Devices, Technology & Applications: A Critique of New Proposals

Sandip Tiwari
st222@cornell.edu

SWOT

Judgments with incomplete information, incomplete model, and incomplete knowledge of environment of the later date!

System 1 and System 2.

Acknowledgements: Collaborators, students, peers, funders (NSF, DARPA, Mellowes Endowment, Hosts), ...

What makes a good digital device?

Large # of applications

General purpose

Deterministic, aligned with current practices

Logic switch

Massive integration
High function/Power
Inexpensive technology
…

e.g., CMOS

Memory

Density
Ultra low power
Speed⇔ function
Embedded/Stand-alone
Inexpensive technology
…
Non-volatility
Sensing/Self-sensing

e.g., sRAM, dRAM, Flash

Small #, but substantive applications

Special purpose

Compelling application where economic constraints can be broken.

Unique attribute!
(non-volatile, soft-error immunity, technology compatibiity, multi-function, …)

e.g., JJ's, SiGe bipolar, BiCMOS, TFT

What makes a good Analog/RF device?

General purpose

Across the electromagnetic spectrum: Hertz to PetaHertz

With digital

- Technology compatibility
- Frequency
- Linearity
- Noise
- Efficiency
- …

e.g., BiCMOS, SiGe Bipolar

App's in integrated processing: Converters, DSPs, LIDARs, PARs, Software radio, …

Stand-alone

Small signal

- Linearity
- Frequency
- Gain
- Noise, sp. Φ
- …

e.g., SiGe Bipolar, LDD MOSFET

Large signal

- Power-frequency
- Non-linearities
- Gain
- Efficiency
- …

e.g., GaN, SiC, GaInAs IIIV HFETs & HBTs, LDD and HV MOSFET

App's: LNA, Power Amps, Active filters, Mixers, Displays, Power, HF Radars, Software radios

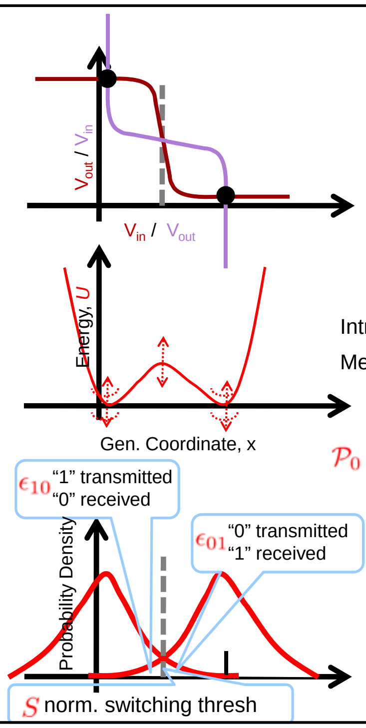
Special purpose

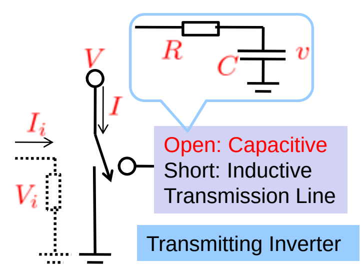
Unique attribute!
(specific frequencies, efficiency, compatibility, …)

e.g., Piezo Converters, CCD, Pixel PD CMOS, THz sources, Organic LED, Plasmonic Antennas, Mixers,

App's: Cameras, Comm. Sensor networks, Reconnaissance, Medical, …

A Switch





Intrinsic minimum energy per logic operation: $U_i = P_i t_i$

Mean thermal energy generated within ckt: $\frac{1}{2}k_B T = \frac{1}{2}C\overline{v}^2$

