Enhancing Fundamental Energy Limits of Field-Coupled Nanocomputing Circuits

Jeferson F. Chaves*†, Marco A. Ribeiro†, Frank Sill Torres† and Omar P. Vilela Neto†
*Centro Federal de Educação Tecnológica de Minas Gerais - CEFET-MG, jeferson@decom.cefetmg.br
†Universidade Federal de Minas Gerais - UFMG, marcoantonio@dcc.ufmg.br, franksill@ufmg.br, omar@dcc.ufmg.br

***Abstract*—Energy dissipation of future integrated systems, consisting of a myriad of devices, is a challenge that cannot be solved solely by emerging technologies and process improvements. Even though approaches like Field-Coupled Nanocomputing allow computations near the fundamental energy limits, there is a demand for strategies that enable the recycling of bits' energy to avoid thermalization of information. In this direction, we propose a new kind of partially reversible systems by exploiting fan-outs in logic networks. We have also introduced a computationally efficient method to evaluate the gain obtained by our strategy. Simulation results for state-of-the-art benchmarks indicate an average reduction of the fundamental energy limit by 17% without affecting the delay. If delay is not the main concern, the average reduction reaches even 51%. To the best of our knowledge, this work presents the first post-synthesis strategy to reduce fundamental energy limits for Field-Coupled Nanocomputing circuits.**

I. Introduction

Recent studies indicate that there might be only four to eight generations left before reaching the energetic boundary for Complementary Metal-Oxide Semiconductor (CMOS) technology [1], [2]. Despite the tremendous progress of integrated technologies, and even with the uprising of promising emerging technologies, the advancement of future systems could be strongly restricted by energy dissipation issues. Surprisingly, the cause of this problem is not a device property, but a direct consequence of the Second Law of Thermodynamics called Landauer's limit [3]. This is the fundamental energy limit and it is related to the information loss in the form of heat, which occurs when bits are irreversibly erased within logic operations. It is on the order of k_BT Joules per erased bit, where k_B is the Boltzmann constant and T is the temperature of the system's thermal environment. Despite the controversies around it, Landauer's idea was experimentally verified, confirming that there is a physical limit in irreversible computation [4]–[7].

CMOS technology can never reach Landauer's minimum due to the Landauer-Shannon limit that is about $50\times$ higher than Landauer's limit [8]. However, several emerging technologies, like Quantum-dot Cellular Automata (QCA) or Nanomagnetic Logic (NML), are not restricted by the Landauer-Shannon bound [6], [9]. Consequently, the fundamental energy limits of these nanotechnologies are defined by irreversible logic operations.

The main contribution of this paper is the exploration of the architecture of Field-Coupled Nanocomputing (FCN) circuits in order to replace conventional logic gates by its reversible equivalents such that the fundamental energy limit is reduced. The proposed algorithm permits the designer to prioritize energy dissipation or delay and has been implemented and tested for QCA circuits. To the best of our knowledge, this work presents the first post-synthesis strategies to reduce fundamental energy limits for FCN designs.

The remaining of this paper is organized as follows. Section II introduces QCA and Reversible Computing techniques. Section III presents the proposed approach, while Section IV discusses the simulation results for state-of-the-art benchmarks [10]. Finally, Section V concludes this work.

