JUMP 2.0 Theme 7: High Performance Energy-efficient Devices for Digital and Analog Applications

Harsono Simka
(h.simka@samsung.com)
Logic Pathfinding Group, Device Lab, Samsung Semiconductors Inc

January 25, 2022
JUMP 2.0 Proposers' Day Workshop

Acknowledgements: Woo-Bin Song, Yong-Seok Kim, Daewon Ha, Ken Rim (SEC); Tim Green & Todd Younkin (SRC)

Outline

- **Summary of Theme 7**
  - Goals and Topics

- **Towards a Strong Research Whitepaper in Theme 7**
  - Scope
  - Throughput and Coverage

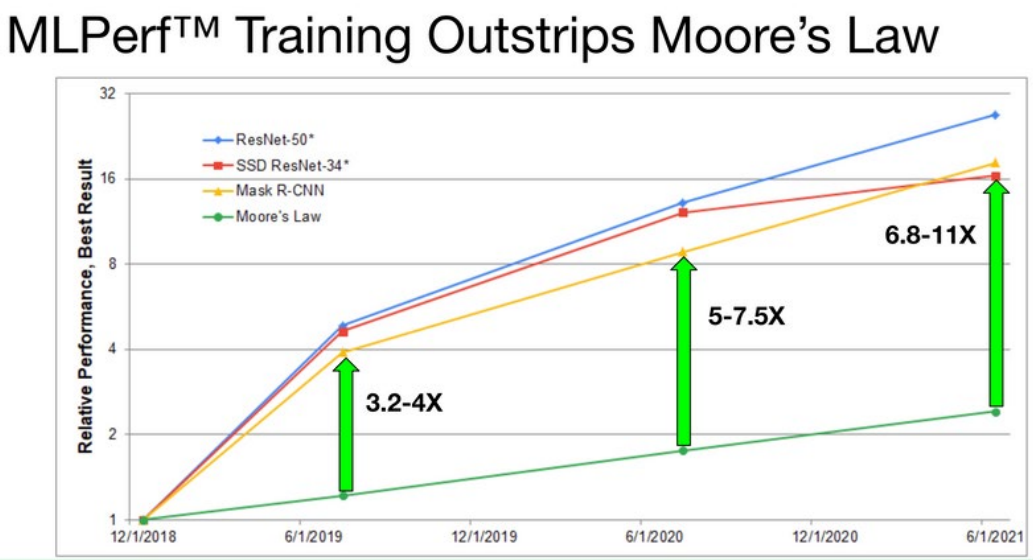
Motivation for Theme 7

AI Training Is Outpacing Moore's Law ›
The new set of MLPerf results proves it

BY SAMUEL K. MOORE | 02 DEC 2021 | 4 MIN READ

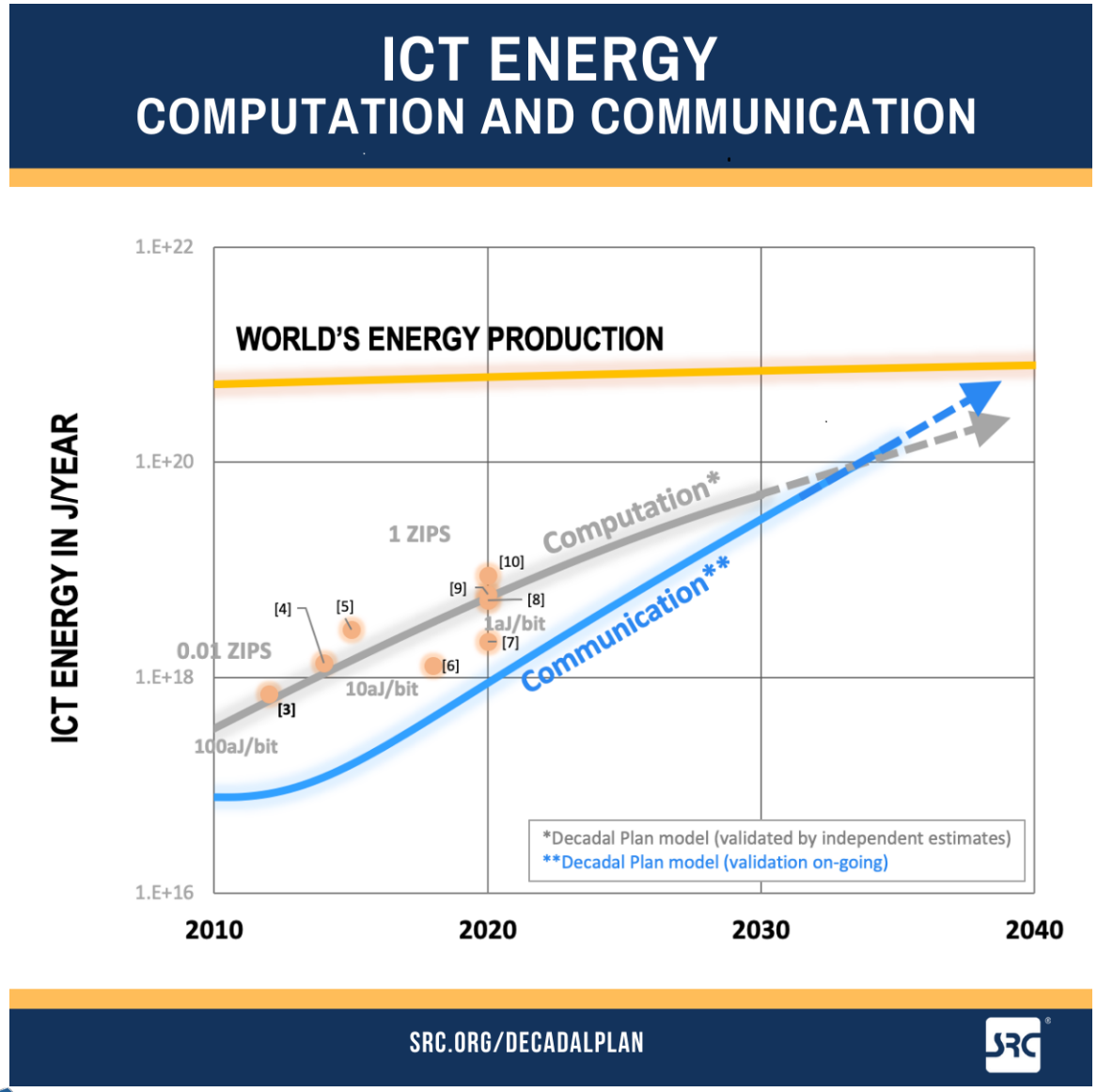
IEEE Spectrum, Dec 02, 2021, AI Training Is Outpacing Moore's Law - IEEE Spectrum

