

*Exceptional service in the national interest*



## Engineering Challenges for Reversible Computing Hardware

Michael P. Frank  
Center for Computing Research  
Sandia National Laboratories

COST Action IC1405, MC plus WGs Meeting 1 of GP4  
College Court, University of Leicester, Leicester, UK  
Tue., Sep. 11<sup>th</sup>, 2018

Approved for Public Release  
SAND2018-9891 PE



Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

### Abstract (hide during talk)



The reversible computing paradigm offers the only physically possible means by which energy efficiency for general computing applications may be extended beyond the Landauer limit, which threatens to stall the progress of conventional computing technology in the not-too-distant future. This fact provides a strong motivation to pursue research and development work on novel hardware technologies focusing on the reversible paradigm. However, to design and build reversible hardware that is, in fact, extremely energy-efficient in practice (and at a reasonable cost) still presents a substantial engineering challenge. It would be highly desirable to identify new low-level device technologies that offer an improved energy-delay product for reversible operation (along with lower standby power), but so far, the device research community has almost completely ignored the reversible paradigm. Furthermore, regardless of the device technology, efficiently synchronizing the operation of a reversible machine presents substantial difficulties, which we are attempting to address in a couple of current research projects at Sandia. Although these and other engineering problems are likely not insoluble, solving them will likely require a much more intensive level of investment dedicated to R&D on low-level hardware technologies for reversible computing than this area has received to date.

# Talk Outline

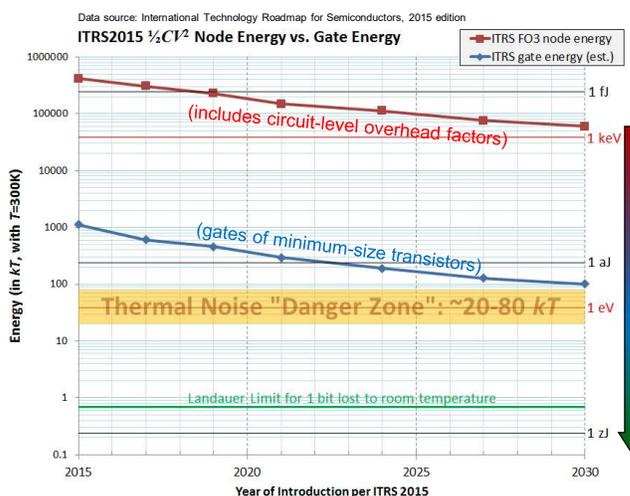


- **Context:** Reversible Computing offers a nascent technology revolution, but its successful emergence is threatened by...
  - Misunderstandings and confusion about fundamental physics issues
  - Failure of mainstream electronics industry to address this challenge
- Fundamental physical issues that still need to be addressed:
  - Broadening RC theory to a nonequilibrium thermodynamics context
  - Exploring how quantum effects might facilitate isentropic dynamics
    - We still need a full buildable, self-contained quantum model of RC!
- Specific engineering challenges that still need to be tackled:
  - Widening search for device types suitable for reversible operation
  - Optimization of device design parameters for reversible operation
  - Efficient resonant energy-recovering supplies for adiabatic circuits
  - Efficient elastic interaction elements for ballistic circuits
- **Conclusion:** Recommendations to funding agencies...

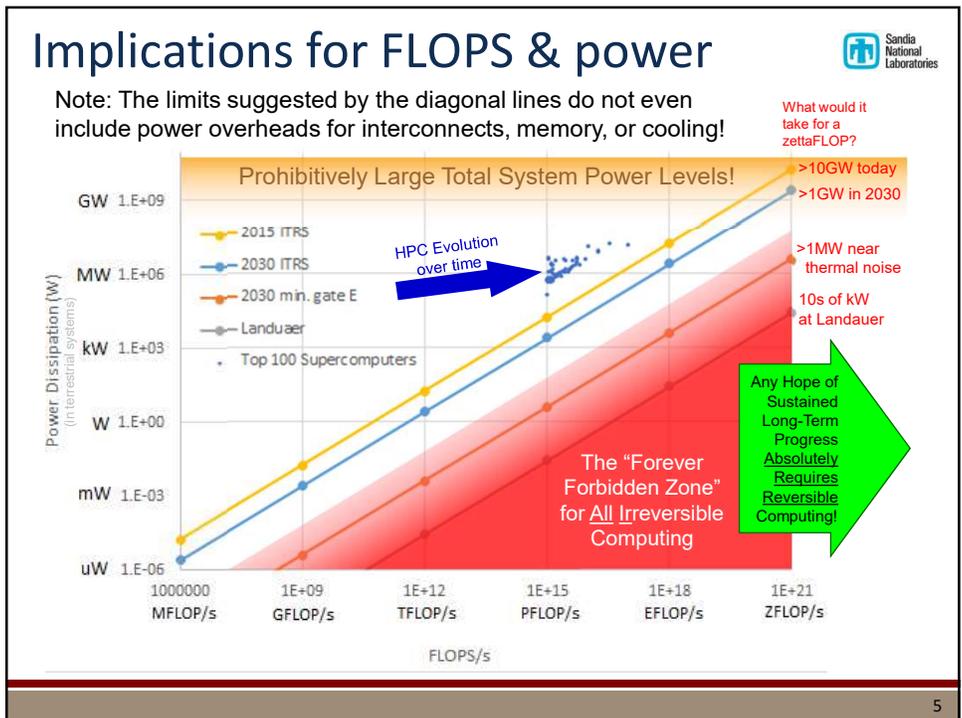
# Semiconductor Roadmap is Ending...



