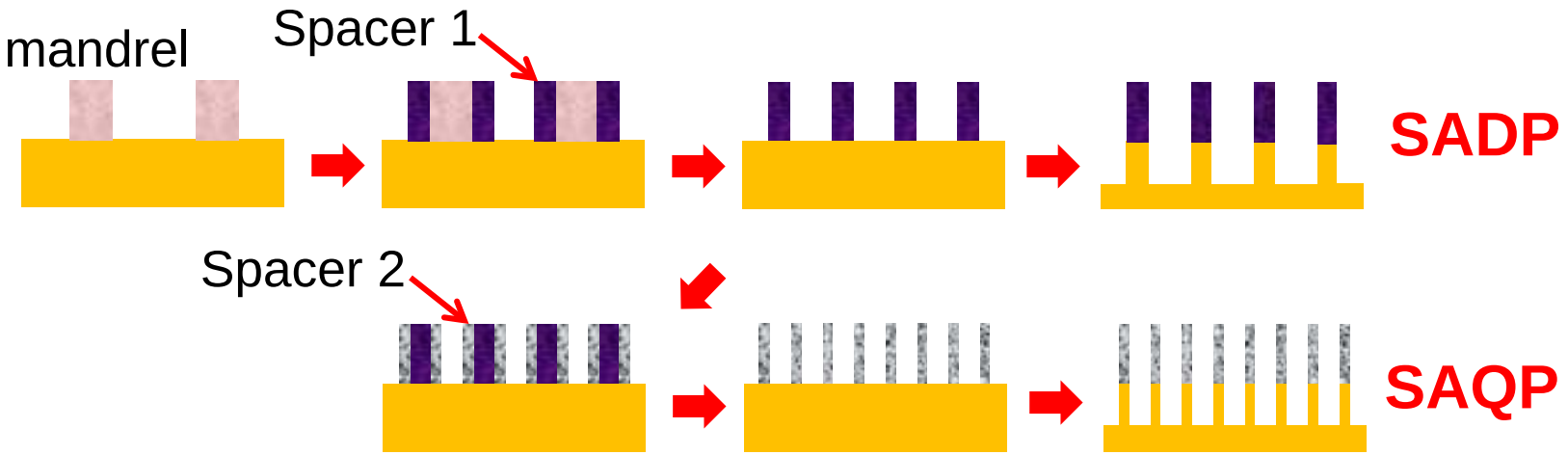
- Extra mask(s) to break line patterns
- Reduced end-to-end spacing
- Limited by overlay in very tight pitch



Arnold *et al.*, ASML [11]
Auth *et al.*, Intel [12]

Spacer-Based Patterning

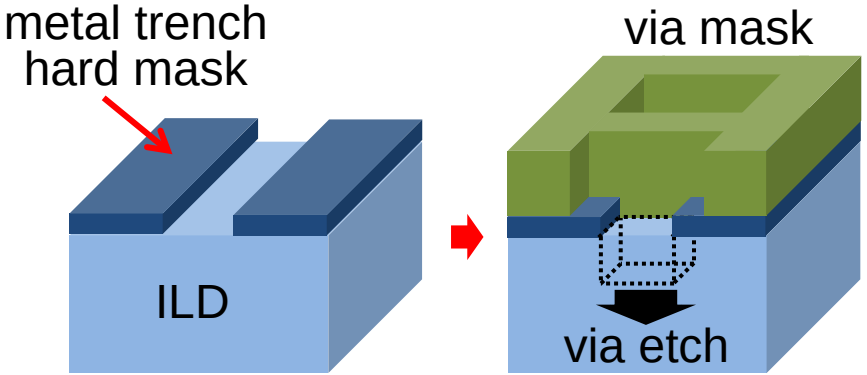
- Pattern fins; now common for gate, MEOL & lower BEOL
- Conformal spacer → correlated LER → less width variation
- Mandrel & spacer width control critical to minimize pitch walking
- Only one feature width, but can be trimmed with extra mask



Self-Aligned Metal Patterning

Self-aligned via

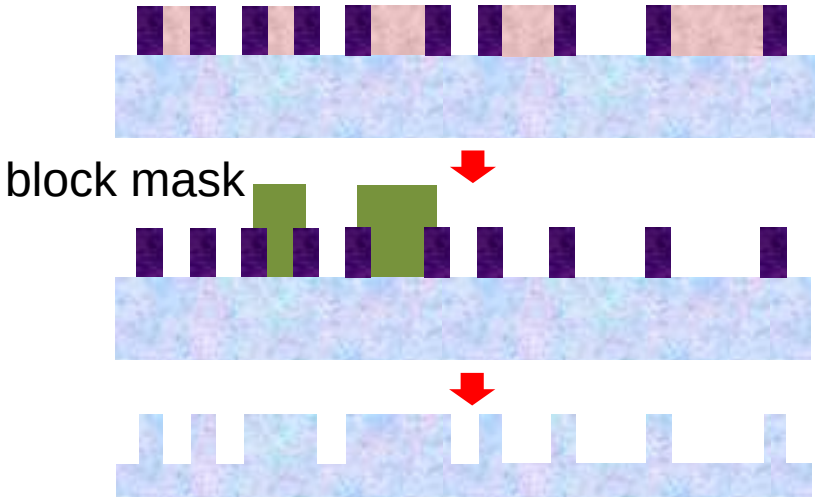
- Dual-damascene metal integration
- Conceptually similar to cuts
- Via etch only at intersection of trench & via masks



Brain *et al.*, Intel [16]

SADP/SAQP + block mask

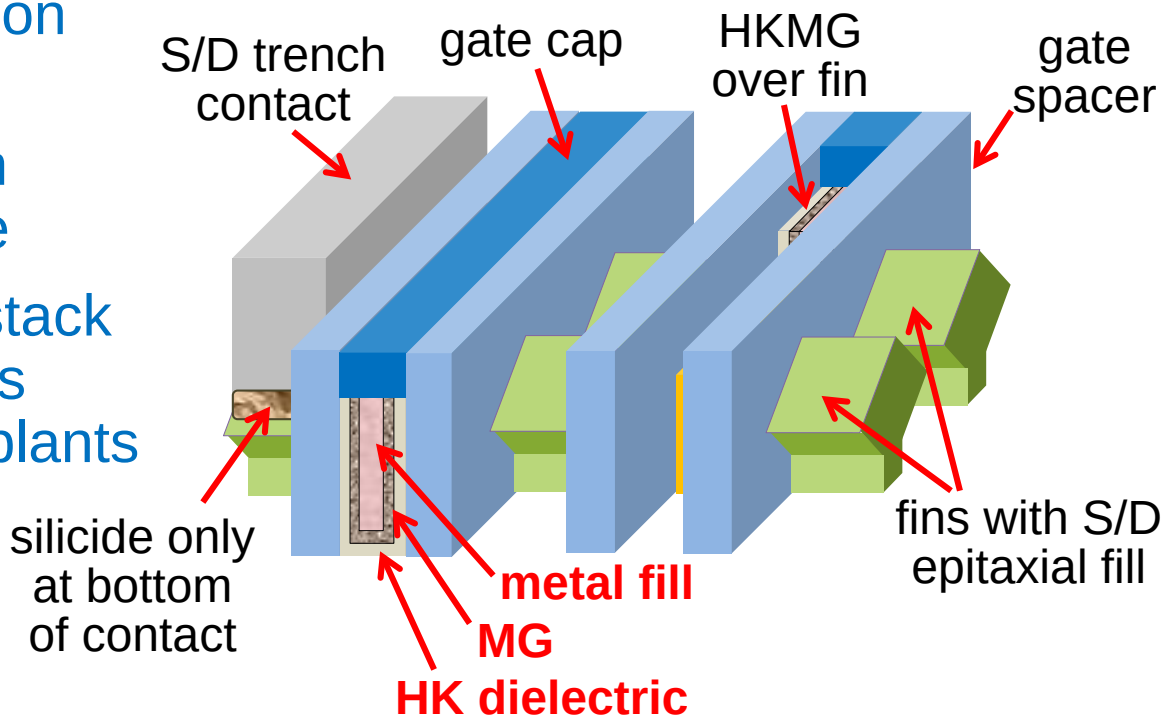
- Block mask to bridge several spacers
- Adjust mandrel width/space for more flexible metal width/space