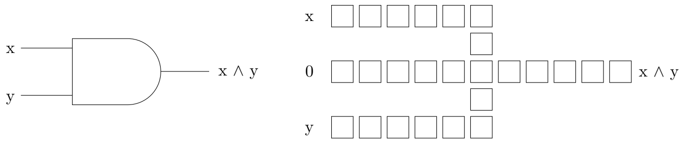
II. Background

A. Quantum-dot Cellular Automata

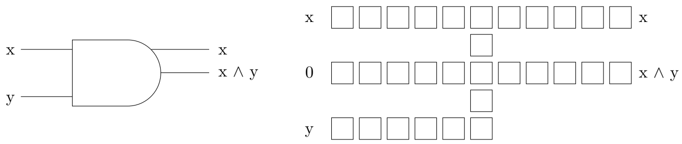
QCA is a Field-Coupled nanotechnology and a potential alternative to traditional CMOS technologies [9], [11]–[21]. Each QCA cell consists of four quantum dots, which are structures able to confine electric charges. These quantum dots are arranged in a square-like fashion such that free mobile electrons can move between them. Since these electrons impose mutual repulsion due to Coulomb interaction [11], they tend to locate themselves at opposite corners of the square. Thus, these electrons assume stable states, called polarizations, which are energetically equal and interpreted as binary 0 and 1. Further, when placed close to each other, the polarization of one QCA cell influences the polarization of the other–again by Coulomb interaction. The exploitation of this effect allows for the realization of logic gates.

B. Reversible Computing Techniques

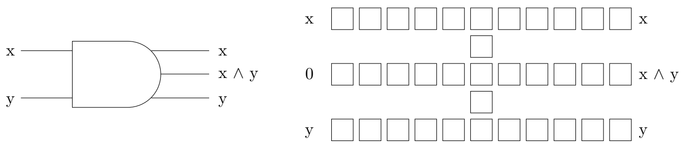
A system is physically reversible if it can return to any previous state in reverse order as explained by Lent et al [12]. Nevertheless, turning a circuit reversible demands modifications of its architecture that usually penalize timing and/or area [22]. Hence, it is relevant to find a balanced trade-off between reduction of energy losses due to irreversible logic operations, and costs regarding area and delay. This requires a formal quantification of those losses, as proposed by Landauer [3]. He showed that the fundamental energy dissipation appears when devices irreversibly erase information, i. e., bits. His observation results from a statistical mechanics' argument, which reveals that this loss is the result of the locally merging of states, happening within the conventional logic gates, e. g., *AND*, *OR* or *Majority*. In these gates, n-bit inputs are merged in a certain way to produce a 1-bit output.



(a) Conventional AND

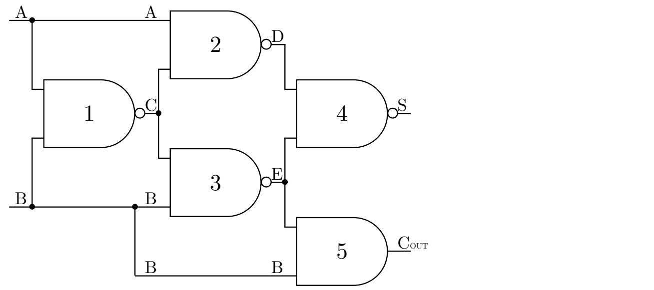


(b) AND recycling energy of one input

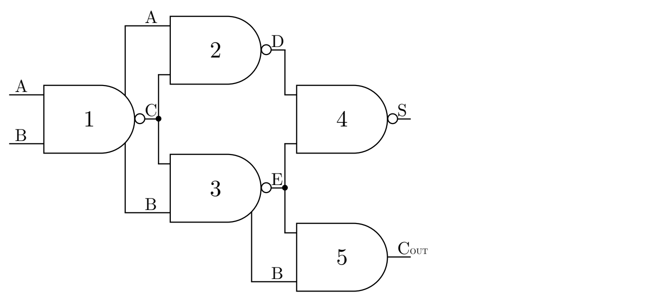


(c) Fully reversible AND recycling energy of both inputs

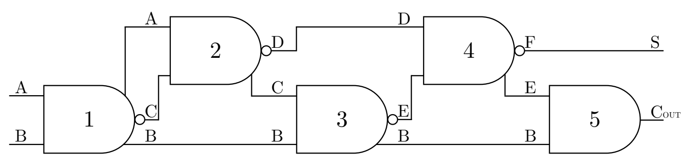
Fig. 1: Symbols and layouts of conventional as well as n-bits recycling QCA gates [12]



(a) Conventional Half-adder



(b) Half-adder exploiting fan-outs without degrading delay



(c) Half-adder exploiting fan-outs with delay penalty

Fig. 2: Application of proposed method for Half-Adder circuit.