For "0" and "1" states, with random fluctuations,

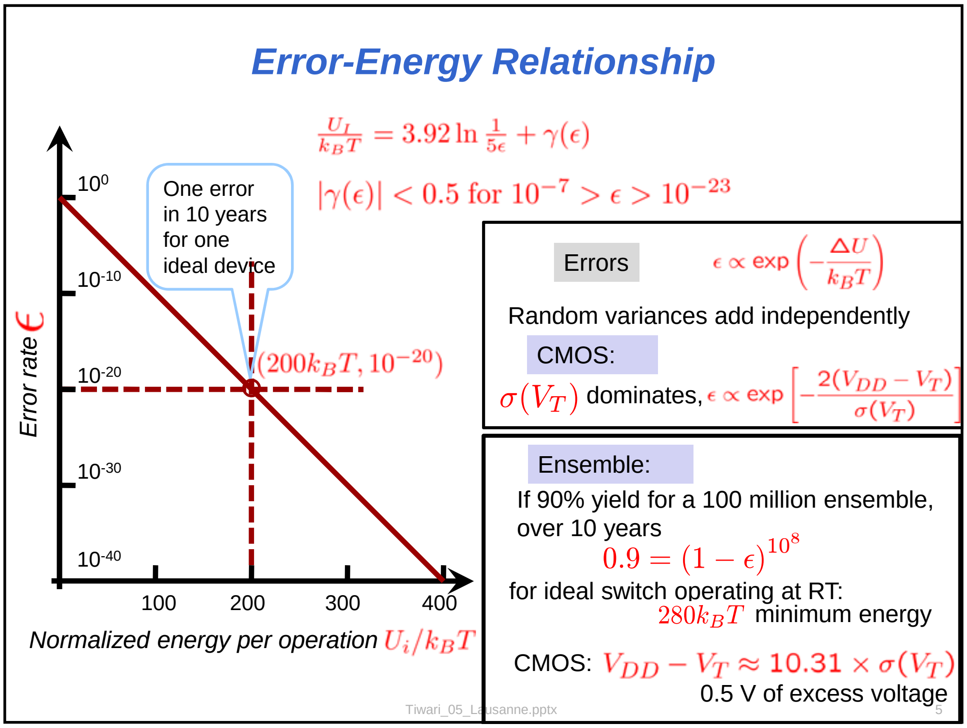
$$\mathcal{P}_0 = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{v^2 C}{2k_B T}\right) \quad \mathcal{P}_1 = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(v-V)^2 C}{2k_B T}\right)$$