Woo *et al.*, Globalfoundries [15]

High-*K* + Metal-Gate (HKMG) → Higher C_{ox}

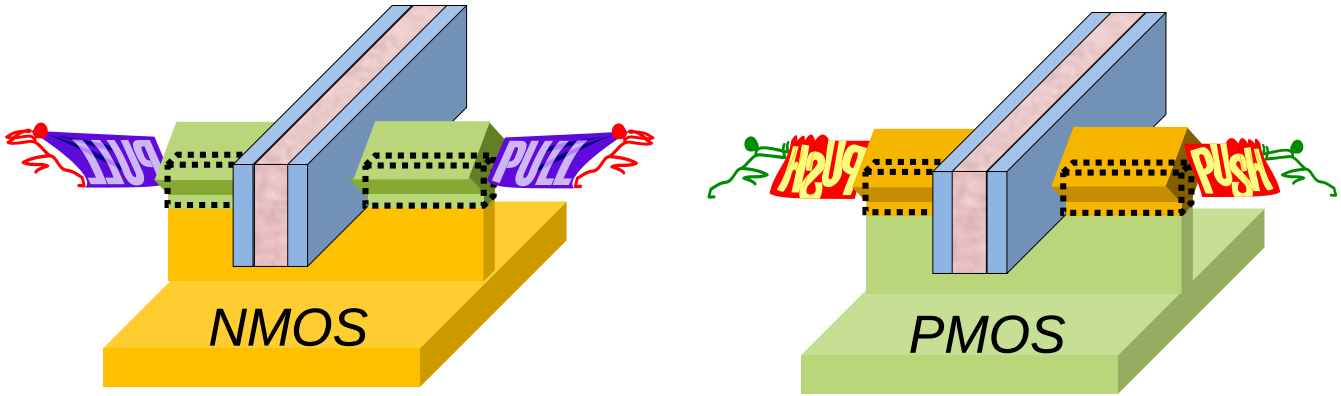
- Less I_{gate}, no poly depletion
- Replacement metal gate (RMG) for stable V_T with delicate HK/MG interface
- V_T tuning with ALD MG stack composition & HK dipoles → less variation than implants
- High gate resistance
- High S/D resistance with silicide last



Hou, TSMC [9]
Auth *et al.*, Intel [14]

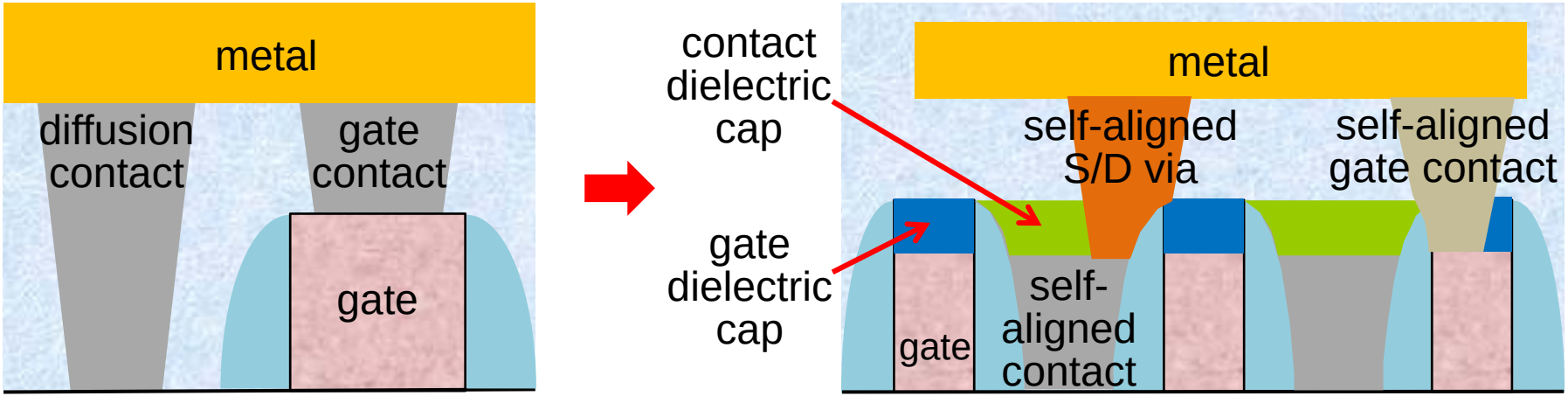
Mechanical Stressors → Mobility Boost

- Induce channel strain along *L* with surrounding stressors
  - Tensile for NMOS, compressive for PMOS (but reality very complicated)
  - Techniques: S/D fin recess & epitaxy, gate stress
- More effective for PMOS, *β* ratio → 1, not scaling well with CGP
- Less effective for longer *L*



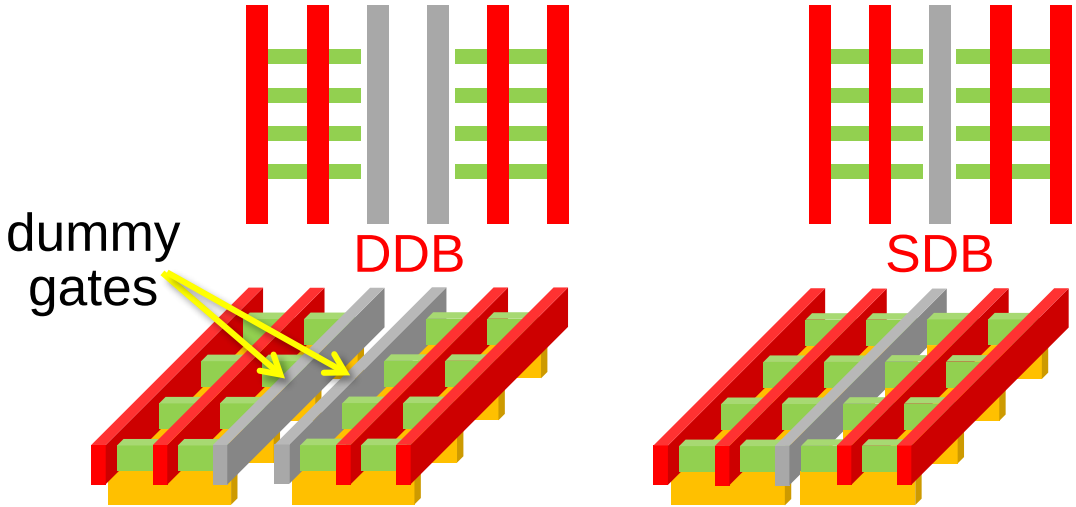
Complex MEOL & Self-Aligned Contacts

- Tight CGP → tough to land diffusion & gate contacts without shorts
- Dielectric caps protect gate & contact against etch
- Self-aligned gate contacts over fins, not restricted to gate overhang
- More interfaces → high S/D, MEOL & lower BEOL resistance



Single vs. Double Diffusion Break

- Dummy gates terminate OD to constrain S/D epitaxy
- SDB eliminates dummy gate waste → saves 10–20% logic area
- Aggressive tensile dielectric isolation for SDB → stress LDE