The gains to AI training performance since MLPerf benchmarks began "managed to dramatically outstrip Moore's Law," says David Kanter, executive director of the MLPerf parent organization MLCommons. The increase in transistor density would account for a little more than doubling of performance between the early version of the MLPerf benchmarks and those from June 2021. But improvements to software as well as processor and computer architecture produced a 6.8-11-fold speedup for the best benchmark results. In the newest tests, called version 1.1, the best results improved by up to 2.3 times over those from June.



- *Are "improvements to software as well as architecture" enough?*
- *Can we keep doing this with existing technology?*

Motivation for Theme 7



https://www.src.org/about/sustainability/ (Oct 2021)

- **Both Compute and Communication are important. Communication / data movement grows faster than Compute**
- **Need technology solutions that address both Compute and Comms. New trajectories needed:**
  - Improve compute by 100-1000X
  - Improve comms / data movement by 100-1000X

Theme 7 is a Technology Theme

SAP: 7 Critical System & Technology Themes

Cognition
- Systems & Architectures for next -gen AI
- Distributed: Heterogeneous, Cloud, Edge

Communications & Connectivity
- Orders of magnitude improvements are needed in capacity, energy -efficiency, security, and resiliency.

Intelligent Sensing to Action
- Analog & Mixed Signal Systems
- Sensor Fusion, Processing, and Interpretation

Systems & Architectures for Distributed Compute
- Distributed = Heterogeneous, Edge, Cloud
- Energy Efficient Compute
- Accelerators & Accelerator Fabric

Intelligent Memory & Storage
- Full-stack optimization of Intelligent Memory systems
- Emerging memory devices & arrays, including enhanced functionalities

Advanced Monolithic and Heterogenous Integration
- Interconnect Fabrics including Photonics
- Architectures and Applications for Advanced Packaging

High-performance energy efficient devices for Digital & Analog Applications
- Novel Materials, Devices, & Interconnect technologies

System Themes

Technology Themes

Novel Materials, Devices, and Interconnect embedded

Common To All

- Energy Efficiency
- Security & Resiliency: Trusted, Private, Robust, Assured, Safe
- Platform Demonstrator(s)
- Prototyping, & Scenarios & Analysis
- CoDesign Integration Capability
- Modeling & Simulation
- Benchmarking
- Metrology & Characterization

SRC Confidential: Select Disclosure

Theme 7 needs to help enable System Themes (1-5) → broad agenda

Theme 7 Overview

- **Disruptive innovations in active and passive devices → *Breakthroughs***
- **Orders of magnitude improvements in scaling, energy efficiency, density, throughput, latency, and novel functionalities**
- **Close coupling between materials and device research**
- **Applications:**
  - Distributed compute
  - Communication
  - Sensing
  - Networking
  - Memory and storage systems
  - …
- **Goal: foundational building blocks from novel materials, processes, devices**

  Common features:
  - Co-design with circuit and architectural considerations, with various FOM and operating characteristics
  - Materials development, innovative metrologies, device demonstration of viable process integration
  - Experimental demonstrations, benchmarking, multi-scale physics-based modeling

Theme 7 Details

Specific need areas include, but are not limited to:

1. High-performance, energy-efficient devices for logic, memory, analog computing, power and sensing
    a. FEOL high-performance devices beyond Si
    b. BEOL compatible high-performance devices
    c. Nonvolatile logic
    d. Low-thermal budget logic devices for 3D
    e. Low power logic
    f. Bio-inspired electronics
    g. Integrated photonics for sensing and computing
    h. Integrated very high efficiency power conversion and delivery devices
2. High-performance passive components
    a. High-density capacitors for power and >300GHz performance
    b. Magnetics/inductors for integrated RF and power
    c. Low-loss interconnect: metallic, optical, new material
3. Novel memory (e.g., scalable NVM that can be integrated with CMOS) and exploratory memory element concepts
    a. Memory materials and selectors for 3D memory/storage: materials, deposition and etching methods enabling stacked-layer configuration and/or 3D-NAND type embodiments of NVM
4. Devices enabling disruptive memory applications including but not limited to BEOL compatible transistors for dense memory arrays and innovative Selector devices
    a. Transistors supporting Ultra-low leakage (< atto-amp) and High on/off ration (>10^6) for memory-centric applications
5. Modeling of technologies and designs for application requirements for synthesis
    a. Predictive materials modeling and nanomechanical structural modeling including interfaces
    b. Full-stack modeling—process relation to structures and interfaces throughout manufacturing, including thermal effect (e.g., self-heating effect, heat dissipation scheme, etc.)
    c. Enabling data science/ML for augmenting Materials & Structural modeling
    d. Benchmarking of technologies and designs for application requirements
6. Advanced Manufacturing Technology & Integration, including advanced patterning
    a. Selective processing (ASD, ALD, ALE, MLD)
    b. Atomically precise processing and interfaces
    c. Ability to control bottom-up process
    d. Processing techniques that do not create surface damage
    e. High-throughput atomic lithography capabilities
7. Innovative metrology, physical-electrical-thermal characterization, test platforms, and from devices to memory systems
8. Material development beyond logic, memory, and interconnect needs:
    a. Materials for > 300GHz RF and analog devices
    b. Materials for photonics: laser integration, photonics processing, interconnect
    c. Thermal materials and materials for extreme environments (thermal shock, high E-field, radiation hard, etc.)
9. Enhanced Materials Discovery. HTE approaches that enable coupling between
    a. High throughput synthesis,
    b. Rapid measurement of the most important 1-2 physical properties, and
    c. AI/ML guided algorithm to learn from the measured data and intelligently identify more promising material compositions for further synthesis, are encouraged.

    *Examples can be found in the state-of-the-art research in clean energy and catalysis, and recently initiated thermoelectric material screening. Detailed description and further examples of HTE applications in electronic, magnetic, optical, and energy-related materials can be found in Green et al, Journal of Applied Physics 113. 231101 (2013), https://aip.scitation.org/doi/10.1063/1.4803530).

10. Enabling the fundamental examination/re-engineering of compute stack via codesign of materials/device/circuit/ architecture/algorithm/programmability. List below provides examples from 6 themes above as a guide; not inclusive or prescriptive.
    a. Optimized device, materials, and architecture to enable for novel algorithms and nontraditional compute, e.g., stochastic computing, high-/hyper-dimensional computing, spiking NN
    b. Device, materials, and architecture to enable small, low-cost compute + memory + sensor systems
    c. Compute/processing-in-memory (CIM/PIM), e.g., digital and analog embedded NVM devices
    d. Compute-near-memory, e.g., digital embedded NVM devices
    e. Hardware for secure computing, e.g., homomorphic encryption, physically unclonable functions, random number generators
    f. Deep learning acceleration, e.g., multi-state and analog synaptic cells with optimal stability, noise, etc. for inference, and update behavior, endurance, etc. for training
    g. 3D deep learning architectures—logic, memory
    h. New device concepts and reimagined technologies that provide for deterministic and large resistive ratio memory devices that are ideally suited for multi-state, toward analog-like performance (may include, but not confined to magnetic, FE, resistive, and PC materials).
        i. Memory technologies for analog and/or brain-inspired compute -> device / novel materials that can reliably achieve deterministic, multi-state functionality
    i. Components and architecture for stochastic computing-based Boltzmann Machines and deep learning network:
        i. Low-cost, high-performance, massively parallel, scalable, tunable stochastic components and circuit implementation concepts
        ii. Architecture- and circuit-level feasibility study for demonstrating orders of magnitude advantage beyond von Neumann-based architecture.

Theme 7 Examples of Desired Topics

1. High-performance, energy efficient devices for logic, memory, analog computing, power, and sensing
2. High-performance passive components
3. Novel memory (e.g. scalable NVM that can be integrated with CMOS) and exploratory memory element concepts
4. Devices enabling disruptive memory applications including but not limited to BEOL compatible transistors for dense memory arrays and innovative selector devices
5. Modeling of technologies and designs for application requirement for synthesis
6. Advanced Manufacturing Technology & Integration
7. Innovative metrology, physical – electrical – thermal characterizations, test platforms, from devices to memory systems
8. Material development beyond logic, memory, and interconnect (e.g. RF, analog, thermals, extreme conditions)
9. Enhanced Materials Discovery (High Throughput Experiments)
10. Enabling the fundamental examination / re-engineering of compute stack via codesign of materials / device / circuit / architecture/algorithm/programmability

