Special Session: Exploring the Ultimate Limits of Adiabatic Circuits

Michael P. Frank  
Cognitive & Emerging Computing Dept.  
Sandia National Laboratories  
Albuquerque, NM, USA  
mpfrank@sandia.gov

Robert W. Brocato  
RF Microsystems Dept.  
Sandia National Laboratories  
Albuquerque, NM, USA  
orcid:0000-0001-9751-1234

Thomas M. Conte  
Schools of Computer Science and Electrical & Computer Eng.  
Georgia Institute of Technology  
Atlanta, GA, USA  
conte@gatech.edu

Anirudh Jain  
School of Computer Science  
Georgia Institute of Technology  
Atlanta, GA, USA  
anirudh.j@gatech.edu

Nancy A. Missert  
Nanoscale Sciences Department  
Sandia National Laboratories  
Albuquerque, NM, USA  
orcid:0000-0003-2082-2262

Karpur Shukla  
Laboratory for Emerging Technologies  
Brown University  
Providence, RI, USA  
orcid:0000-0002-7775-6979

Alexander H. Hsia  
Sandia National Laboratories  
Albuquerque, NM, USA  
requescat in pace

Brian D. Tierney  
Radiation Hard CMOS Technology Dept.  
Sandia National Laboratories  
Albuquerque, NM, USA  
bdtierney@sandia.gov

Abstract—The field of adiabatic circuits is rooted in electronics know-how stretching all the way back to the 1960s and has potential applications in vastly increasing the energy efficiency of future computing. But now, the field is experiencing an increased level of attention in part due to its potential to reduce the vulnerability of systems to side-channel attacks that exploit, e.g., unwanted EM emissions, power supply fluctuations, and so forth.

In this context, one natural question is: Just how low can the energy dissipation from adiabatic circuits, and the associated extraneous signal emissions, be made to go? We argue that the ultimate limits of this approach lie much farther away than is commonly appreciated. Recent advances at Sandia in the design of fully static, fully adiabatic CMOS logic styles and high-quality energy-recovering resonant power-clock drivers offer the potential to reduce dynamic switching losses by multiple orders of magnitude, and, particularly for cryogenic applications, optimization of device structures can reduce the standby power consumption of inactive devices, and the ultimate dissipation limits of the adiabatic approach, by multiple orders of magnitude as well.

In this paper, we review the above issues, and give a preliminary overview of our group’s activities towards the demonstration of groundbreaking levels of energy efficiency for semiconductor-based logic, together with a broader exploration of the ultimate limits of physically realizable techniques for approaching the theoretical ideal of perfect thermodynamic reversibility in computing, and the study of the implications of this technology direction for practical computing architectures.

Keywords—Adiabatic circuits, CMOS, physical limits of computing, low-power design, reversible computing

I. INTRODUCTION

In 1961, Rolf Landauer of IBM observed that there is a fundamental physical limit to the energy efficiency of logically irreversible computational operations, meaning, those that lose information to their thermal environment [1]. This observation, now enshrined as Landauer’s Principle, is today understood (with some important qualifications) as a rigorous theorem of fundamental statistical physics and information theory [2], and is widely considered to be one of the most important results in the fundamental physics of computing.

During the 1970s, Landauer’s protégé Charles Bennett realized that an alternative paradigm for computation based on logically reversible operations, which could in principle sidestep Landauer’s efficiency limit, was theoretically coherent [3], and went on to develop a more sophisticated understanding of the thermodynamics of computation built on this insight [4], [5]. Early concepts for engineering implementations of reversible computing (e.g., [6], [7]) eventually led to Younis & Knight’s discovery (in the early 1990s) that it could even be achieved for pipelined sequential CMOS, using their CRL [8] and SCRL [9] logic styles. CRL was the first complete sequential CMOS logic family that is what I call truly, fully adiabatic, meaning that, at least in an idealized, no-leakage limit, and when parasitic capacitances are negligible, in principle the dissipation per logic operation can be made arbitrarily small—even below Landauer’s limit of $E_{\text{diss}} \geq kT \ln 2$ for irreversible bit erasures. Due to Landauer’s Principle, we know that any physical mechanism capable of beating this limit must be logically reversible—although what that means, precisely, is misapprehended in much of the literature (see [10] for a correct account).