The calculation of the Landauer limit is often erroneously understood. This led to recent efforts by the research community towards clarification of this energy assessment [23]–[26]. The minimum loss is the addition of the difference between Shannon's entropies of gate's input(s) and output(s). The main implication is that irreversible functions can be conditionally reversible with an adequate input combination subset (one that guarantees a 1-to-1 relation between input(s) and output(s)) [26].

A decade after Landauer's groundbreaking work, Bennett presented a way to handle Landauer's limit [27]. He showed that any logically irreversible function can be embedded in a logically reversible circuit. Lent and collaborators implemented this idea in the QCA technology by changing the clock timing [12]. However, this technique comes with the cost of strong degradation of the design throughput. Based on this initial approach, Ottavi et al. [15] proposed an intermediate solution that allows balancing energy dissipation and throughput by adding memory stages and a pipeline-like control scheme. Nevertheless, there remains a considerable performance penalty.

In their work, Lent et al. indicated a different approach in order to exploit Bennett's idea [12]. The authors proposed echoing the inputs of QCA gates to the output, turning these gates reversible. This shall be explained by help of Fig. 1, which depicts three different implementations of a QCA AND gate: conventional, recovering information of one input (1-bit recycling gate) and fully reversible (2-bit recycling gate).

Lent et al. showed in [12] that the conventional QCA AND (see Fig. 1a) dissipates energy above $k_B T \ln(2)$ Joules, when the value of one input differs from the output. However, if the input that differs from the output is echoed, as in the case of the 1-bit recycling (Fig. 1b) and fully reversible (Fig. 1c) versions, the dissipation remains below $k_B T \ln(2)$ Joules. The drawbacks of this method are an increased complexity of the circuit and the requirement to treat intermediate results accordingly.

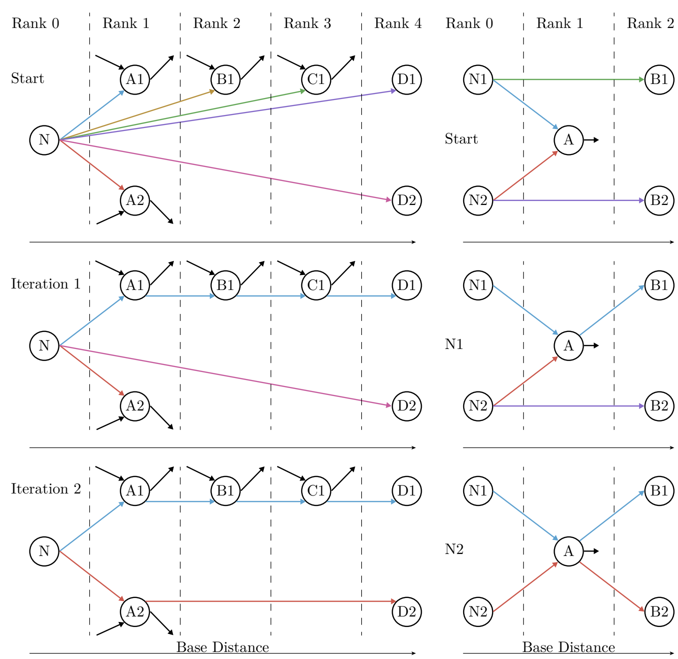
III. Reducing Fundamental Energy Limits

Unfortunately, the evaluation of Landauer's limit comes at high computational cost with exponential time and space complexity, turning it ineligible for comprehensive designs. Thus, we propose a simpler but practicable approach. The main idea of our method is the summation of the upper loss for each gate in the design, i. e., the maximum entropy that a gate can receive. Note that by this choice we are overestimating the losses and underestimate the gains since we are not considering conditionally reversible cases. We consider that this is a fair trade since we are now solving a polynomial problem instead of an exponential one and, thus, enable the evaluation of large circuits.