From "JUMP 2.0 Research Needs V3", released Dec 2022

Outline

- Summary of Theme 7
  - Goals and Topics

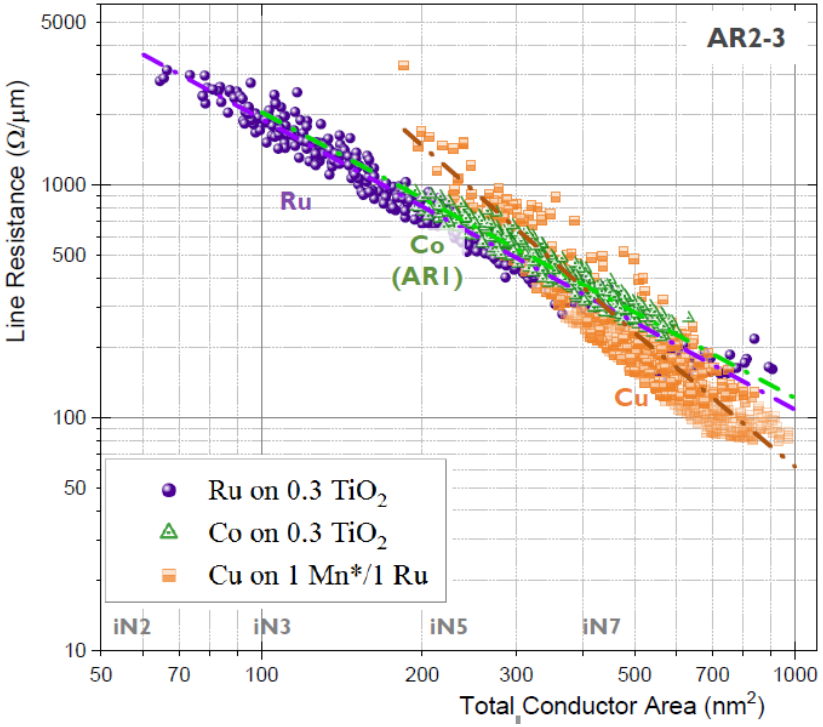
- **Towards a Strong Research Whitepaper in Theme 7**
  - Scope
  - Throughput and Coverage

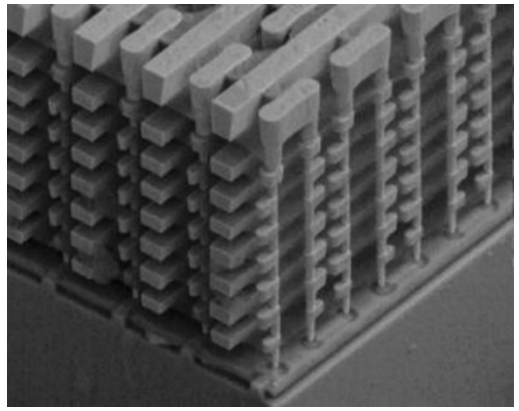
Defining the Scope: Interconnect as Example (1)

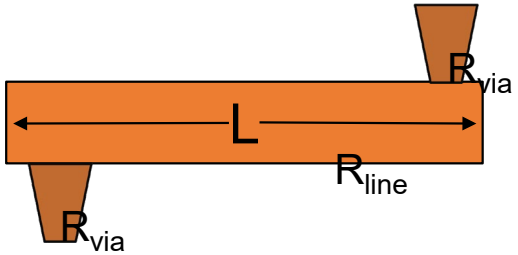
- **Start with benchmarking and define a relevant metric**