Note that avoiding energy dissipation includes avoiding energy emissions associated with side-channel attacks such as Differential Power Analysis (DPA); this led to attention being paid to the possible security applications of adiabatic circuits and reversible computing as early as 2006 [11]. However, existing implementations of adiabatic circuits typically remain far from the reversible ideal, for several reasons that we will discuss.
In 2017, an effort informally termed the Adiabatic Circuits Feasibility Study (ACSF) started up at Sandia National Laboratories with the goal of pushing the limits of energy efficiency in physical implementations of adiabatic CMOS circuits, as a part of an effort to assess the long-term feasibility of these techniques for possible use in future high-performance computing (HPC) systems. Although the main goal of this effort is to attain more power-efficient (and ultimately, more cost-efficient) computations, improving security can be a potentially beneficial side effect, as previously mentioned.

II. TECHNICAL APPROACH

As mentioned above, to avoid limiting energy efficiency (i.e., to approach zero energy dissipation per operation), the physical operation of our circuits must be “truly, fully adiabatic,” which requires us to apply the principles of reversible computing, albeit in the generalized sense discussed in [10]. Ignoring leakage for the moment, even just to approach reversible, adiabatic operation in a switching circuit requires that the following fundamental “design rules” must be obeyed [12]:

1) Avoid passing current through diodes. This is due to the “diode drop,” an intrinsic voltage drop that exists in a diode, which is unavoidable for fundamental thermodynamic reasons.

2) Avoid turning a switch (such as a FET) “on” when there is a (non-negligible) voltage difference \( V_{DS} \) between its drain and source terminals. Such an event would result in a substantial, non-adiabatic flow of current.

3) Avoid turning a switch “off” when there is a non-negligible current \( I_{DS} \) between drain and source, unless these nodes remain connected along an alternate path. The reasoning behind this rule is more subtle, and so it is not widely recognized, and is broken by many designs. Nevertheless, obeying this rule is essential for approaching physical reversibility [13].

Rules 2 & 3 together imply that transitions in fully adiabatic circuits must be driven by quasi-trapezoidal supply waveforms (i.e., having flat tops and bottoms), since the supply voltage presented to a device must remain stable while the device is being turned on or off. But many so-called “adiabatic” circuit designs in the literature invoke sinusoidal supply waveforms, ergo aren’t truly, fully adiabatic. However, to supply a flat-topped waveform resonantly (with high-quality energy recovery) presents a significant RF engineering challenge.

But, since our goal in this project is to push the limits of energy efficiency in adiabatic circuits, we are tackling the challenges head-on. Therefore, our overall technical strategy is:

1) Implement truly, fully adiabatic (ergo reversible) CMOS logic families, obeying all of the adiabatic design rules.

2) Design & implement high-quality resonant oscillators generating multiple near-ideal, flat-topped, phase-locked quasi-trapezoidal waveforms to serve as the power-clock supplies.

3) Test the adiabatic logic, driven by the resonator, measure the system power dissipation, and further refine the design.

Further, we aim to characterize the energy efficiency gains achievable in our approach across a range of available CMOS processes, to better understand which processes will afford the best possible energy efficiency using our methods.

In parallel with this core engineering effort, the scope of our effort has expanded over the last year to also include two university-based collaborations on related topics:

1) Karpur Shukla of Brown University is applying theoretical tools from modern non-equilibrium quantum thermodynamics to derive fundamental, technology-independent lower limits on energy dissipation as a function of delay for general quantum systems performing classical reversible computations. This study may help us develop new insights into the design of innovative “Beyond CMOS” device technologies for reversible computing, so as to aspirationally achieve breakthrough reductions in energy-delay product, and take the cost-efficient performance of reversible computing far beyond the limits of any possible non-reversible technology.

2) Tom Conte and Anirudh Jain of Georgia Tech aim to carry out a study of processor architectures based on reversible technology, including an assessment of the potential system-level impact of hypothetical new reversible device technologies for computational science applications.

III. PROGRESS TO DATE AND EARLY RESULTS

A. Logic Family, Simulations, and Test Chip Layouts

As mentioned in the previous section, our approach is based on implementing truly, fully adiabatic CMOS logic styles. The first of these, historically, was Younis & Knight’s Charge Recovery Logic (CRL) [8]. A number of generalizations and variations on the techniques innovated by CRL were developed in subsequent literature (e.g., [9], [14], [15]). We selected for our study a version called 2LAL (two-level adiabatic logic) which is effectively just a slight generalization of CRL but was reinvented by one of us (Frank) in 2000. It is an attractive target for study because of its simple clock waveforms, which have the shortest possible period for a fully adiabatic sequential logic family (only 4 adiabatic transition times per clock cycle), and thus, 2LAL’s trapezoidal waveforms are the closest to an ordinary sinusoidal wave, facilitating design of resonant oscillators.