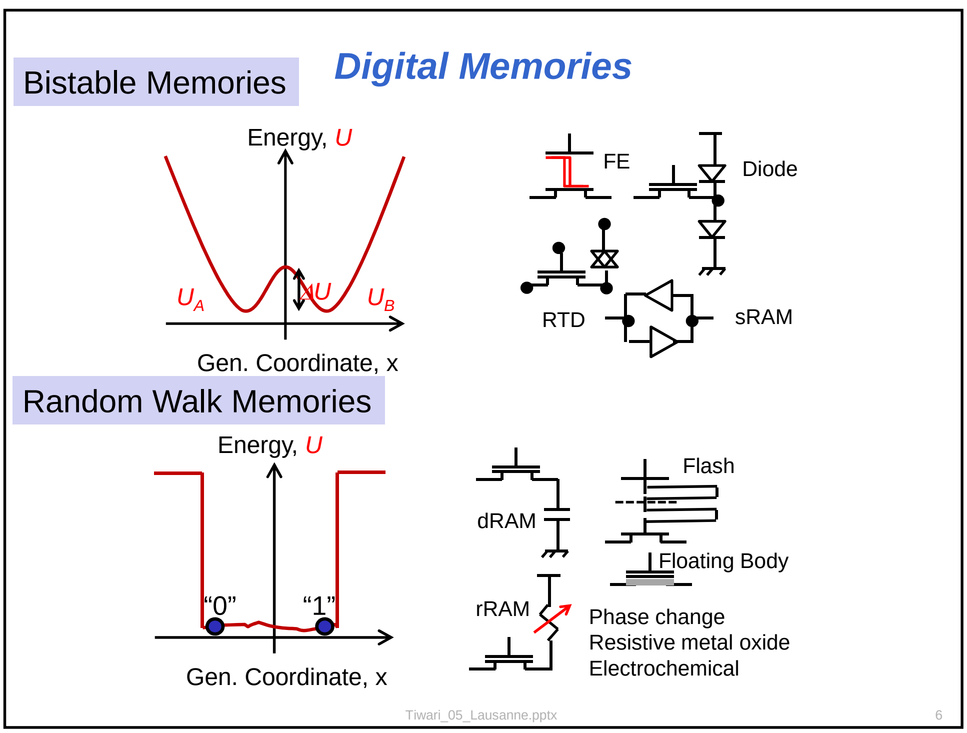
$$S = \frac{V}{2\sqrt{k_B T/C}}$$

Error rate: $\epsilon_{01} + \epsilon_{10} = 2\epsilon$

Total Error Rate:

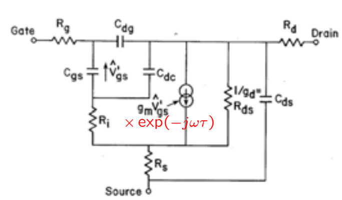
$$\epsilon = \frac{2}{\sqrt{2\pi}}\int_S^\infty \exp\left(-\frac{v^2 C}{2k_B T}\right) d\left(\frac{v}{\sqrt{2k_B T/C}}\right)$$

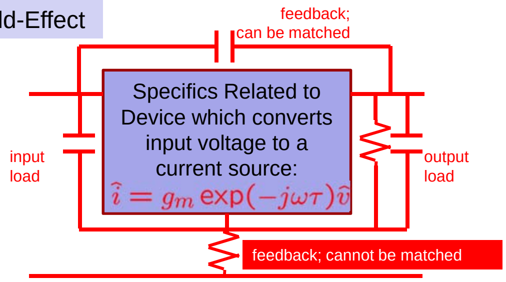

Error-Energy Relationship
$\frac{U_I}{k_BT} = 3.92 \ln \frac{1}{5\epsilon} + \gamma(\epsilon)$
$|\gamma(\epsilon)| < 0.5$ for $10^{-7} > \epsilon > 10^{-23}$
One error in 10 years for one ideal device
$(200k_BT, 10^{-20})$
Error rate ϵ
10^0
10^{-10}
10^{-20}
10^{-30}
10^{-40}
100
200
300
400
Normalized energy per operation U_i/k_BT
Errors
$\epsilon \propto \exp\left(-\frac{\Delta U}{k_BT}\right)$
Random variances add independently
CMOS:
$\sigma(V_T)$ dominates, $\epsilon \propto \exp\left[-\frac{2(V_{DD}-V_T)}{\sigma(V_T)}\right]$
Ensemble:
If 90% yield for a 100 million ensemble, over 10 years
$0.9 = (1-\epsilon)^{10^8}$
for ideal switch operating at RT:
$280k_BT$ minimum energy
CMOS: $V_{DD} - V_T \approx 10.31 \times \sigma(V_T)$
0.5 V of excess voltage
Tiwari_05_Lausanne.pptx
5


Digital Memories
Bistable Memories
Energy, U
U_A
ΔU
U_B
Gen. Coordinate, x
FE
Diode
RTD
sRAM
Random Walk Memories
Energy, U
"0"
"1"
Gen. Coordinate, x
dRAM
Flash
Floating Body
rRAM
Phase change
Resistive metal oxide
Electrochemical
Tiwari_05_Lausanne.pptx
6

RF/Analog; Voltage Driven Current Source

Field-Effect





For really high frequencies, $g_m/2\pi C$ is inadequate as a measure

Frequency for unity current gain

Frequency for unity unilateral gain (Power)