SRC nCORE kickoff meeting, Feb 2020
SRC JUMP ASCENT review, Aug 2021

Data: M. van der Veen et al (imec), IITC 2018

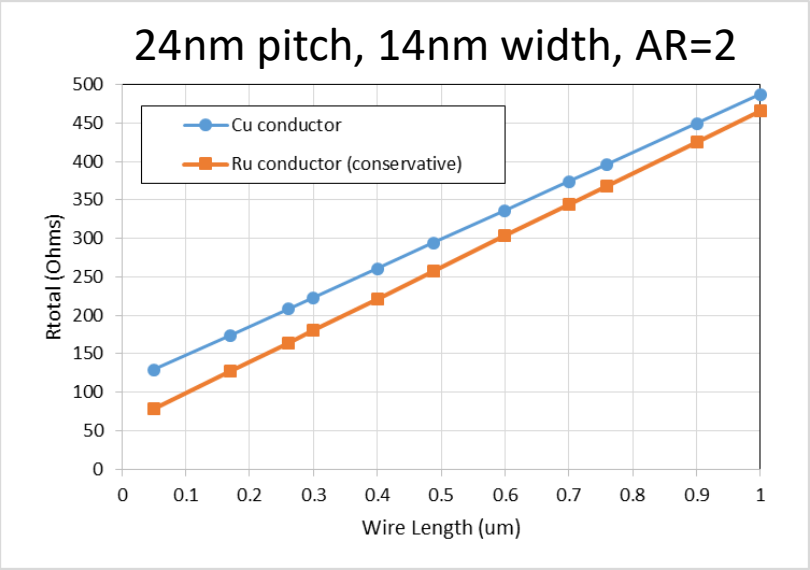






A simple metric for interconnect

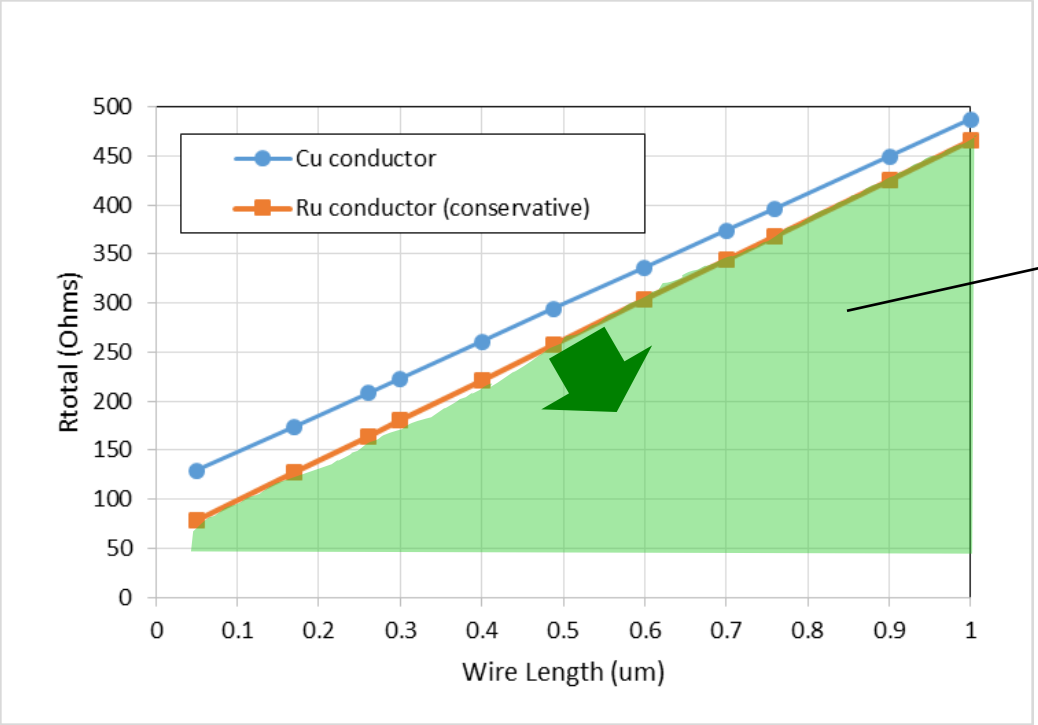
$$R_{total} = R_{line} + 2\,R_{via}$$



Defining the Scope: Interconnect as Example (2)

- **Use the metric to set research goals**

SRC nCORE kickoff meeting, Feb 2020
SRC JUMP ASCENT review, Aug 2021



Systematic research desired:

- New physics
- New materials

Defining the Scope: Interconnect as Example (3)

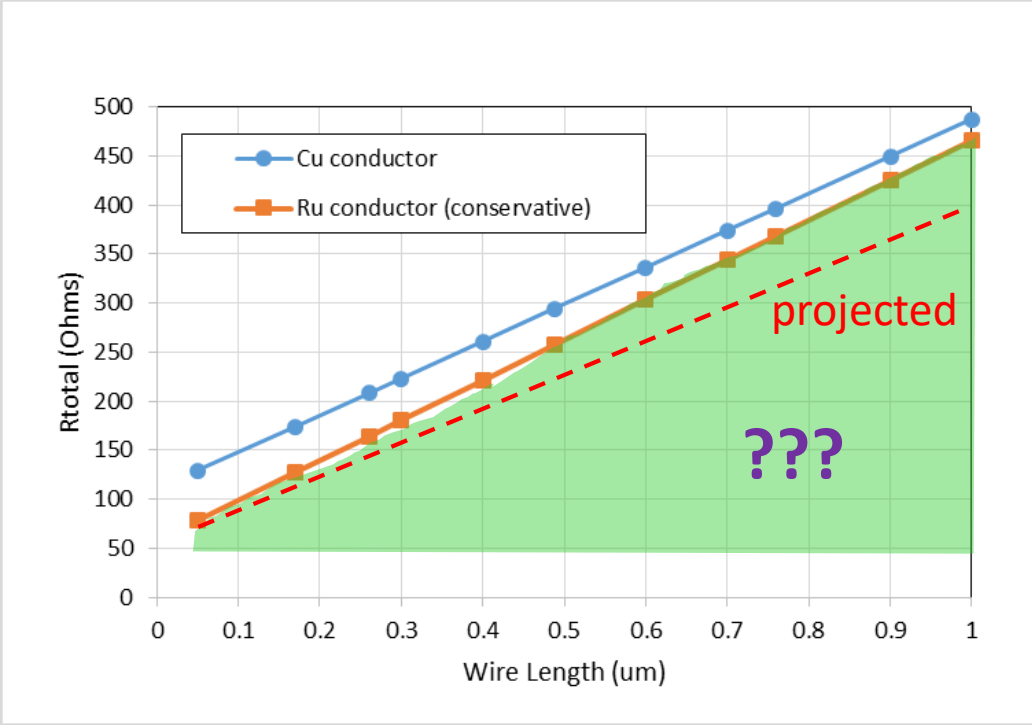
- **Evaluate and improve**

SRC research progress in 2020-2021:

SRC JUMP ASCENT review, Aug 2021

- Successful **precursor synthesis** of Ru, Co, Ti (Wayne State, ASCENT 2776.057)
- Improved nucleation Ru ALD led to larger grains and **bulk-like resistivity** (UCSD, ASCENT 2776.010)
- EE-ALD with background reactive gas led to **low temperature growth** of smooth TiN and Ru films (Colorado, ASCENT 2776.009)
- Anneal of Ru with Graphene cap and/or 2D liner led to **lower resistivity** (Purdue, nCORE NEWLIMITS 2819.020)
- Synthesized **MoP nanowire** resistivity may be **lower than Ru** at <1000nm^2 (Yale, nCORE IMPACT 2966.005)**
- **CoSi** (a semimetal) grown on GaN led to **very low resistivity** for 20nm film (Notre Dame, nCORE IMPACT 2966.004)
- A few ternary and quarternary material options identified (Stanford, nCORE IMPACT 2966.001; RPI, nCORE IMPACT 2966.002)

** Prof Judy Cha won the 2021 SRC Young Faculty Award for her work on MoP → industry acknowledgement of importance of new interconnect materials research



- **Need to consider stability, resistance to oxidation, reliability, fabrication, cost → Multiple options needed**