In case of conventional gates, e. g., the one depicted in Fig. 1a, we count the number of variable input bits as losses. In the other cases we consider each embedded fan-out as a recycled bit and are not considered as losses. Based on this method, we propose the exploitation of signals that are used more than once in the circuit. This gives us the ability to add gates that recycle input bits, in a way that we can reduce the

Algorithm 1: Building chains of n-bits recycling gates

```
  data : netlist ← circuit's netlist
1 foreach fo ∈ sort_ranks_desc(netlist.fanouts) do
2     repeat
3         chain ← ∅
4         foreach rank ∈ fo.ranked_children_asc do
5             children ← choose(rank)
6             chain ← chain ∪ children
7             rank ← rank ∖ children
8         fo.make_chain(chain)
9     until |chain| ⩽ 1
```



(a) Generation of chains of 1-bit recycling gates (b) Integration of fully reversible gates

Fig. 3: Operation of algorithm for two circuits

fundamental energy limit.

Following example shall detail the application of this approach.

Example 1. *Fig. 2a shows a half-adder, which solely applies conventional QCA gates. The fundamental energy loss of the whole circuit results from the summation of the information loss of each gate. This circuit dissipates 3.76 bits following Landauer's method [23], while our approach approximates a loss of 10 bits. The circuit delay, which is directly related to the number of stages, is equal to three.*

Looking at the circuit, one can notice that some signals, i. e., A *and* B*, feed gates in different stages. Hence, instead of using a fan-out structure that multiplies the signal, one can feed it into gates that recycle its input's energy and propagate the signal throughout them to the following stages. The resulting circuit is shown in Fig. 2b. Here, gate* 1 *has been exchanged by its fully reversible equivalent that passes both inputs* A *and* B *to the respective following gates* 2 *and* 3*. Further, gate* 3 *has been exchanged by its* 1*-bit recycling version, which passes the input signal* B *to the following gate* 5*. According to our method, the fundamental energy loss reduces to* 7 *bits compared to* 1.88 *bits estimated by Landauer's technique. Comparing these results with those for Fig. 2a, one sees that the gain with Landauer's evaluation is 1.46 against 1.42 in the approximated evaluation. The delay of this circuit is identical to the initial version.*

If the delay is of lower priority, one can exploit that still some signals are inputs to more than one gate, i. e., signals C *and* E*. Thus, instead of processing a signal in parallel one can serialize the access and apply it as an echoed input to gates that recycle inputs. Fig. 2c depicts the resulting circuit, in which signal* C *passes through the* 1*-bit recycling gate* 2 *before entering gate* 3*. Also, signal* E *traverses gate* 4 *before being an input to gate* 5*. This additional modification comes at the costs of an increased delay from three to five stages. However, this reduces the fundamental energy loss down to* 5 *bits (our method) and* 0.69 *bit (Landauer's approach), respectively. Comparing these results with those for Fig. 2a, one sees that the gain with Landauer's evaluation is 5.44 against 2.11 in the approximated evaluation.*

Algorithm 1 implements the proposed method for synthesized circuits given as netlist. It starts by iterating over all fan-outs, which are sorted in descending order by its distance from the primary inputs, i. e., its rank (line 1). For each of these fan-outs, the algorithm builds chains of gates by selecting unique children from the fan-out's outputs sorted by their rank in ascending order (lines $4 \sim 7$).

Function `choose` enables the definition of gate's priorities, e. g., 1-bit recycling reversible gates over conventional ones, and the restriction of modifiable gates. If energy shall be prioritized, this function may select multiple gates with same rank. After selection of the children, the algorithm converts them to a chain by linking their inputs and outputs together (line 8). The executions stops when the chain has only one node left (line 9), since at this point the algorithm cannot build any other chain.

The examples in Fig. 3a demonstrate the operation of the algorithm for two benchmark circuits developed by Zografos et al. [28].