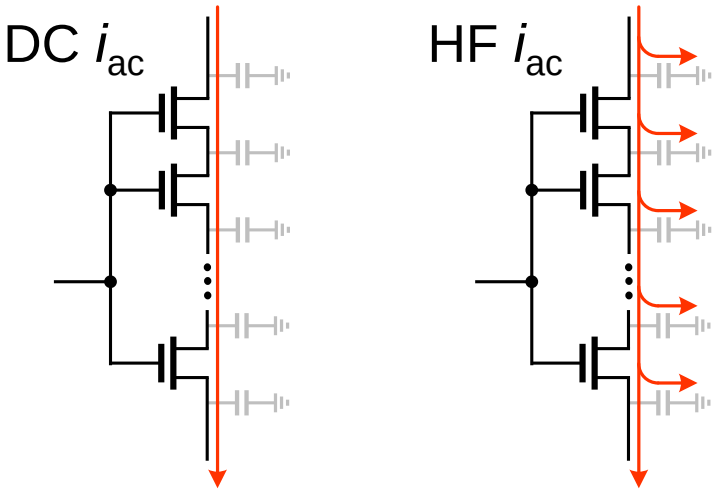


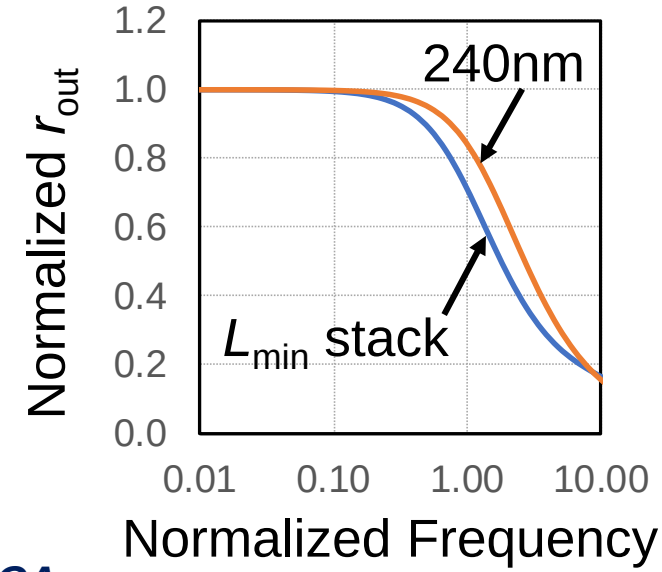
Outline

- Introduction
- Technology Scaling Enablers
- **AMS Device Palette**
  - **Transistors**
  - **Passives**
  - **Diodes**
- AMS Design Impact
- Concurrent Technology/Design Development
- Conclusion

Stacked FET for Higher r_{out}

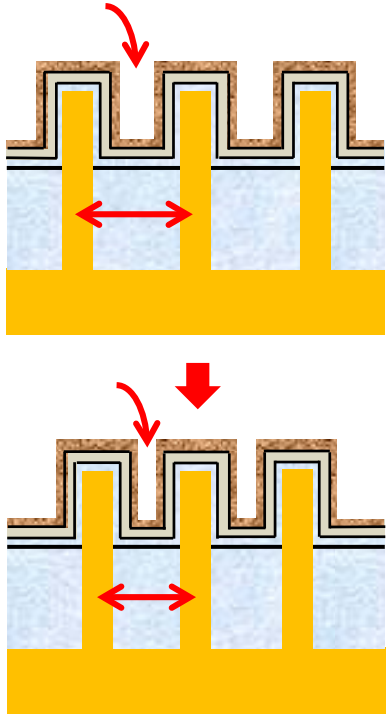
- L_{max} limited by gate litho/etch loading & HKMG CMP
- Short L has most μ boost → potentially less area
- Intermediate diffusions degrade HF r_{out} (gain, CMRR, …)





Thick-Oxide I/O FET

- GPIOs still use 1.8V swing despite reduced core V_{DD}
  - Talk to peripheral ICs made in lower cost nodes
- Challenging to keep 1.8V I/O devices
  - Tighter fin pitch → tough ALD gate fill
  - Complex level shifters to handle larger ΔV_{DD}
  - Many links stopped supporting legacy modes to enable higher data rate & lower power
  - Improve power & area with thinner I/O oxide, e.g., 1.2V

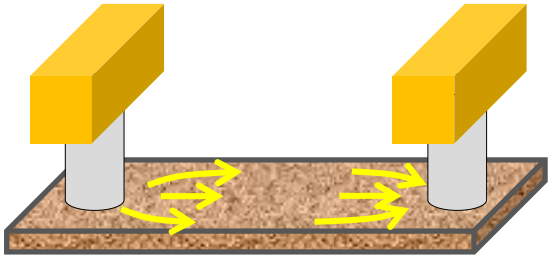


tighter fin pitch

Passives (*RCL*)

Resistors

- MEOL thin film resistor
  - HKMG → poly resistor obsolete
  - Variation limits area scaling
- Gate resistor – unusable, high variation

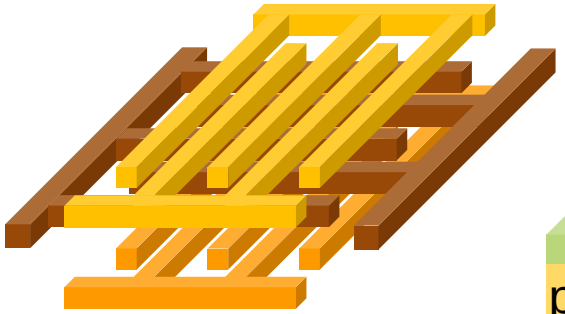


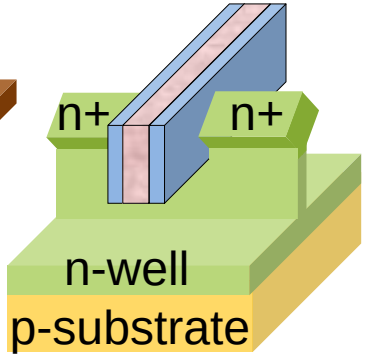
Inductors

- Upper BEOL layers
- Small impact from scaling, but more fill

Capacitors

- BEOL MOM (Metal-Oxide-Metal)
  - High density with metal pitch scaling
  - Reduced AC coupling efficiency
- Accumulation-mode varactor
  - Steeper *C-V* transition
- Upper-BEOL MIM – uncommon in mobile

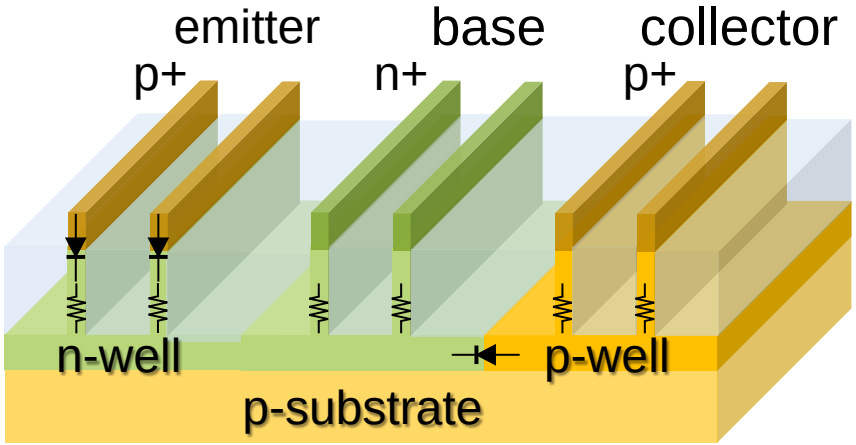




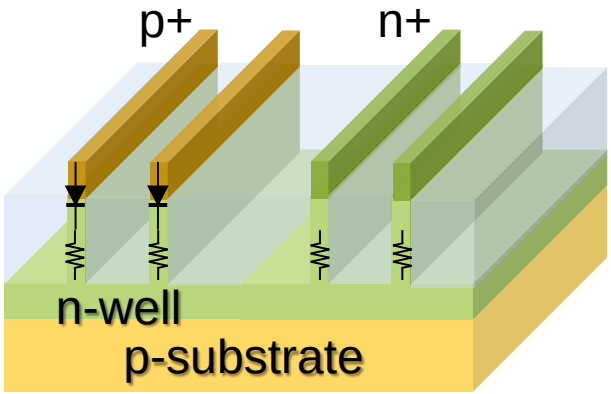
Diodes

- High R_{well} → stricter well tie, guard ring & latch-up rules
- ESD & latch-up guidelines immature during technology development

PNP-BJT (Analog Diode)