Throughput and Coverage: Challenges

- **A typical research group can experimentally evaluate only ~2 materials per year**
  - Most options left unexplored (Billions of materials can be made from ~70 readily available elements)
  - Each multi-year program can only look at several materials → expensive in time and opportunity cost
  - Tend to "lock-in" research direction, even when data shows chosen materials are not feasible
  - Throughput too small to enable effective coupling between modeling and experiments
- ***Can we do this differently? What about using High Throughput Experiments?***

http://dx.doi.org/10.1063/1.4803530

JOURNAL OF APPLIED PHYSICS **113**, 231101 (2013)

APPLIED PHYSICS REVIEWS—FOCUSED REVIEW

Applications of high throughput (combinatorial) methodologies to electronic, magnetic, optical, and energy-related materials

Martin L. Green,[1] Ichiro Takeuchi,[2] and Jason R. Hattrick-Simpers[3]
[1]*Materials Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA*
[2]*Department of Materials Science and Engineering, University of Maryland, College Park, Maryland 20742, USA*
[3]*Department of Chemical Engineering, University of South Carolina, Columbia, South Carolina 29208, USA*

(Received 10 August 2012; accepted 16 April 2013; published online 17 June 2013)

High throughput (combinatorial) materials science methodology is a relatively new research paradigm that offers the promise of rapid and efficient materials screening, optimization, and discovery. The paradigm started in the pharmaceutical industry but was rapidly adopted to accelerate materials research in a wide variety of areas. High throughput experiments are characterized by synthesis of a "library" sample that contains the materials variation of interest (typically composition), and rapid and localized measurement schemes that result in massive data sets. Because the data are collected at the same time on the same "library" sample, they can be highly uniform with respect to fixed processing parameters. This article critically reviews the literature pertaining to applications of combinatorial materials science for electronic, magnetic, optical, and energy-related materials. It is expected that high throughput methodologies will facilitate commercialization of novel materials for these critically important applications. Despite the overwhelming evidence presented in this paper that high throughput studies can effectively inform commercial practice, in our perception, it remains an underutilized research and development tool. Part of this perception may be due to the inaccessibility of proprietary industrial

TABLE OF CONTENTS

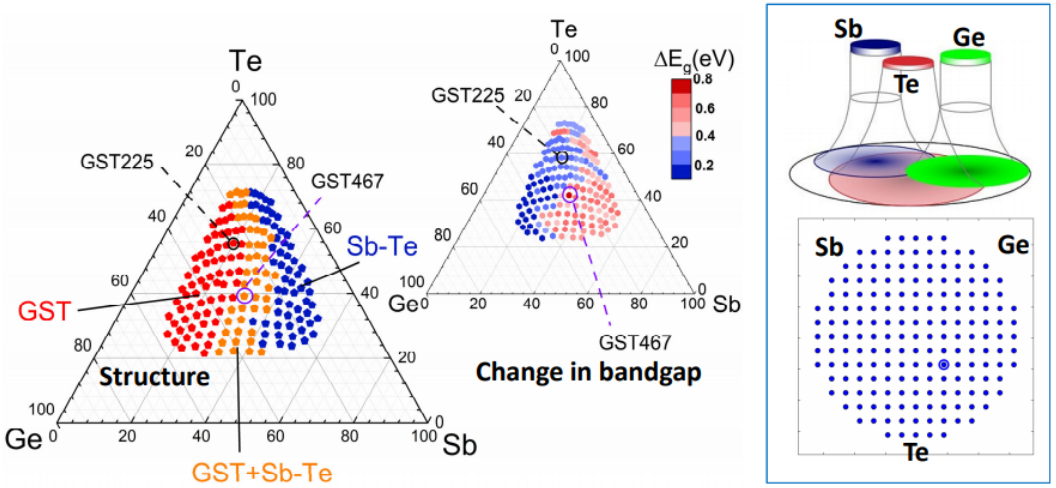
Key features:

- **Automated, parallel synthesis of materials (hundreds to thousands at a time)**
- **Rapidly test each individual material or composition for 1-2 properties**

Throughput and Coverage: Phase Change Materials Screening Example

- **Materials screening for novel PCM (SRC IMPACT 2966.010. Ack: Prof Ichiro Takeuchi, U. Maryland)**
  - High Throughput Experiments: using high vacuum sputtering with 3 targets
  - Fast characterization of bandgap, using laser beam scanning ellipsometry
  - Extended to explore a new class of ternary PCM (Ti-Sb-Te)

Combinatorial screening and discovery of GST467



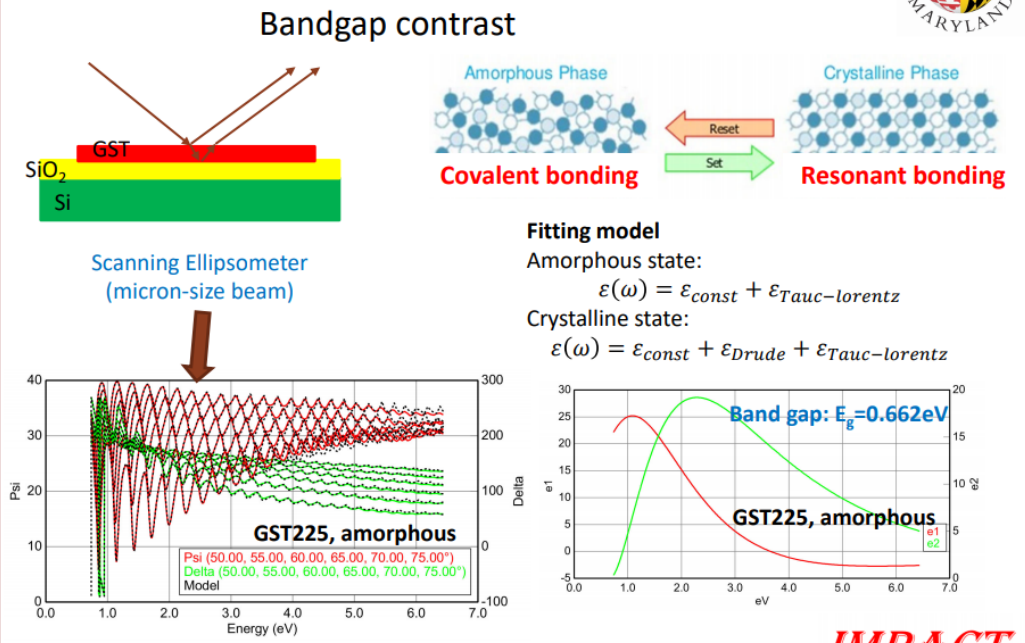
Rapid screening of compositions based on structural phases and change in bandgap:

- ☐ Map the phase diagram quickly via active learning
- ☐ Discovered **GST467** with largest bandgap difference and optical contrast
- ☐ Reproducible in multiple composition spreads

Kusne *et al.*, Nat. Commun. **11**, 5966 (2020)

IMPACT
Innovative Materials and Processes for Accelerated Compute Technologies

Key property screening: bandgap determination



IMPACT
Innovative Materials and Processes for Accelerated Compute Technologies

Throughput and Coverage: Solar Cell Materials Discovery Example

- https://www.technologyreview.com/2021/04/27/1021753/fast-forward-2/ : "Speeding up the development of solar cell materials, using machine learning, robots, and good old-fashioned teamwork" (Ack: Prof Tonio Buonassisi, MIT)
- Goals: to find viable alternatives to silicon for use in capturing solar energy. "The sheer number of perovskites is exciting—there could be thousands out there that match up well with different applications."

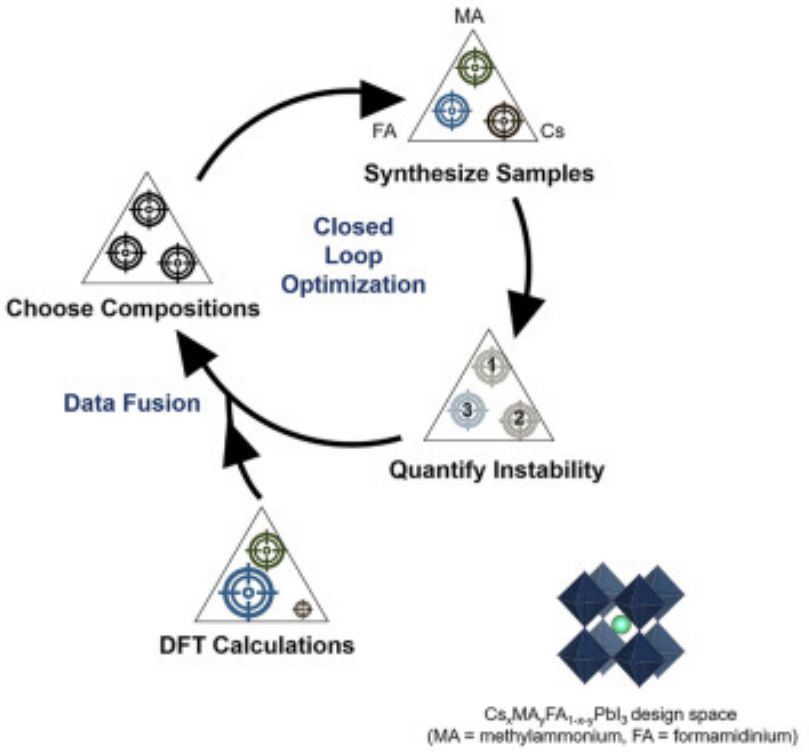
A data fusion approach to optimize compositional stability of halide perovskites

Shijing Sun [7] • Armi Tiihonen [7] • Felipe Oviedo • ... Vladan Stevanovic • John Fisher III • Tonio Buonassisi [8] • Show all authors • Show footnotes

Published: February 01, 2021 • DOI: https://doi.org/10.1016/j.matt.2021.01.008 • Check for updates

Highlights

- Physics-informed machine learning enables accelerated search of stable perovskites
- Closed-loop Bayesian optimization takes the human out of the decision-making loop
- >17-fold higher stability achieved within a combinatorial space of $Cs_xMA_yFA_{1-x-y}PbI_3$
- Improved perovskite thin-film stability translates into enhanced solar cell reliability



- Three required components:
  (1) High-throughput experiments
  (2) Automated analysis to run tests in parallel
  (3) Algorithms trained on theoretical knowledge and previous results to make smarter guesses

Throughput and Coverage: Device Research Example

- **Increase throughput by using Machine Learning to find optimal process conditions faster**

Embedding physics domain knowledge into a Bayesian network enables layer-by-layer process innovation for photovoltaics