- Thermal noise on gates of minimum-width segments of FET gates leads to channel PES fluctuations when  $E_g \lesssim 1-2 \text{ eV}$ 
  - Increases leakage, impairs practical device performance
    - Thus, ITRS has minimum gate energy asymptoting to  $\sim 2 \text{ eV}$
- Also, real logic circuits incur many further overhead factors:
  - Transistor width  $10-20 \times$  min.
  - Parasitic (junction, etc.) transistor capacitances ( $\sim 2 \times$ )
  - Multiple ( $\sim 2$ ) transistors fed by each input to a given logic gate
  - Fan-out to a few ( $\sim 3$ ) logic gates
  - Parasitic wire capacitance ( $\sim 2 \times$ )
- Due to all these overheads, the energy of each bit in real logic circuits is many times larger than the min.-width gate energy
  - $375-600 \times$  (!) larger in ITRS'15
    - $\therefore$  Practical bit energy for irreversible logic asymptotes to  $\sim 1 \text{ keV}$ !
- Practical, real-world logic circuit designs can't just magically cross this  $\sim 500 \times$  architectural gap!
  - $\therefore$  Thermodynamic limits imply much larger practical limits!
  - The end is near!



Only reversible computing can take us from  $\sim 1 \text{ keV}$  at the end of the CMOS roadmap, all the way down to  $\ll kT$ .



## Reversible Computing – Why? What?

Sandia National Laboratories

- Fundamental microphysics is *reversible*—it conserves information!
  - Therefore, losing information from a digital system (by erasing/overwriting it) necessarily implies ejecting that information into the system’s environment
    - Once thermalized by the environment, information that was previously known (correlated) becomes entropy (unknown/uncorrelated information)
      - ...and this implies dissipation of  $kT \ln 2$  of organized energy (work) to heat at temperature  $T$  per bit of information lost (Landauer’s Principle)
- Unfortunately, in the conventional (irreversible) computing paradigm, we discard computational information all the time...
  - Every active conventional logic gate destructively overwrites its output node on every clock cycle, losing the information embodied in the previous output
    - Similarly for line drivers, on every bus cycle for every interconnect wire
    - And for memory cells/lines, every time a cell is written, read out or refreshed
- How can we compute without losing information? (And please note that “computing” includes driving interconnects, accessing memory, etc. as needed!)
  - *Reversibly transform states, instead of destructively overwriting them!*
    - This then allows avoiding the Landauer principle’s limit on energy efficiency

There is no known fundamental (technology-independent) limit on computational energy efficiency, but only if the reversible computing principle is used!

## International Roadmap for Devices and Systems (IRDS), 2017 edition

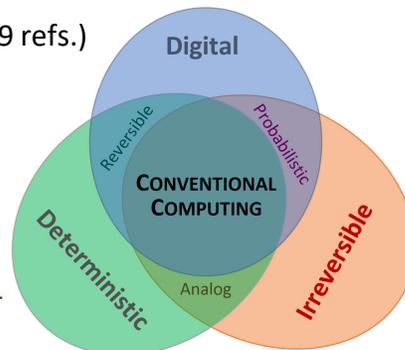


irds.ieee.org

- *Beyond CMOS* chapter (123 pp., 1079 refs.)

- *Sec. 5, Emerging Device-Architecture Interaction*

- Focuses on unconventional computing paradigms (besides quantum)
      - (Quantum computing will have its own dedicated chapter in the next edition)
    - A relevant paragraph from the section's Introduction is quoted below...
      - I'll elaborate on some of these points.



**Reversible (adiabatic and/or ballistic) computing** (§5.4) – Computing paradigms that approach logical and physical reversibility offer the potential to greatly exceed the energy efficiency of all other approaches to general-purpose digital computation. Primitive devices for reversible computing may include devices having fairly conventional functions (such as switches or oscillators). These devices would need to be optimized differently to use quasi-reversible physical processes such as near-adiabatic state transitions, near-ballistic signal propagation, highly elastic interactions, and highly underdamped oscillations. Reversible devices also must be organized into circuits and architectures in tightly constrained ways, for reversibility at the logical as well as physical level.<sup>1050</sup> Careful fine-tuning and optimization of analog circuit characteristics (e.g., resonator quality factors or elasticity of ballistic interactions) remains a difficult and crucially important engineering challenge that must be met in order for this paradigm to realize its promise.