STI ESD Diode

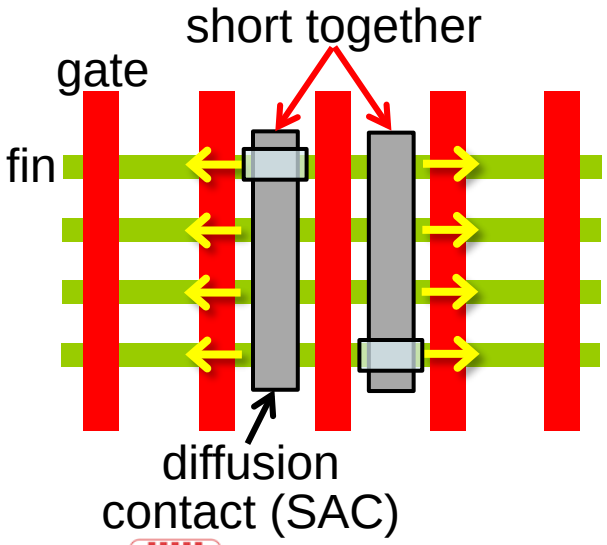


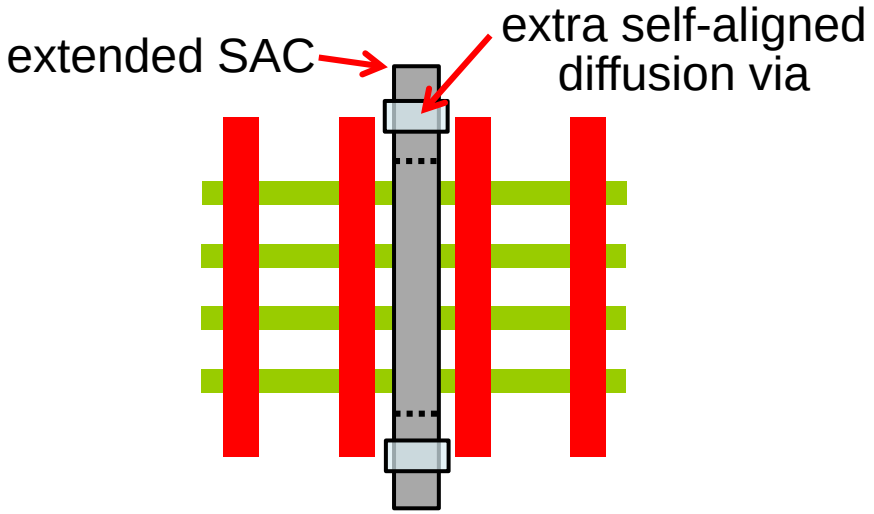
Outline

- Introduction
- Technology Scaling Enablers
- AMS Device Palette
- **AMS Design Impact**
  - Parasitics
  - Layout-Dependent Effects (LDEs)
  - Layout Considerations
- Concurrent Technology/Design Development
- Conclusion

Diffusion & MEOL Resistance

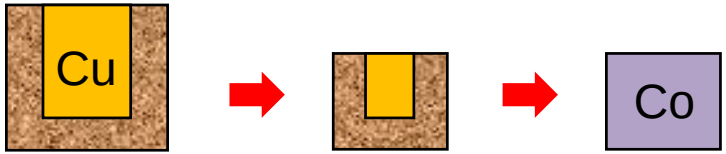
- Challenging for high-current circuits, e.g., I/O drivers, clock buffers
- Double-source layout halves S/D $R_{contact}$ (despite higher C)
- Extend SAC to land extra diffusion via



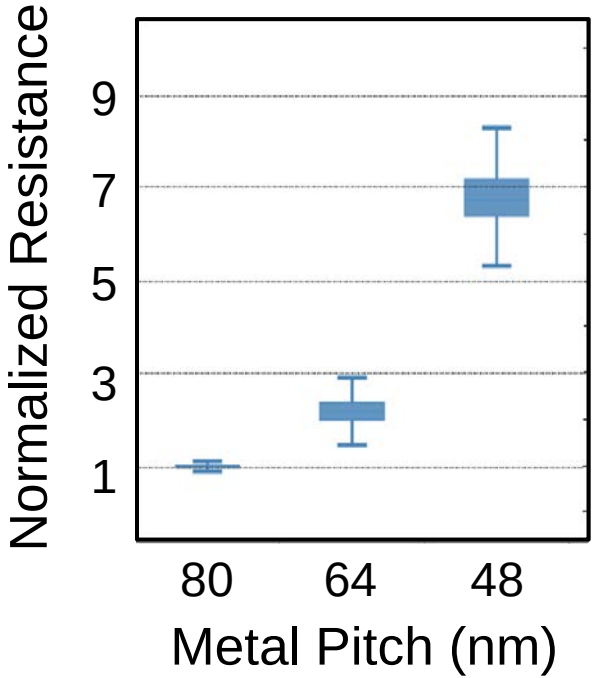


BEOL Resistance

- Aggressive M1-M3 pitch scaling for dense logic routing → less die area & cost
- Meticulous pitch optimization for PPAC
- High-ρ TaN barrier cladding Cu wire not scaling with metal pitch

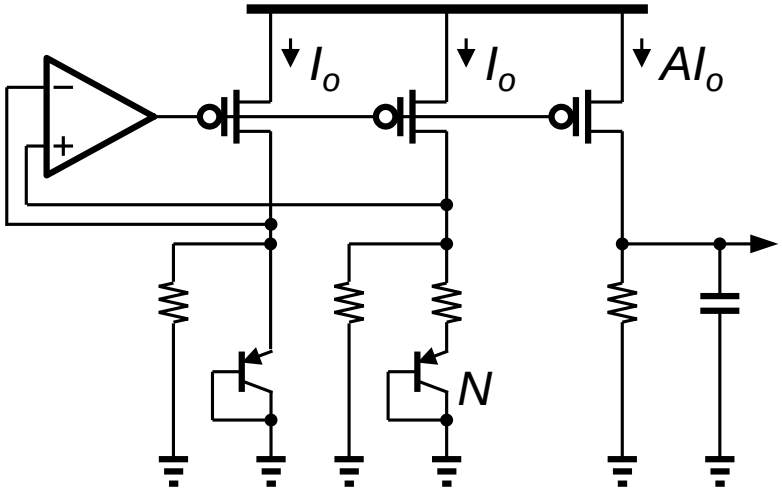


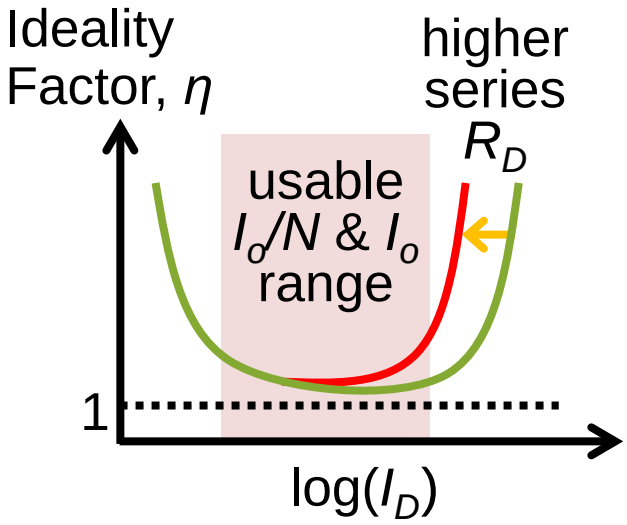
- Barrier-less cobalt & ruthenium with higher ρ are promising material enablers



Auth *et al.*, Intel [14]
Yang *et al.*, Qualcomm [22]