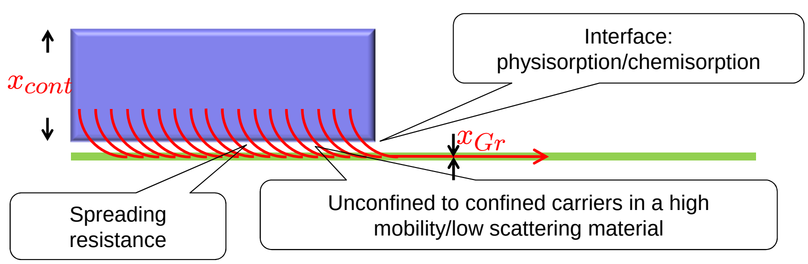
$$f_T = \frac{g_m/2\pi(C_{gs}+C_{dg})}{\left(1+\frac{R_d+R_s}{R_{ds}}\right)+C_{dg}g_m(R_d+R_s)}$$

$$f_{max} = \frac{g_m/2\pi(C_{gs}+C_{dg})}{\sqrt{4\frac{R_g+R_s+R_i}{R_{ds}}\left[1+4\pi f_T R_{ds} C_{gs}\left(1+\frac{R_s}{R_g+R_s+R_i}+\frac{2\pi\tau}{(R_g+R_s+R_i)C_{gs}}\right)\right]}}$$

High output resistance important!

Low delay factors, RC constants (parasitics), transit delays important!

Contacts!



Surface potential in contact

$$\phi_s = \frac{E_j x_{cont}}{\pi}\left[\ln\frac{t^2-1}{t^{\star 2}-1}+\ln\frac{1-s^2t^{\star 2}}{1-s^2t^2}\right] \qquad R_{c,sprd}W = \frac{2}{\pi}R_{\square,cont}x_{cont}\ln\left(0.75\frac{x_{cont}}{x_{Gr}}\right)-\frac{0.21}{\pi}R_{\square,Gr}x_{Gr}$$

Solution matches expectations of Si inversion layers

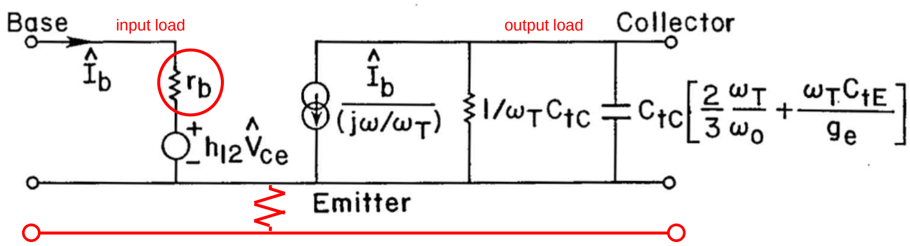
Graphene	Edge: R_c (Ω.µm)	Areal*: ρ_c (Ω.cm^2)
Nagashio et al. (U. Tokyo)	~10^3-10^4	10^{-4} – 10^{-3}
Xia et al. (IBM)	~100-200	10^{-6}
Silicon	~1-10	Mid-10^{-8}

* estimated

In Graphene this is a considerably more complicated problem.
Spreading resistance provides a lower bound.

RF/Analog; Current Driven Current Source

Charge Density/Current Density Effect



a simple model:

$$f_T = \frac{1}{2\pi\left[\tau_B + \tau_c' + (C_{tE} + C_{tC})/g_e\right]}$$

$$f_{max} = \sqrt{\left(\frac{f_T}{8\pi r_b C_{tC}}\right)}$$

+ emitter resistance's feedback

This is the equivalent of $g_m/2\pi C$ field-effect, but now time constants of charging and current flow

Input resistance losses important!

Note in all this, the current gain is a very poor measure of the device for high frequency. Power gain is; device is useful for rf if it can convert dc to ac.

Recap

Digital:

- Noise
- Variance
- Defects
- Energy
- Compactness

Analog:

- Signal transit delays
- Charging/discharging delays
- Noise
- Linearity
- Power
- Gain
- Parasitics

MEMS has its own additional set arising from 3D and mechanical properties and their interactions with others