7

## Basic Physics of Computing Issues



- The literature is full of fundamental misunderstandings about Landauer's principle and reversible computing theory...
  - Constantly, people are generating false "disproofs" of these concepts...
    - Simply beating back all of the misinformation would be a full-time job!
- **Needed:** A reformulation of the basic thermodynamics of computation and reversible computing theory in the language of nonequilibrium/stochastic thermodynamics & stat. mech.
  - Landauer's principle itself can *already* be proved without having to make any essential equilibrium assumptions... *However:*
    - Reformulating a complete theory of Landauer's principle / reversible computing in standard non-equilibrium language should substantially help dispel confusion
      - E.g., the precise role of the fluctuation/dissipation theorem should be more thoroughly addressed
        - » With clear proofs demonstrating why the theorem does not prevent us from asymptotically approaching reversible operation
- **Also needed:** A *full* quantum-mechanical model of reversible computing. (*Self-contained, complete, realistic, buildable.*)

8

## Computer Science issues

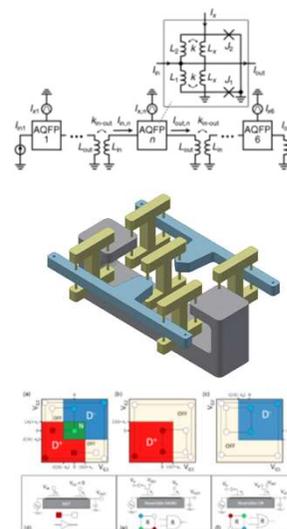


- The CS research community (within its reversible computation subfield) already has begun to address these topics, but more work is still needed in a number of important areas, such as:
  - More space/time efficient reversible algorithms for important problems
  - Broadening reversible logic theory & synthesis efforts to include more general classes of models of reversible computation, including:
    - Generalized (Conditional) Reversible Computing (topic of my RC17 paper)
      - Appropriate for adiabatic circuit design; c.f. collab. with Wille & Zulehner
    - Asynchronous (Ballistic) Reversible Computing (topic of my ICRC17 paper)
      - Basis of a \$1.5M internal superconducting circuit design effort at Sandia
  - Hardware description languages for adiabatic/reversible circuit design
    - Ongoing dialogue with Wille/Zulehner @ JKU, & Perumalla @ ORNL
  - **Perhaps most important:** Systems engineering of novel computer architectures that trade off energy savings via reversibility vs. realistic cost metrics in key areas including hardware efficiency and serial performance, while accounting for real nonidealities and parasitic losses
    - This is one is more engineering than CS, but it is nevertheless essential!

## Device Technology Issues

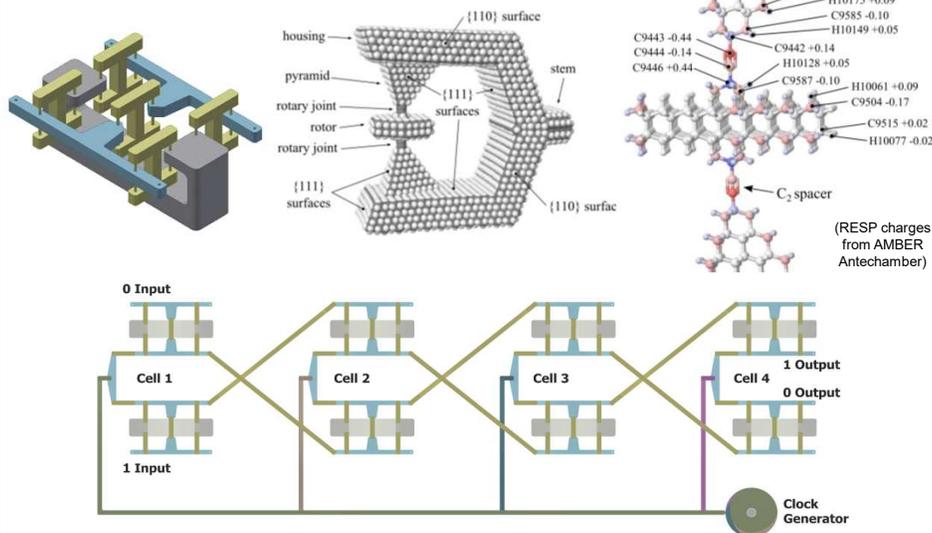


- Some of the recently-active research areas/groups in terms of device technologies for reversible computing include:
  - Reversible adiabatic superconducting logics
    - Vasili Semenov (& student Jie Ren) at SUNY Stonybrook
      - Results near  $kT$ , but this line of work is not currently active
    - AQFP/RQFP group at Yokohama National University (Japan)
      - Yuki Yamanashi, Noboyuki Yoshikawa, Naoki Takeuchi, etc.
  - Nanomechanical rod logic
    - Ralph Merkle and colleagues at IMM
      - Improving upon earlier work by K. Erik Drexler
    - More large-scale modeling still needed
    - Still very far from manufacturability
  - Valleytronics (China/Singapore-based collaboration)
    - Newly emerging research area
  - There are a few scattered others (QDCA, etc.), but this field is not very well unified/coherent...
    - **Needed:** A workshop dedicated to device & circuit engineering for reversible computing!