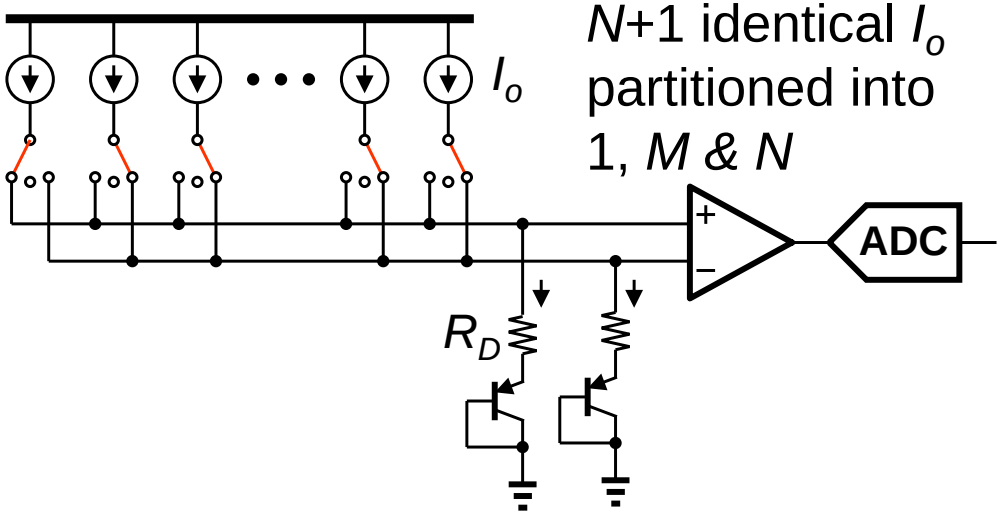
Low-Voltage Bandgap Reference





- Higher R_D → smaller N → variation sensitivity
- Higher V_D from high well doping → higher V_{DD} (e.g., 1.2V) for headroom
- Variation dominated by PMOS mirror mismatch → trimming

Thermal Sensor with R_D Cancellation



$$\Delta V_{BE,M} = \frac{\eta kT}{q} \ln M + (M\text{-}1)\, I_o R_D$$

$$\Delta V_{BE,N} = \frac{\eta kT}{q} \ln N + (N\text{-}1)\, I_o R_D$$

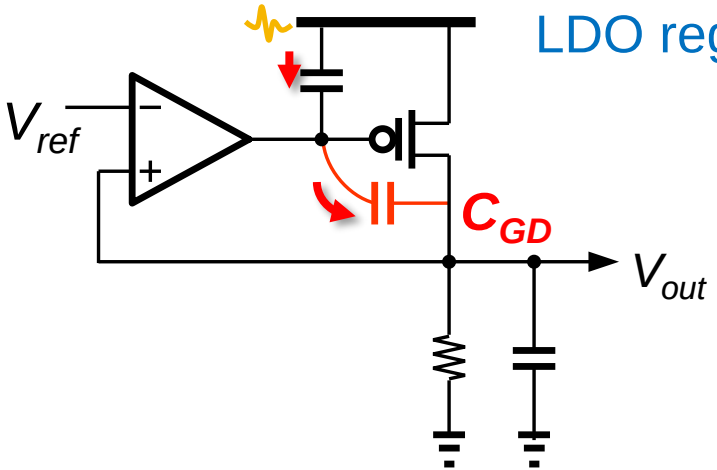
$$\frac{\eta kT}{q} = \frac{(N\text{-}1)\Delta V_{BE,M} - (M\text{-}1)\Delta V_{BE,N}}{(N\text{-}1)\ln M - (M\text{-}1)\ln N}$$

- Measure ΔV_{BE} at M:1 & N:1 → cancel R_D
- DEM → cancel I_o mismatch
- Swap amp inputs → cancel diode mismatch

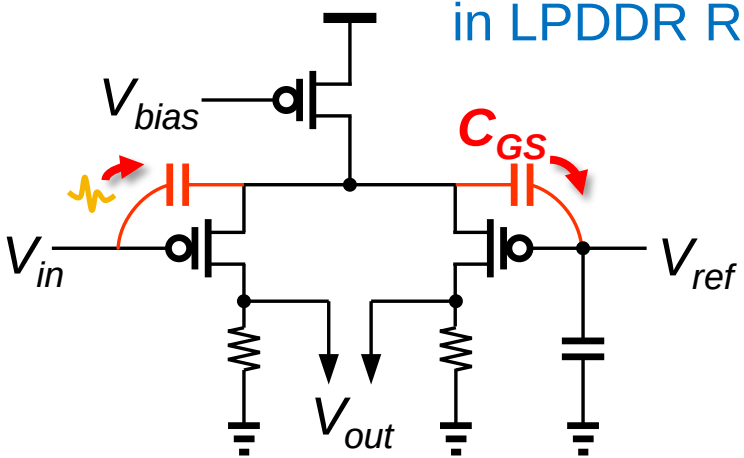
Parasitic Capacitance Impact on Analog

- S/D trench contacts & gate form vertical plate capacitors
- Adding capacitance increases area & wake-up time (burst-mode)

Worse PSRR in LDO regulator

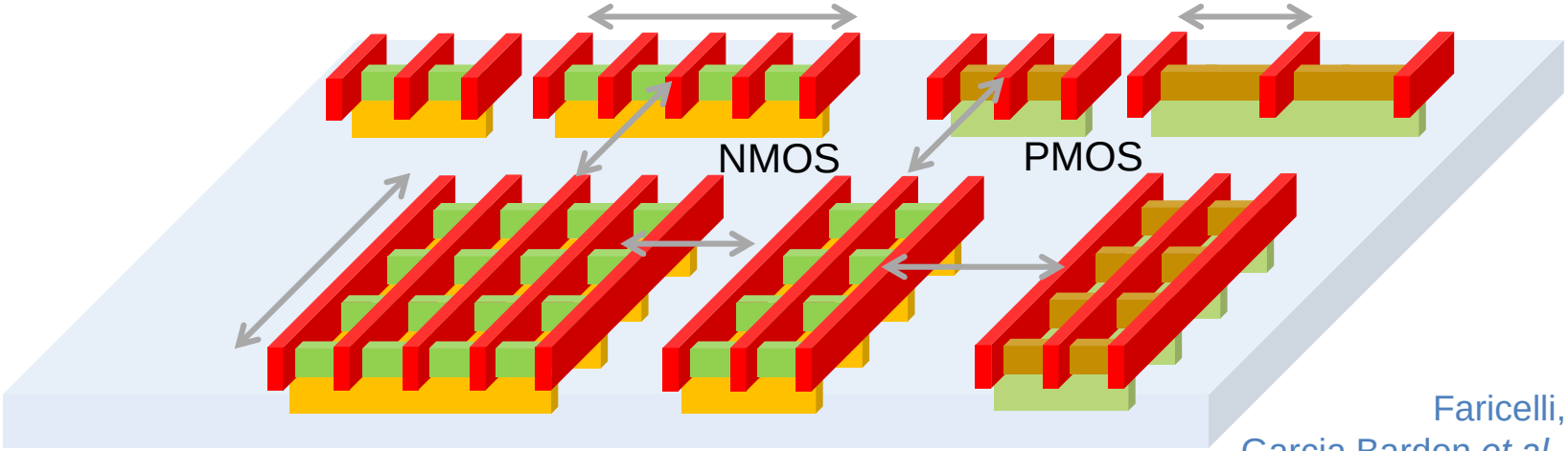


Kickback noise in LPDDR RX



Stress LDEs

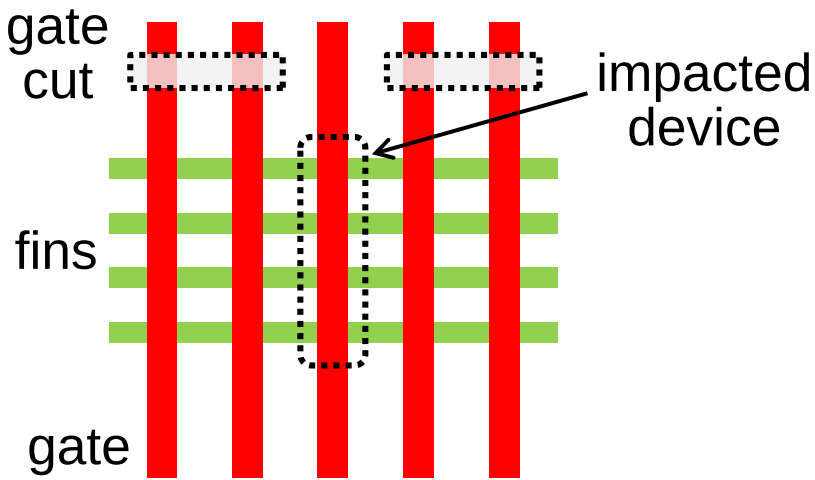
- Stronger stressors & layout effects → schematic/layout Δ
- Stress build-up in longer OD, I_D per fin not constant vs. # fins
- Interaction with surrounding tensile STI & ILD stress
- NMOS/PMOS stress mutually weaken each other