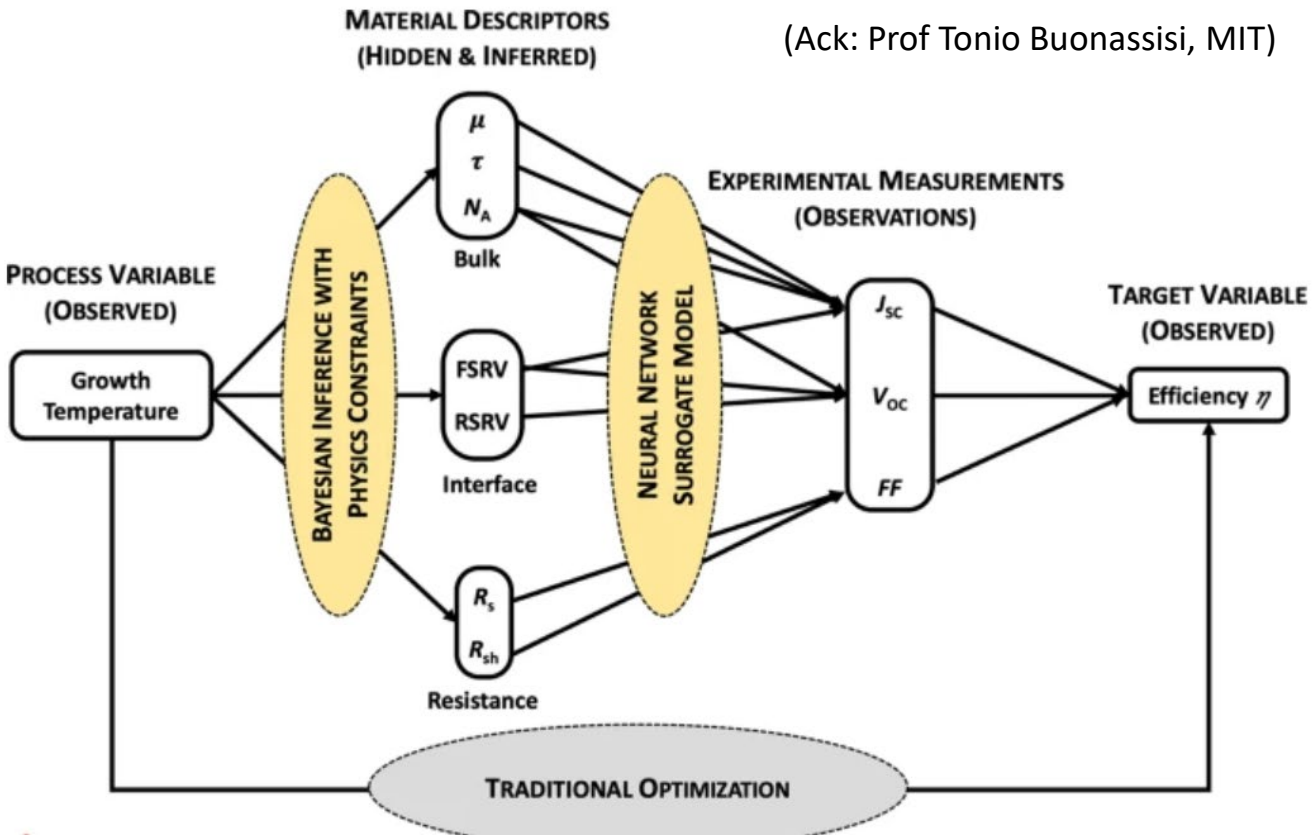
Zekun Ren, Felipe Oviedo, Maung Thway, Siyu I. P. Tian, Yue Wang, Hansong Xue, Jose Dario Perea, Mariya Layurova, Thomas Heumueller, Erik Birgersson, Armin G. Aberle, Christoph J. Brabec, Rolf Stangl, Qianxiao Li, Shijing Sun, Fen Lin, Ian Marius Peters & Tonio Buonassisi

npj Computational Materials **6**, Article number: 9 (2020) | Cite this article

We developed and applied a Bayesian network approach to GaAs solar cell growth optimization. This approach enables us to exceed our baseline efficiency by 6.5% relative, by tuning process variables layer by layer, in just six MOCVD experiments. Our approach is enabled by implementing physics-informed relations between process variables and materials descriptors, and embedding these into a Bayesian network. We link these material descriptors to device performance using a neural network surrogate model, which is 100× faster than numerical device simulation. The small number of growth (MOCVD) runs necessary to implement this layer-by-layer process-optimization scheme translate into significant cost and time reductions compared to common black-box optimization methods. We believe this approach can generalize to other solar cell materials,[55,56] as well as other systems with physics-based or black-box relations between process variables and materials descriptors, and physics-based device performance models. Our surrogate model can replace common models in closed-loop black-box optimization, such as a Gaussian process regression in Bayesian optimization, providing good functional fitting and physical insights.

Fig. 1: Schematic of our Bayesian network-based process-optimization model, featuring a two-step Bayesian inference that first links process conditions to material descriptors, and then the latter to device performance.

(Ack: Prof Tonio Buonassisi, MIT)



- **New examples appearing in publications (CNT, 2D transistor device, …)**

Suggestions to Proposers

- **Theme 7 has a broad agenda. Focus on key ideas and try not to cover all topics**

- **Center proposers:**
  - Stay focused on what your center wants to do, and listen to the collective inputs from industry & DARPA
  - Articulate and co-develop links to other Themes
  - Define the right scope and be aggressive, strive towards orders of magnitude improvements
  - Balance and cross-pollinate between device and materials research
  - Think "100X" in throughput and coverage for materials screening, and device optimization
  - Create collaboration opportunities and shared direction among center researchers

- **Individual proposers:**
  - Understand the theme goals and how your proposal helps the center goal
  - Articulate how your proposal can expand the scope, or increase the throughput