## Nanomechanical Link Logic

Merkle et al., IMM Report 46 and Hogg et al., DOI: [10.1039/C7ME00021A](https://doi.org/10.1039/C7ME00021A)

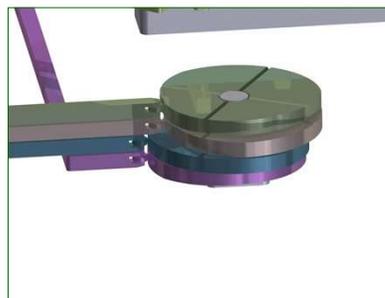


11

## Link Logic Lock Operation



- Videos animate schematic geometry of a pair of locks in a shift register
- Molecular Dynamics modeling/simulation tools used for analysis include:
  - LAMMPS, GROMACS, AMBER Antechamber
- Simulated dissipation:
  - $\sim 4 \times 10^{-26}$  J/cycle at 100 MHz
    - 74,000  $\times$  below the Landauer limit for irreversible ops!
- Speeds up into GHz range should also be achievable



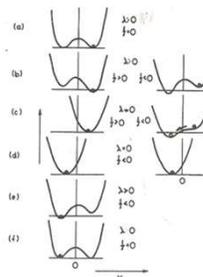
12

# Adiabatic Reversible Computing



A general class of implementation techniques for reversible computing that relies on controlled adiabatic transformations of the information-bearing degrees of freedom.

- Has been explored in various physical systems:
  - Superconducting electronics (Likharev '77, etc.)
  - LC switching circuits (Fredkin & Toffoli '78)
  - Adiabatic CMOS (Seitz '85, etc.)
  - Molecular nanomechanical logic (Drexler '91, etc.)
  - Single-electron quantum dots (Lent '92, etc.)
- Some drawbacks of this class of approaches:
  - Every logic transition must be explicitly driven by a power-clock
    - Numerous clocks are required in combinational and sequential designs
    - Substantial design complexity overhead to distribute clocks to every gate
  - Challenging to design finely-tuned, high-Q power-clock resonators
    - Problems with load balancing in long-range global clock distribution networks with large parasitics, avoiding data-dependent back-action



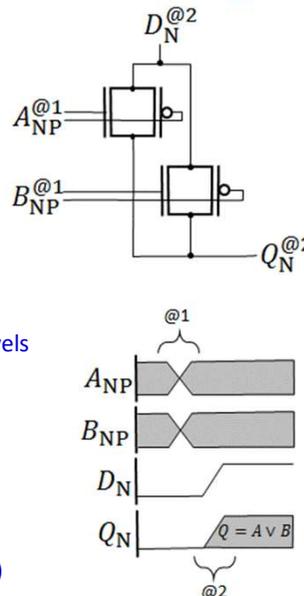
# Conditionally-Reversible Boolean Logic in Adiabatic Circuits



- This simple CMOS structure can be used to do/undo latched reversible **OR** operations
  - Example of 2LAL logic family (Frank '00)
    - Based on CMOS transmission gates
    - Uses dual-rail complementary signals (PN pairs)
- Computation sequence:
 

Precondition: Output signal  $Q$  is initially at logic 0  
By design, driving signal  $D$  is also initially logic 0

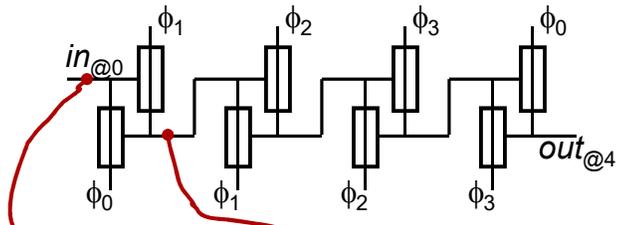
  1. At time 1 (@1), inputs  $A, B$  transition to new levels
    - Connecting  $D$  to  $Q$  if and only if  $A$  or  $B$  is logic 1
  2. At time 2 (@2), driver  $D$  transitions from 0 to 1
    - $Q$  follows it to 1 if and only if  $A$  or  $B$  is logic 1
    - Now  $Q$  is the logical OR of inputs  $A, B$
- Reversible things that we can do afterwards:
  - Restore both  $A, B$  to 0 (latching  $Q$  in place), or,
  - Undo above sequence (decomputing  $Q$  back to 0)



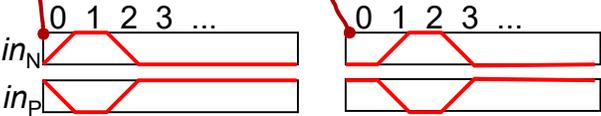
## 2LAL Shift Register Structure



- 1-tick delay per logic stage: Animation: <http://y2u.be/c18mDI0q1IQ>



- Logic pulse timing and signal propagation:

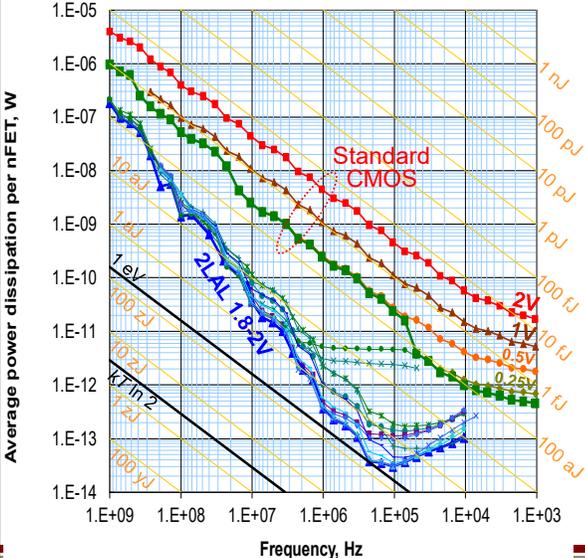


9/14/2018

## Simulation Results (Cadence/Spectre)



**Power vs. freq., TSMC 0.18, Std. CMOS vs. 2LAL**  
 2LAL = Two-level adiabatic logic (invented at UF, '00)



- Graph shows per-FET power dissipation vs. frequency
  - in an 8-stage shift register.
- At moderate freqs. (1 MHz),
  - Reversible uses  $< 1/100^{\text{th}}$  the power of irreversible!
- At ultra-low power levels (1 pW/transistor)
  - Reversible is  $100 \times$  faster than irreversible!
- Minimum energy dissipation per nFET is  **$< 1$  electron volt!**
  - $500 \times$  lower dissipation than best irreversible CMOS!
  - $500 \times$  higher computational energy efficiency!
- Energy transferred per nFET per cycle is still on the order of 10 fJ (100 keV)
  - So, energy recovery efficiency is on the order of 99.999%!
    - Quality factor  $Q = 100,000!$ 
      - Note this does not include any of the parasitic losses associated with power supply and clock distribution yet, though

16

## Design Automation for Adiabatic Circuits—work w. Wille & Zulehner (JKU)



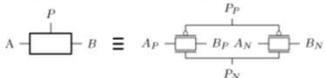
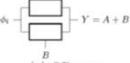
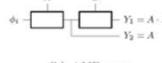


Fig. 1. Transmission gate for dual-rail signals

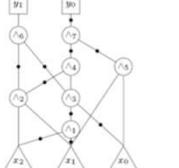


(a) OR gate

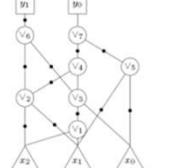


(b) AND gate

Fig. 2. Adiabatic gates

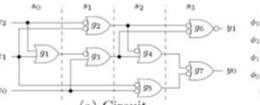


(a) AND-Inverter graph

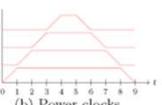


(b) OR-Inverter graph

Fig. 3. Graph representations for Boolean functions



(a) Circuit



(b) Power clocks

Fig. 4. Synthesized retractile circuit

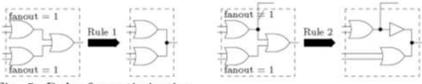


Fig. 5. Rules for optimization

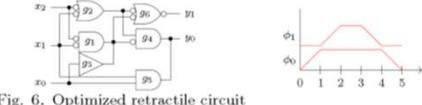
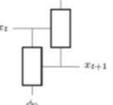


Fig. 6. Optimized retractile circuit

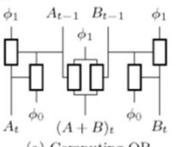


(a) Transmission gates

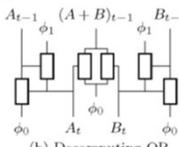


(b) Clocks

Fig. 7. Buffer element for fully-pipelined circuits



(a) Computing OR



(b) Decomputing OR

Fig. 8. OR gate for fully-pipelined circuits

17

## Resonant Energy-Recovering Power Supplies for Adiabatic Circuits



- An extremely nontrivial, and extremely under-emphasized engineering challenge!
- All existing adiabatic schemes for reversible computing (including the superconducting ones!) rely on a (typically unspecified) external system to deliver precisely-conditioned AC waveforms to drive their adiabatic transitions...
  - Ignoring the problem of how to design these systems to work efficiently (as almost everyone in the adiabatic circuits field does!) essentially just sweeps the entire real energy dissipation problem under the rug!
    - It's *extremely difficult* to design a supply that actually recovers almost the entire signal energy... Engineering-wise, this is almost *the entire problem!*
      - We *already know* (ever since Younis & Knight's CRL, 1993) in principle how to design fully-adiabatic switching circuits; that's not even the hard part... It's the energy recovery part that's hard!