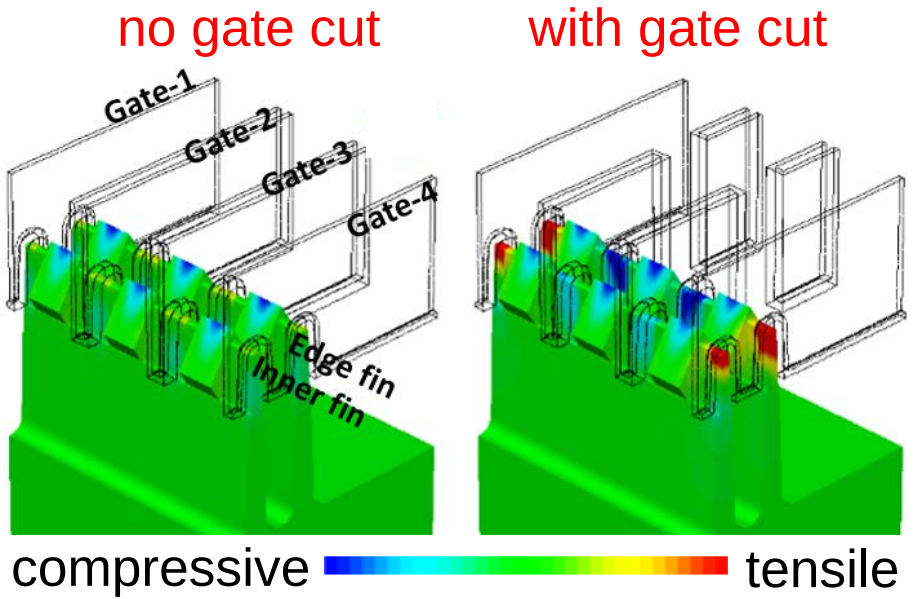


Faricelli, AMD [29]
Garcia Bardon *et al.*, imec [30]
Bianchi *et al.*, STMicroelectronics [31]

Gate Cut Stress LDE

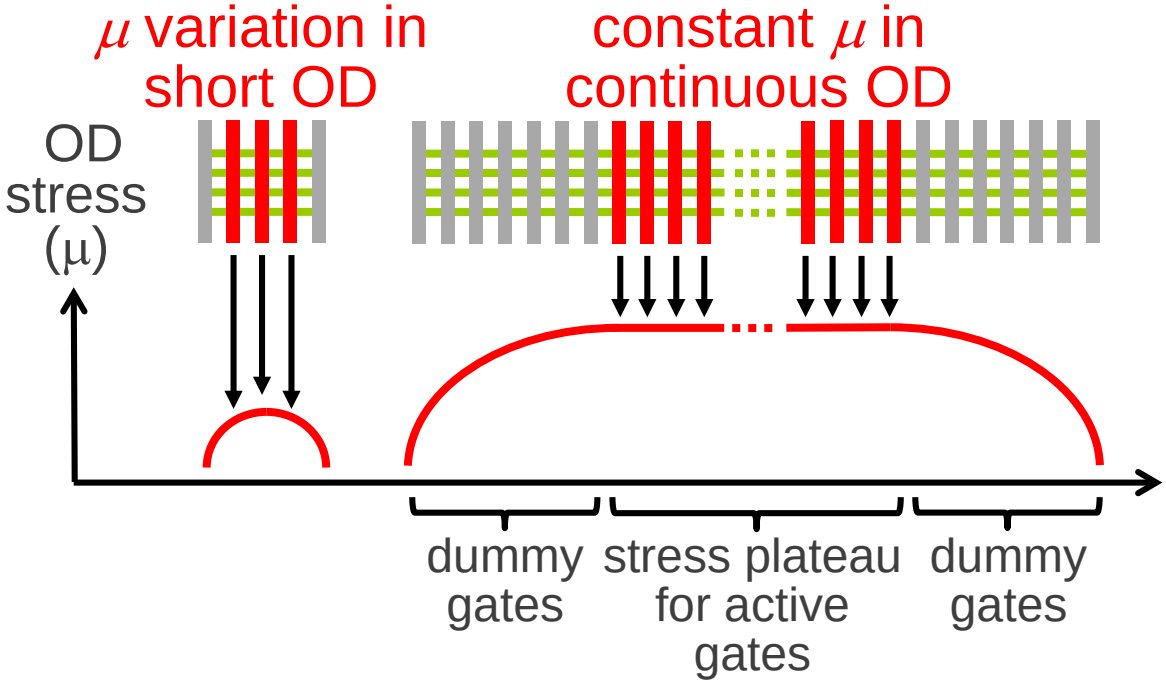
- Gate cut disrupts mechanical support of continuous gate
- Modulate stress near cut → $\Delta\mu$ & ΔV_T, modeled in post-layout netlist





Continuous OD for Performance & Matching

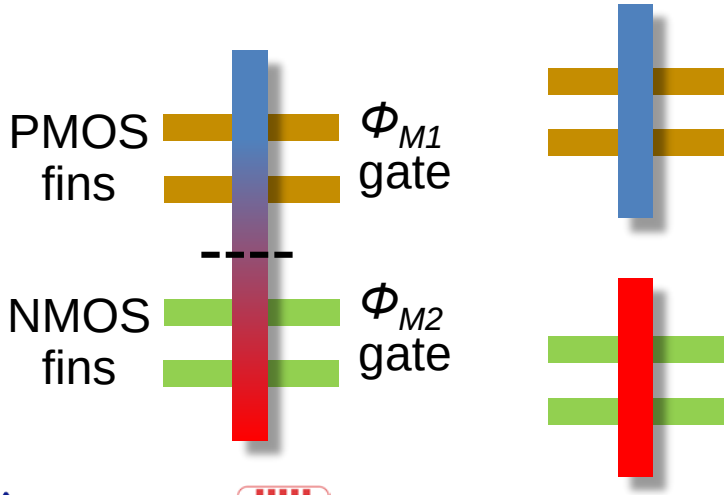
- Build up stress plateau for higher μ
- Desensitize FET from μ variation in short OD
- Most critical for short *L* with strongest LDE
- Matched FETs also need matched spacing to surrounding devices



HKMG LDEs

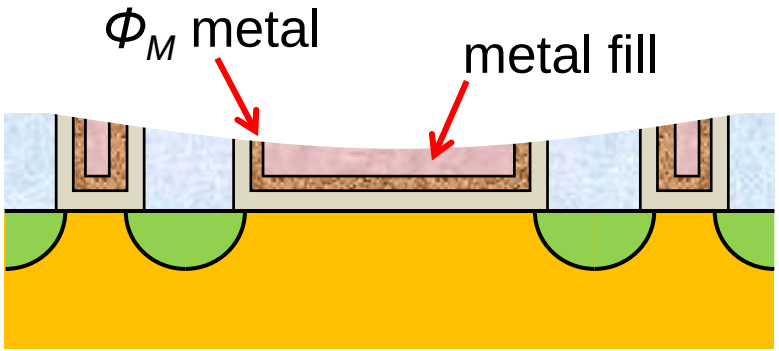
Metal Boundary Effect

- ΔV_T near border of different Φ_M
- Interdiffusion of Φ_M
- Modeled in post-layout netlist