SWOT

	Strength	Weakness	Opportunity	Threat
Graphene	Electrostatics, ideal mobility	Charge outside graphene, Resistance to dimensionality, No BGap	3D integration, Thz nonlinearities, Broad λ photoabsorption	Contact R, Reproducibility
Nanotubes	Electrostatics, high mobility	Variable bandgap, metal & semiconductor	Conductiivty, mechanical strength	Contact R, variability, no substrate
Tunnel transistor	Only works at small spacing	Tunneling is size dependent, high C	High current	Variability, Contact density
Mott transistor	Insulator to conduction transition	Poor mobility	New principle, size independence	Contact R, region transitions
BISFET	Potential low energy	Coherence and negative R	New principle	RT, variability,
Nanowires	Electrostatics	Surfaces	New fabrication techniques	Contact R, Variability
III-V's	Bulk mobility	Surface states	New materials	Variability, Poor inversion,
Topological Insulators	Surface conduction, bulk insulation	Low T, Small bandgap materials	New physical mechanisms	Small bandgap, low currents, Low T
Electrochem Memory	Atomic scale cond path	Stochasticity, interfaces, energy	Simple structure	Randomness, high field traps
Electromechanical	Mechanics, zero off Curr	Size, cycling	Configurability, memory, dynamic	Stiction, reproducibilty,
Spin-Semicond	No charge movement	Coherence lengths, temp	Less soft errors	Not RT, too much energy in conversions, loss
Spin transfer, magnetics	Speed,	Current, integration	Non-volatile, soft error immunity	Density, robustness in large # of thin films

Even now design is a limiting factor in CMOS integration density. What about new technologies?

Design as a limitation in CMOS integration density is a self-inflicted wound.

It is this way because of two principal reasons:

(a) Hierarchy – separation of design from technology, the ineffective abstractions, the fast cycle of design for products means that the design tool changes are fixes so that a creeky infrastructure can continue. **Design at system level needs stronger foundation**. Foundations of 70's can't last into 5 decades. *Just look at progress of 3D integration in past decade!*

(b) Energy constraint, higher order capabilities – adaptation, evolution, codesign, functional description are not intrinsic part of the approach.
Today's design approach is a Byzantiane balkanized process.

If this same style continues, the impact of new technologies will be small.

Responses to Charge to the Panel

- Which types of applications will be the drivers for Beyond CMOS?
  - Ultra-low power microsystems
    - Networks for healthcare diagnostics, environment monitoring, infrastructure safety, … , merging of heterogeneous technologies – mechanical is still in infancy
  - Machines with learning capabilities
    - Robotics for elderly, repetitive tasks, unsafe environments, …
  - Machines with inference capabilities
    - Merging of learning with inferencing on a physical platform that is not an HPC machine (Dr. Watson! Capabilities in compact forms)
  - … don't know what form it will take, *but effective education platforms – cost of education, like health care, is unsustainable and technology may provide answers*
- Will they all compete for the same killing application, or will they share the market?
  - Depends.
  - Common hardware blocks may exist since learning, inferencing is common to many of these applications, as is lower power.
  - Memory – low power, archival – will continue to expand in demand.

Responses to Charge to the Panel (cont'd)

- Will design challenges be different for different applications?
  - Yes and no.
  - Design for very compact low power human held instruments will be different from mini scale to large scale machines.
- Can design tools be the discriminating factor for the success of one specific technology?
  - Yes. There was a period in 80's when DEC (Digital)'s minicomputers competed with IBM's even though IBM technology was 3 generations ahead.
  - Look at the success of Apple, design tool is codification of a design process and its mathematical translation