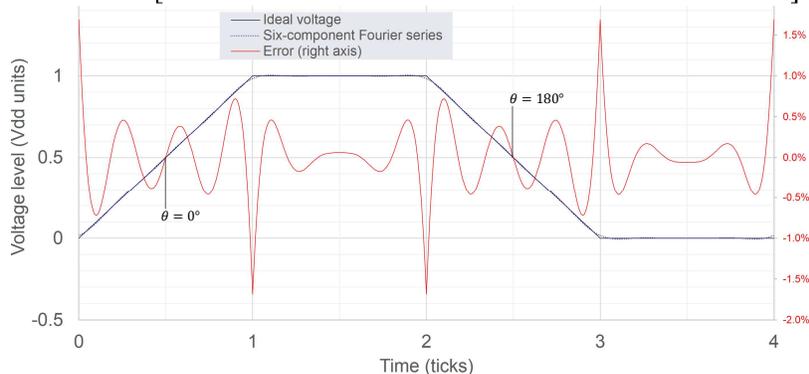
18

## Spectrum of Trapezoidal Wave



- Relative to mid-level crossing, waveform is an odd function
  - Spectrum includes only odd harmonics  $f, 3f, 5f, \dots$
- Six-component Fourier series expansion is shown below
  - Maximum offset with  $11f$  frequency cutoff is  $< 1.7\%$  of  $V_{dd}$

$$v_{f6}(t) = V_{dd} \left[ \frac{1}{2} + \frac{4\sqrt{2}}{\pi^2} \left( \sin \theta + \frac{\sin 3\theta}{3^2} - \frac{\sin 5\theta}{5^2} - \frac{\sin 7\theta}{7^2} + \frac{\sin 9\theta}{9^2} + \frac{\sin 11\theta}{11^2} \right) \right]$$



19

## Resonator design effort, in progress...



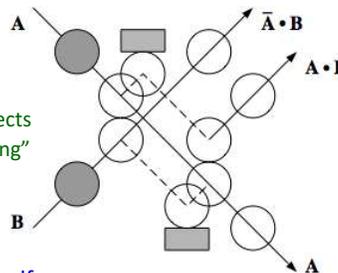
Funding source: DOE ASC (Advanced Simulation and Computing) program

- Goal of this effort:**
  - Design & validate in simulation (and, stretch goal: with a physical prototype) a high-efficiency resonant oscillator (for low-to-medium RF frequencies) that approximates a trapezoidal output voltage waveform
- Initial design concept:**
  - Coupled assemblage of LC tank circuits with resonant frequencies corresponding to odd multiples of the fundamental frequency, excited in the right relative amplitudes to approximate the target wave shape
- Some detailed requirement specifications:**
  - Initial target operating point: 230 kHz, 1.8V (optimal point for minimum dissipation in the UF study) **(MET.)**
    - Explore a wider range of parameter values as the project proceeds
  - Tops and bottoms of trapezoidal wave should be within  $\leq 5\%$  of flatness throughout  $\frac{1}{4}$  clock period. **(MET.)**
  - The 10-90% rise/fall time should be between 75 & 100% of its nominal value (80% of  $\frac{1}{4}$  clock period) **(MET.)**
  - Efficiency goals:**
    - Quality factor of resonator during unpowered ring-down should be  $\geq 1,000$ . **(MET. Measured value:  $\sim 19,550$ .)**
    - Total energy dissipation per cycle during steady-state powered operation should be  $\leq 1\%$  of magnetically-stored energy in the resonator, when the oscillator is running in isolation.
    - Total energy dissipation per cycle during steady-state powered operation should be  $\leq 10\%$  of the capacitively-stored energy on an appropriately-sized model (RC) load, when the oscillator is coupled to the load.
- A number of significant design challenges that have been encountered so far:**
  - How to tune the relative amplitudes of the component resonant modes **(Solved.)**
  - How to prevent phase drift and transfer of energy between modes **(Solved.)**
  - Identifying/tailoring components to have precise-enough  $L, C$  values
  - Designing a driver circuit that meets efficiency goals during steady-state operation
- We have already solved a number of the problems encountered, but still have a ways to go...**
  - We have only spent 1 year/\$250 K on this effort so far.
  - Currently requesting \$330K for next FY.  $\rightarrow$  **Goal for next FY:** Get to a publishable result.

# Ballistic Reversible Computing



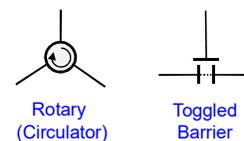
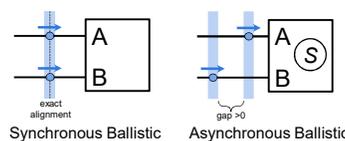
- Original concept:
  - Fredkin & Toffoli's *Billiard Ball Model of computation* ("Conservative Logic," 1982)
    - Based on elastic collisions between moving objects
    - Spawned a subfield of "collision-based computing"
      - Localized pulses/solitons in various media
- No power-clock signals needed!
  - Devices operate when data signals arrive
  - The operation energy is carried by the signal itself
    - Most of the signal energy is preserved in outgoing signals
- However, existing design concepts for ballistic computing invoke implicitly *synchronized* arrivals of ballistically-propagating signals...
  - Making this work in reality presents some serious difficulties, however:
    - Unrealistic in practice to assume precise alignment of signal arrival times
      - Thermal fluctuations & quantum uncertainty, at minimum, are always present
    - Any relative timing uncertainty leads to chaotic dynamics when signals interact
      - Exponentially-increasing uncertainties in the dynamical trajectory
- Can we come up with a ballistic model that avoids these problems?