Gate Density Induced Mismatch

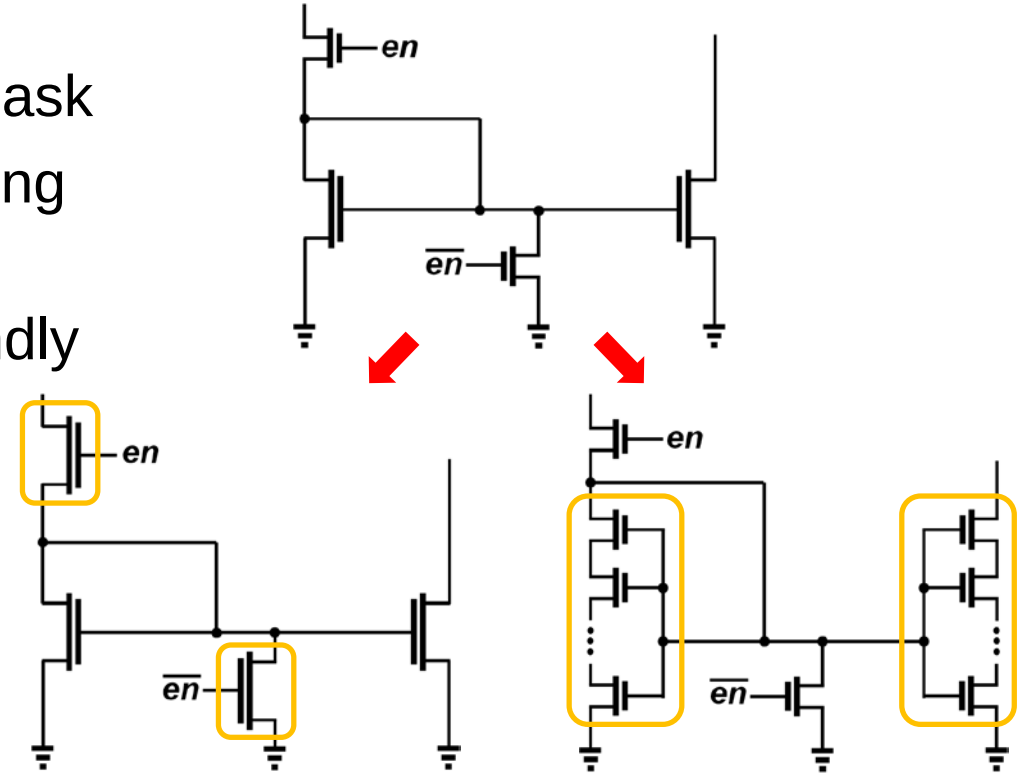
- ΔV_T from RMG CMP dishing
- Φ_M influenced by metal fill & sidewall MG
- Not modeled, contained with DRC



Yang et al, Qualcomm [24]
Hamaguchi *et al*., Toshiba [33]

Current Mirror with Enable Devices

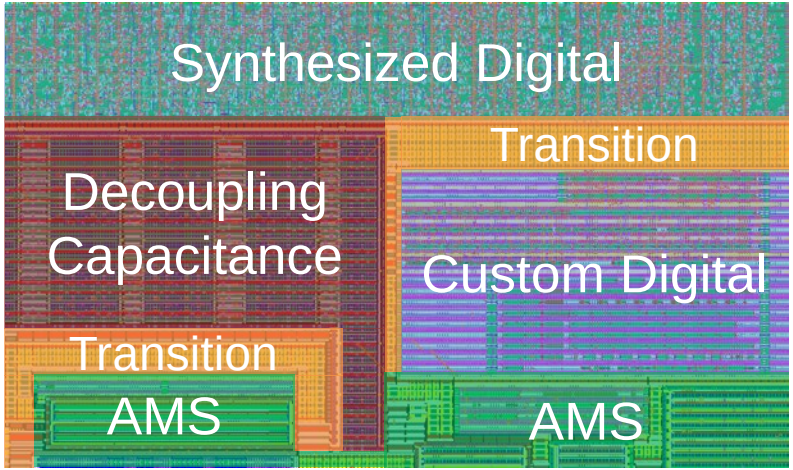
- Short *L* patterned by SADP; long *L* with conventional 2nd mask
- SADP prone to litho/etch loading effects
- Consistent *L* more SADP-friendly
- Avoid mixing short- & long-channel FETs



Layout Density & Floorplan Considerations

- Tougher DRCs → AMS layout resemble logic arrays
- Density checks to reduce long-range pattern variation
  → iterative rework of smaller cells
  - Contacts, vias, cuts, tight-pitch metal
  - Area, perimeter, gradient
  - Larger checking windows
  - Density union of multiple metal levels
- Floorplanning more tedious & bloated
  - More dummy gates, well taps, guard rings
  - Transitions between different device types & pattern densities

ADPLL partial floorplan

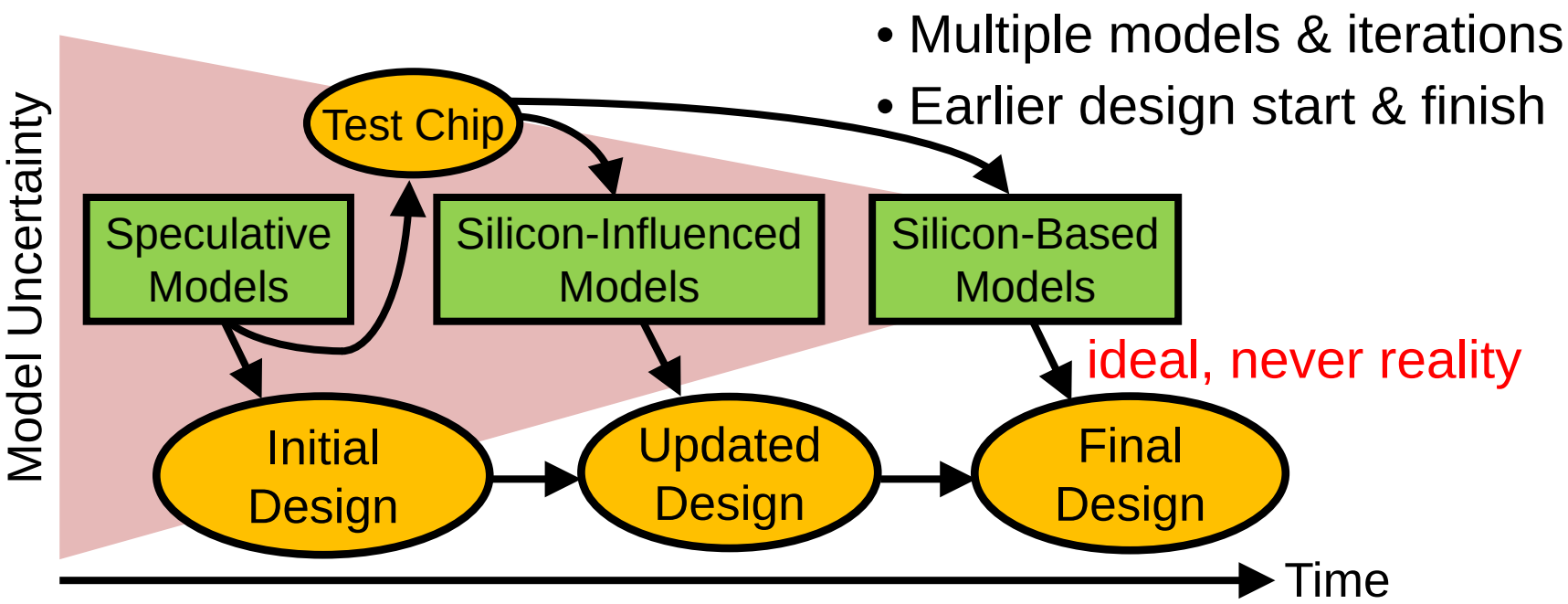


Outline

- Introduction
- Technology Scaling Enablers
- AMS Device Palette
- AMS Design Impact
- **Concurrent Technology/Design Development**
  - Perspectives
  - Dealing with Model Uncertainties
- Conclusion