In addition, Cadence/Spectre simulations carried out at the University of Florida in 2004 for a TSMC 180 nm process suggested that 2LAL could be extremely efficient, capable of dissipating as little 1 eV (0.16 aJ) per device per cycle. Our initial aim was to try to replicate that result using contemporary device models for similar processes available today. However, careful reexamination of files from the earlier study revealed that it had used BSIM3 MOSFET models that neglected gate leakage, and thus likely underestimated the true minimum energy dissipation for 2LAL circuits implemented in that process. This then led us to consider new ideas for overcoming the leakage problem, which we will discuss in the next section.

In the meantime, we decided to update the 2LAL simulation results using newer models (e.g., BSIM3S01, BSIM4) for available processes, while also cross-validating those models against detailed finite-element (TCAD) models, where possible. So far, our simulation efforts have focused on the in-house 350 nm and 180 nm processes provided by the MESA fabrication facility at Sandia and a 180 nm SOI process from GlobalFoundries. ADS simulations suggest a minimum energy dissipation for the 180 nm processes on the order of ~1 keV (160 aJ) per cycle per FET (at a 50% activity factor) when running at speeds around 1 MHz.
To further validate our simulation results, we decided to lay out, fabricate, and measure energy dissipation of an actual test chip. To our knowledge, physical chips based on CRL/2LAL have never been built,\(^1\) so this seemed worthwhile. We laid out a 720-stage 2LAL shift register in our internal 180 nm process (Figs. 1 & 2), to be taped out to a shuttle run in Aug. 2020.

One issue encountered in our simulation work was that the dynamic nature of logic styles such as 2LAL (and the original CRL) creates difficulties for the correct and energy-efficient operation of these circuits. In general, floating nodes are vulnerable to voltage drift (due to leakage) and sag (due to capacitive voltage-divider effects involving parasitic couplings). If, due to drift/sag, a floating node voltage diverges by \(\Delta V\) from its nominal level, then, upon reconnection of that node to its reference level, there will be a sudden, non-adiabatic energy dissipation of \(C(\Delta V)^2/2\). Since voltage drift tends to increase linearly with time, this leads to a component of total per-cycle energy dissipation that increases quadratically with clock period. In addition, drift causes logical errors at sufficiently low frequencies, or in any node that remains in the floating state for sufficiently many cycles due to a low activity factor in that part of the circuit.

Confronting these issues makes plain that a fully static, fully adiabatic (“perfectly adiabatic”) logic style is required to maximize energy efficiency in adiabatic circuits. This also facilitates testing, since, in a static style, one can even pause the clocks. In June 2020, we invented one such technique, called S2LAL (for static 2LAL), which will be presented in a separate paper.

B. Power-Clock Resonator Circuit

One of the key challenges being tackled in this project is the design and development of a high-quality resonant oscillator suited to drive a fully adiabatic circuit. Our initial target is the simple 4-phase trapezoidal waveform \(\tilde{\phi}_4\) in Fig. 1. Early concepts for resonantly generating trapezoidal waveforms [16] were based on combining subcircuits that resonate at respective frequency components of the desired signal. For example, the first six terms of the Fourier series for \(\tilde{\phi}_4\) are

\[
v_{f_0}(t) = V_{DD} \left[ \frac{4}{\pi^2} \sin(\omega t) + \frac{4}{3\pi^2} \sin(3\omega t) + \frac{4}{5\pi^2} \sin(5\omega t) - \frac{4}{7\pi^2} \sin(7\omega t) + \frac{4}{9\pi^2} \sin(9\omega t) + \frac{4}{11\pi^2} \sin(11\omega t) \right].
\]