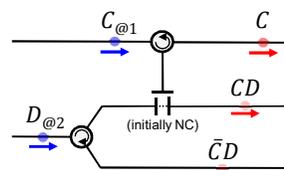
# Asynchronous Ballistic Reversible Computing



- To avoid the problems with dynamical chaos that are inherent to collision-based computing,
  - We must avoid any **direct** interaction between ballistically-propagating signals
- Instead, require temporally-localized pulses to arrive at *distinct*, non-overlapping times
  - Device's dynamical trajectory then becomes *independent* of the precise pulse arrival time
    - Timing uncertainty per logic stage now accumulates *only linearly*, not exponentially
      - Only occasional re-synchronization will be needed
    - To do logic, devices now must have **internal state**
- No power-clock signals, unlike adiabatic designs
  - Devices simply operate whenever data pulses arrive
  - The operation energy is carried by the pulse itself
    - Most of the energy is preserved in outgoing pulses
      - Signal restoration can be carried out incrementally
- A new project has started at Sandia which aims to implement ABRC in superconducting circuits
  - 3-year, \$1.5M internally-funded project



Example ABR device functions



Example logic construction

## Physical realizations of ABRC?



- Of course, to be useful, this model needs to be realized in a specific physical implementation technology that actually provides (nearly) thermodynamically-reversible operation.
  - Need some kind of soliton-like, near-ballistically-propagating pulse,
    - or some sort of particle or quasiparticle (or a larger bound object).
  - Need some physical state variable that can stably maintain at least binary state within the devices
    - for the toggling devices
  - Need a means of physically interacting the pulses with the device states...
    - in ways that can reliably, and almost physically-reversibly, implement at least a universal subset of (probably 2- and 3-terminal) primitive devices.

23

## ABRC in superconducting circuits



- One intriguing possible candidate implementation technology is to use superconducting circuits...
  - SFQ (single flux quantum, or fluxon) pulses on appropriately constructed superconducting transmission lines can carry info. with relatively low dispersion and high propagation velocity (e.g.  $2/3 c$ )
    - Fluxons are naturally quantized by the SQUID-like circuits that produce them, and are naturally polarized (carry 1 bit's worth of +/- polarization state information per pulse)
      - Need to select suitable ABRC primitives operating on arity-2 signals
  - Fluxons trapped in loops (SQUID-like structures) can hold data quiescently
    - Generally, loops hold integer numbers of fluxons in some small range:
      - ..., -2, -1, 0, +1, +2, ...
  - How exactly to implement the reversible interactions?
    - A 3-year, internally-funded project is just starting at Sandia to investigate this...

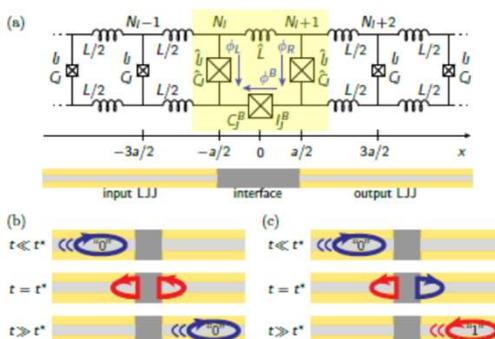
24

# A Very Recent Advance!



Wustman (LPS) & Osborn (JQI) '17 (preprint), "Efficient reversible logic gates without adiabatic constraint: Fluxon resonant scattering with polarity changes"

- The circuit shown at right can be considered as a 2-terminal ABRC device for binary pulses (fluxons)
  - The specified function is to preserve or flip the polarity of a fluxon passing through, depending on device parameters
- Here, the "wires" are LJJ transmission lines
  - Major loss mechanism is resonant plasmon emission
    - With lattice spacing  $0.4\lambda_J$ , fluxon decay time is  $\sim 10^7$  junction switching times given initial  $v = 0.6c$ .