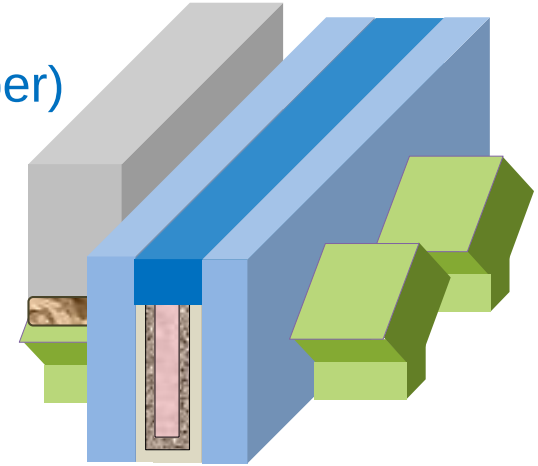
Bleeding-Edge Product Development

- Design while technology being developed to shorten time-to-market



Dealing with Target-Based Model Uncertainty

- Process at tapeout more immature in each new node
  - More masks & longer fab cycle time → fewer cycles of silicon learning
- Process development areas, even after tapeout
  - HKMG stack & RMG optimization to tune multiple V_T
  - S/D epitaxy, MEOL modules (contacts, vias & metal)
  - Logic & SRAM area-saving constructs (SDB, S/D jumper)
- Most vulnerable (unstable) model parameters
  - FET V_T, μ, LDEs
  - Long *L* & I/O FETs usually impacted, not priority #1
  - *RC* parasitics in S/D & MEOL
- Trade incremental AMS area for reduced exposure



Overcoming Process/Model Immaturity

Layout Guideline	Reduced Exposure
Use continuous OD stress plateau	Stress LDEs (S/D epitaxy, STI)
Attach dummy FETs to OD ends	
Avoid single-diffusion break	
Use only one Φ_M in each gate	Metal boundary LDE
Avoid using gate as interconnect	Gate cut LDE
Contact gate on both sides	Gate resistance, Φ_M tuning to adjust V_T
Use groups of fewer fins	
Use double-source layout for high-*I* nets	S/D contact resistance & epitaxy
Extend S/D contact to land extra via	MEOL & S/D resistance
Do not push DRCs to limit	DRC updates

Conclusion

- AMS design in remaining CMOS nodes is tedious & about managing technology-imposed non-idealities
- AMS area scaling 0.8-0.9x per node vs. 0.5x for logic/SRAM
- Parasitics & LDEs only get worse, will ultimately limit scaling
- Digital-friendly AMS design inspires new performance, power & integration levels
- Implementation just requires a lot more perspiration

References (1/3)

[1] L. Bair, “Process/product interactions in a concurrent design environment,” in *IEEE CICC*, 2007.

[2] C. Auth *et al*., “A 22nm high performance and low-power CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors,” in *IEEE Symp. VLSI Technology*, 2012.

[3] S.-Y. Wu *et al*., “A 16nm finFET CMOS technology for mobile SoC and computing applications,” in *IEEE IEDM*, 2013.

[4] H.-J. Cho *et al.*, “Si finFET based 10nm technology with multi Vt gate stack for low power and high performance applications,” in *IEEE Symp. VLSI Technology*, 2016.

[5] S.-Y. Wu *et al*., “A 7nm CMOS platform technology featuring 4th genera-tion finFET transistors with a 0.027um^2 high-density 6-T SRAM cell for mobile SoC applications,” in *IEEE IEDM*, 2016.

[6] P. Packan *et al.*, “High performance 32nm logic technology featuring 2nd generation high-k + metal gate transistors,” in *IEEE IEDM*, 2009.

[7] L. Wei *et al*., “Exploration of device design space to meet circuit speed targeting 22nm and beyond,” in *Int. Conf. SSDM*, 2009.

[8] F.-L. Hsueh *et al*., “Analog/RF wonderland: circuit and technology co-optimization in advanced finFET technology,” in *IEEE Symp. VLSI Technology*, 2016.

[9] C. Hou, “A smart design paradigm for smart chips,” in *IEEE ISSCC*, 2017.

[10] D. Yang, “SoC scaling challenges in the era of the single digit technology nodes,” in *Int. Workshop Advanced Patterning Solutions*, 2017.

[11] W. Arnold *et al.*, “Manufacturing challenges in double patterning lithography,” in *IEEE ISSM*, 2006.

References (2/3)

[12] C. Auth *et al.*, “45nm high-k + metal-gate strain-enhanced transistors,” in *IEEE Symp. VLSI Technology*, 2008.

[13] Y.-K. Choi, T.-J. King, and C. Hu, “A spacer patterning technology for nanoscale CMOS,” *IEEE Trans. Electron Devices*, vol. 49, no. 3, 2002.

[14] C. Auth *et al.*, “A 10nm high performance and low-power CMOS technology featuring 3rd generation finFET transistors, self-aligned quad patterning, contact over active gate and cobalt local interconnects,” in *IEEE IEDM*, 2017.

[15] Y. Woo *et al*., “Design and process technology co-optimization with SADP BEOL in sub-10nm SRAM bitcell,” in *IEEE IEDM*, 2015.

[16] R. Brain *et al*., “Low-k interconnect stack with a novel self-aligned via patterning process for 32nm high volume manufacturing,” in *IEEE IITC*, 2009.

[17] S. H. Yang *et al*., “28nm metal-gate high-K CMOS SoC technology for high-performance mobile applications,” in *IEEE CICC*, 2011.

[18] C. Y. Kang *et al.*, “The impact of La-doping on the reliability of low Vth high-k/metal gate nMOSFETs under various gate stress conditions," in *IEEE IEDM*, CA, 2008.

[19] L. Chang *et al*., “Gate length scaling and threshold voltage control of double-gate MOSFETs,” in *IEEE IEDM*, 2000.

[20] V. Chan *et al*., “Strain for CMOS performance improvement,” in *IEEE CICC*, 2005.

[21] M. Rashed *et al.*, “Innovations in special constructs for standard cell libraries in sub 28nm technologies,” in *IEEE IEDM*, 2013.

References (3/3)

[22] S. Yang *et al.*, “10 nm high performance mobile SoC design and technology co-developed for performance, power, and area scaling,” in *IEEE Symp. VLSI Technology*, 2017.

[23] F.-L. Hsueh, “Device challenges for scaled analog-RF,” in *IEEE Symp. VLSI Technology*, Short Course, 2017.

[24] S. Yang *et al.*, “High-performance mobile SoC design and technology co-optimization to mitigate high-K metal gate process variations,” in *IEEE Symp. VLSI Technology*, 2014.

[25] A. Wei *et al.*, “Challenges of analog and I/O scaling in 10nm SoC technology and beyond,” in *IEEE IEDM*, 2014.

[26] H. Banba *et al.*, “A CMOS bandgap reference circuit with sub-1-V operation,” *IEEE J. Solid-State Circuits*, vol. 34, no. 5, 1999.

[27] “ADT7461 ±1°C temperature monitor with series resistance cancellation,” *ON Semiconductor Pub. No. ADT7461/D*, 2014.

[28] E. Terzioglu, “Design and technology co-optimization for mobile SoCs,” in *IEEE ICICDT*, Keynote, 2015.

[29] J. Faricelli, “Layout-dependent proximity effects in deep nanoscale CMOS,” in *IEEE CICC*, 2010.

[30] M. Garcia Bardon *et al.*, “Layout-induced stress effects in 14nm & 10nm finFETs and their impact on performance,” in *IEEE Symp. VLSI Technology*, 2013.

[31] R.A. Bianchi *et al.*, “Accurate modeling of trench isolation induced mechanical stress effects on MOSFET electrical performance,” in *IEEE IEDM*, 2002.

[32] X. Xi *et al.*, BSIM4.3.0 MOSFET Model User’s Manual, *Regents of Univ. California at Berkeley*, 2003.

[33] M. Hamaguchi *et al.*, “New layout dependency in high-K/metal gate MOSFETs,” in *IEEE IEDM*, 2011.