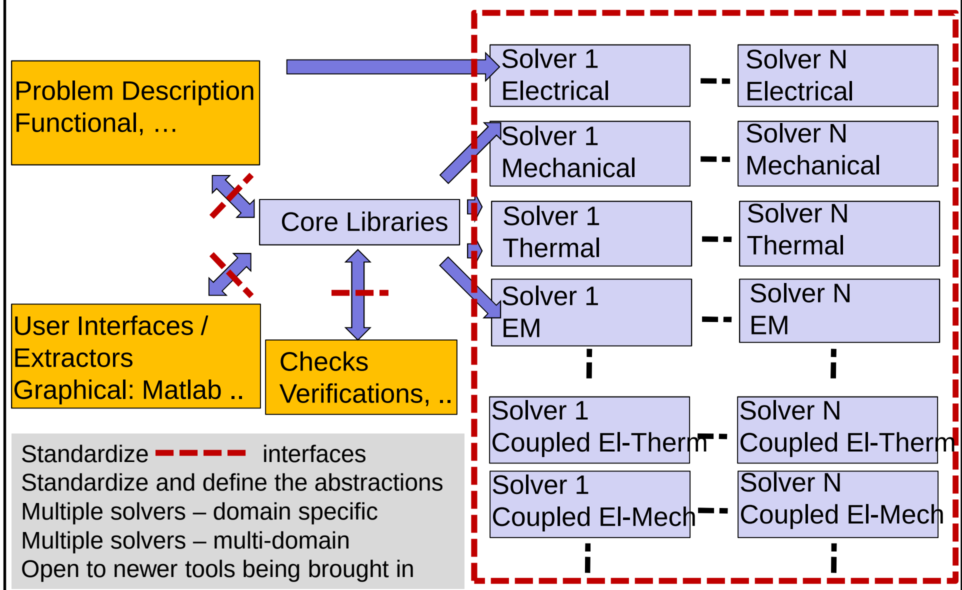
Responses to Charge to the Panel (cont'd)

- Present design tools are a huge legacy: what can trigger the investment needed for new tools?
  - The energy challenge. Present implementations of information processing could be at least 10x better in power, possibly up to 1000x
  - The application challenge. Society well-being challenge. Health-care, High school and college education are incredible economic burdens on society, and do not have to be. They are a diagnostic and treatment decision making problem, or knowledge exchange and development problem that information tools should be good at.

How?

- Semiconductor, electronic, computer, … technology industry can not do this because of economics and time cycles.
- This is an international -national scale problem that crosses frontiers of many disciplines that must come together (EE, Appl. Math, Phys, Comp. Sci., Mechanics, in particular, and Biology, Chemistry, …)
  - Needs a cooperative effort (national laboratories, universities, …) under a unified leadership (European community, DARPA-Industry, …). Requires many disciplines working together, much thinking at start, and a longer project duration.
  - It must be open, so that experts can contribute and has wider interdisciplinary impact.
  - Start with definition of what is critical information to pass between different specializations appropriate to modern systems needs (electronics, optics, magnetics, mechanical, … , to function of the system). System architects to engineers to scientists need to do this under strong direction and control so that it is scientifically write and open. **End result is abstractions and format of the interface that the community agrees to.**
  - People can work on design components using their own favorite languages, approach, …, so long as the interface uses a common language, so the right abstractions and their format that are openly defined. Industry (robustness, production oriented) and Academia (intellectually challenging, use-at-your-own-risk) can contribute to this without breaking economic models.

Build A New Open Infrastructure



New Technology Challenges

Molecular Electronics

- 100 C effects - reliability, >100 nA/µm currents, …

Solid-State Quantum Computing

- Define the system and why and how. If prime factorization is the only application, it is insufficient. Once factorized, it is a problem of look up table.
- *Spintronics*
- µA write, Integration at length scale, spin-to-electrical conversion, …

Nanowires

- V_T control, Multi-V_T and sigma, heat, tunneling is depth sensitive – dopant tails, …

~~Memristors~~ – Variable Resistors

- Synaptic architectures not understood; refresh-decay, isn't this "hysteresis", …

Graphene

- Contact resistances 1 order of magnitude better than silicon needed. For digital, suppressing leakage, variability, surface effects, …

New Technology Challenges

When the answer to improvement in a subset of properties is negative in a fundamental form at the lowest strata, the conclusion is easy.

When it passes that test, it is much harder.

The Challenge is that if the system use is of many components interacting together

Global Optimization versus Local Optimization under system constraints is a hard problem that requires a thorough design incorporating the abstractions.

Example: In late 70's, IBM stayed with bipolar because CMOS was too slow. The answer under power constraints was in architecture (similar to the multicore today).

So, system-scale design very critical.

Today: Look at ARM versus Intel. There are x10-100 more ARM processors in the world

19