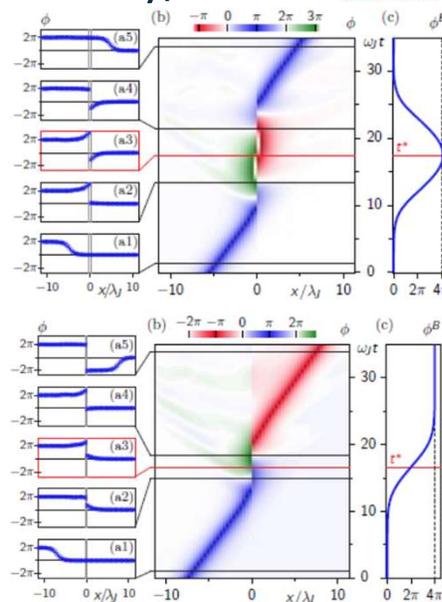


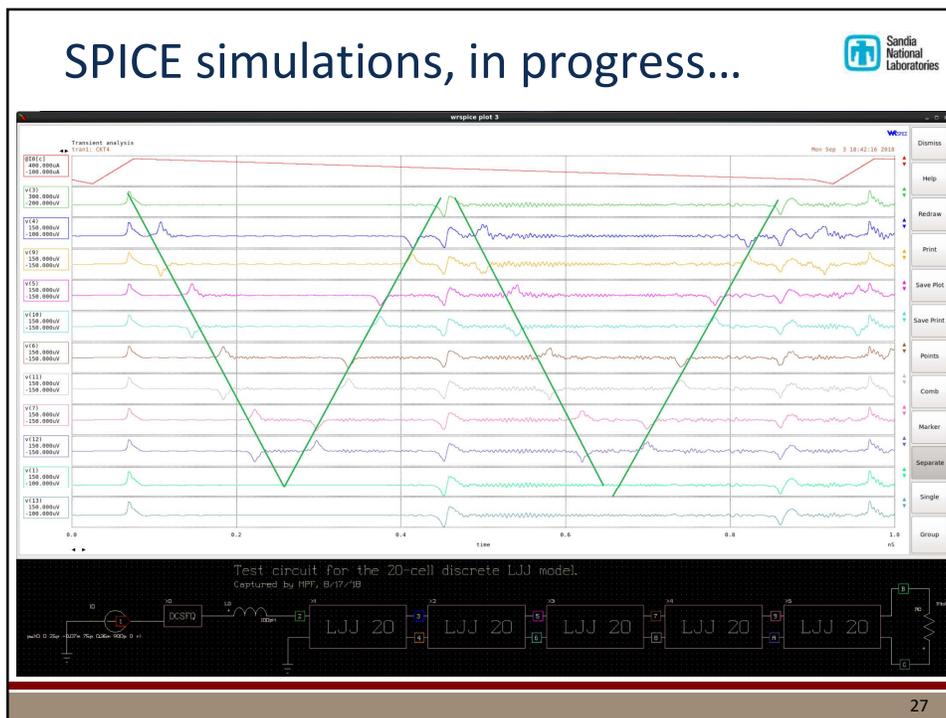
- W&O's paper also describes some more complex (4-terminal) devices
  - Synchronous so far, but they are now starting to explore asynchronous

# W&O's simulation of identity/NOT



- Direct numerical integration of JJ circuit's equations of motion
  - Lagrangian:
 
$$\mathcal{L} = \left(\frac{\Phi_0}{2\pi}\right)^2 \left[ \sum_n \frac{C_{J,n}^A}{2} (\dot{\phi}_n)^2 + \frac{C_{J,n}^B}{2} (\dot{\phi}_n^B)^2 \right] - \left(\frac{\Phi_0}{2\pi}\right) \left[ \sum_n I_{J,n}^A (1 - \cos \phi_n) + I_{J,n}^B (1 - \cos \phi_n^B) \right] - \frac{1}{2} \sum_n [L_n^A (I_n^A)^2 + L_n^B (I_n^B)^2]$$
  - Gives a discrete approximation to sine-Gordon equation:
 
$$\ddot{\phi} - c^2 \phi'' + \omega_J^2 \sin \phi = 0$$
- Scattering interaction at interface is nearly elastic
  - Loss in fluxon velocity of only 4%
    - Loss in energy of 2.1-2.5%





## Conclusion



- A mature reversible computing technology is a prerequisite if we wish to sustain practical performance growth of digital systems over the long term
  - This is guaranteed by irrefutable facts of fundamental physics...
- However, the *engineering* of thermodynamically efficient physical implementations of reversible computing is a field that still very much in its infancy, and, as a research area, is still extremely poorly organized...
  - Far, far more focused work is needed in key areas such as novel device physics for RC, resonator design for adiabatic circuits, and elastic circuits for ballistic computing...
  - The mainstream electronics industry does not appear interested in even attempting to tackle these kinds of problems...
    - Perhaps due to a general perception that approaching RC is too difficult, or even impossible
- The present slow rate of progress would likely be significantly ameliorated by:
  - Improved understanding of the fundamental physics of reversible computing
  - Working demonstrations of useful computations at very low energy dissipation levels
    - Important: While taking the power supply into account!
  - Workshops in key underdeveloped research areas such as reversible device physics
  - Increased support for basic physics & engineering research for reversible computing
- I would encourage funding agencies to dedicate substantial resources to these areas, if they ever want a reversible computing revolution to happen...
  - It's definitely *not* going to happen if everyone just sits around and waits for it!