\[\text{Fig. 1. Shift register schematic and timing sequence in the 2LAL logic style (a variant of CRL). Each signal } x_i \text{ in the schematic is implicitly a dual-rail pair of complementary signals, an active-high signal denoted } \bar{x}_i \text{ and active-low } x_i. \text{ Each rectangle denotes a parallel pair of CMOS transmission gates. A timing diagram for power-clocks } \phi_i \text{ and data signals } S_i \text{ is shown at right. This shift register transfers a single logical symbol (e.g., logic 1) and an identical, parallel structure would be needed to also transfer an alternate symbol (e.g., logic 0).}
\]

which together suffice to come within 2% of the ideal waveform, in theory. But, simple approaches such as those in [16] are insufficient to phase-lock the various sinusoidal components together and maintain an appropriate amplitude distribution. So, we invented an improved circuit technique that solves these problems. In simulations, our design attains \(Q\) values of ~3,000. Sample output waveforms after oscillator startup are shown in Fig. 3 below.

IV. NEXT STEPS

Important next steps for this project, over the coming year, include the following:

1) Simulations, layouts and test chip fabrication for the newly invented fully static version of 2LAL (S2LAL).

2) Further development of the trapezoidal resonant oscillator circuit, including experimental verification that prototype implementations can drive the 2LAL test chip, and measurements of system power dissipation.

3) Exploration of additional available processes as implementation candidates, including MOSIS MPW offerings.

---

\(^1\) In [14], Athas et al. described a shift register test chip for a logic style that is very similar to, but not exactly the same as, CRL/2LAL; also, that one was a much smaller design for a significantly older process.
For the longer term, to maximize the energy efficiency of adiabatic circuits, it will be crucial to reduce leakage. We see two potentially viable routes to this:

1) *Use an older process.* Perhaps ironically, *older* CMOS technologies can actually be *more* energy efficient than newer ones when perfectly adiabatic design styles are used and energy dissipation is aggressively minimized. This is because subthreshold leakage is much lower (since thresholds are higher), and also gate leakage is much lower (since gate oxides are thicker). These are both exponential effects, and thus overwhelm the polynomial factors in adiabatic minimum energy dissipation that result from increased $C$ and $V$ values.

2) *Operate at cryo temperatures.* Recent research [17]–[19] shows that modern CMOS processes perform quite well at deep cryogenic temperatures (e.g., 4 K, and perhaps lower), and moreover, subthreshold leakage is much lower than at room temperature (due to increased subthreshold slope). Due to this, device stacks can be reoptimized in several ways to dramatically reduce overall minimum energy dissipation at cryogenic temperatures, particularly when adiabatic switching is used. Although more research is needed, it is plausible that efficiency gains achievable via such methods could outweigh the substantial power overhead for cryogenic cooling ($\approx 400 \times$ to go from 4 K to room temperature). Also, cryogenic operation raises the possibility of building ultra-high-$Q$ *superconducting* resonators, or simply leaving the resonators at room temperature to reduce cooling overhead. Finally, superconducting *interconnects* offer another means to greatly improve energy efficiency of cryogenic CMOS-based systems with adiabatic switching.

V. CONCLUSION

Careful consideration of both the fundamental and practical limits on the levels of energy efficiency achievable by means of adiabatic circuit techniques makes it very clear that these limits are, today, still very far from being reached. Our work is illuminating several clear and necessary steps along the path towards maximizing the energy efficiency of adiabatic CMOS:

1) *Utilize truly, fully adiabatic reversible logic styles, and beyond that, fully static, “perfectly adiabatic” logic styles.*

2) *Develop high-$Q$ trapezoidal resonators, which are perfectly suited for driving perfectly adiabatic logic circuits.*

3) *Minimize leakage aggressively, through e.g., cryogenic operation and associated reoptimization of the device design.*

In the long term, as adiabatic CMOS becomes ever more efficient, there will be increasing demand for 3D fabrication processes that provide ever-more layers of active devices at an ever-lower cost per device, resulting in ever-greater levels of overall system cost efficiency, taking both manufacturing cost and power dissipation related costs into account.

Further, in the very long run, exploration of the fundamental quantum thermodynamic limits of reversible computing can lead to the discovery of completely new types of physical devices and mechanisms of operation that could lead to breakthrough levels of energy efficiency as a function of speed, perhaps many orders of magnitude beyond anything achievable with CMOS.

Landauer’s Principle assures us that achieving breakthroughs of such a magnitude will require reversible computing.

ACKNOWLEDGMENT

Thanks are due to Erik DeBenedito for calling our attention to the recent line of research [17]–[19] on cryogenic CMOS.

REFERENCES