Example 2. *In the initial version of the circuit depicted in Fig. 3a, the output of node* N *connects to all remaining nodes. During the first iteration, the algorithm selects and merges nodes* A1*,* B1*,* C1 *and* D1 *into a chain that starts at node* N *(marked as blue lines). Further, nodes* A1*,* B1 *and* C1 *are changed to its* 1*-bit recycling versions. In the following iteration, a second chain consisting of nodes* A2 *and* D2 *is created (marked as red lines), and node* A2 *is modified to a* 1*-bit recycling gate.*

The circuit depicted in Fig. 3b illustrates how fully reversible gates are integrated. In the first iteration, the connection between nodes N1 *and* B1 *(blue lines) is placed between nodes* A *and* B1 *and node* A *is modified to a* 1*-bit recycling gate. In the second iteration, the connection between nodes* N2 *and* B2 *(red lines) is moved between nodes* A *and* B2 *and node* A *is changed to a fully reversible gate.*

TABLE I: Benchmark results

Benchmark	Size	Inputs	Outputs	Original		Delay-oriented		Energy-Oriented			
				Delay^a	Energy^b	Energy^b	Diff.^{c}	Energy^b	Diff.^{c}	Delay^a	×Increase^c
router	113	60	30	9	242	212	−12%	171	−29%	14	1.56
priority	498	128	8	10	1060	920	−13%	620	−42%	21	2.10
voter	14 078	1001	1	60	29 313	25 400	−13%	15 464	−47%	132	2.20
arbiter	2179	256	129	14	4945	3792	−23%	2518	−49%	88	6.29
sqrt	52 344	128	64	690	111 752	84 935	−24%	56 889	−49%	6027	8.73
max	7202	512	130	27	16 056	13 397	−17%	8121	−49%	1216	45.04
i2c	1108	147	142	9	2285	1847	−19%	1140	−50%	114	12.67
bar	3054	135	128	14	6108	6108	0%	3061	−50%	1027	73.36
sin	7895	24	25	91	17 265	13 982	−19%	8678	−50%	607	6.67
mem_ctrl	10 362	1204	1231	18	21 067	17 158	−19%	10 445	−50%	457	25.39
square	19 201	64	128	36	39 711	34 702	−13%	19 922	−50%	391	10.86
log2	37 584	32	32	181	78 889	63 860	−19%	39 790	−50%	1654	9.14
adder	2982	256	129	12	7332	5813	−21%	3611	−51%	72	6.00
mult	41 936	128	128	111	92 915	77 351	−17%	45 130	−51%	1477	13.31
int2float	246	11	7	9	529	420	−21%	256	−52%	53	5.89
cavlc	745	10	11	11	1625	1212	−25%	775	−52%	119	10.82
ctrl	127	7	26	5	264	198	−25%	110	−58%	37	7.40
dec	420	8	256	4	840	840	0%	172	−80%	259	64.75
AVERAGE	11 226	228	145	73	24 011	19 564	−17%	12 048	−51%	765	17.34

^{a} Delay is given as number of stages.
^{b} Energy means the approximation of the fundamental energy limit, calculated with the proposed method.
^{c} Compared to the original netlist.

IV. Simulation Results

We applied our approach to the EPFL benchmarks used by Testa [29]. The results are summarized in Table I. Here, the first column show the benchmark names. The second, third and fourth columns display the size of the circuit (number of gates) and the amount of inputs and outputs, respectively. The fifth column lists the delay of the original circuits, which is identically to the number of stages. The sixth column presents the energy dissipation of the original circuits, which means the fundamental energy limit calculated with the method proposed in section III. The next two columns refer to the delay-oriented results, i. e., when no delay penalty is allowed. The seventh column shows the achieved fundamental energy limits, while column eight lists the difference to the original versions. The remaining four columns relate to the energy-oriented results, listing the new fundamental energy limits and the reduction compared to the original version as well as the new delay and the increase.

The results indicate that the fundamental energy limits could be reduced by up to 25% and on average by 17% if the delay had to remain constant. If delay was not restricted, fundamental energy limits decreased by to 80% and on average by 51%, at the costs of a maximum delay increase of 73.36× and on average by 17.34×. For both benchmarks *bar* and *dec* no energy reduction was possible if delay was prioritized. In these cases, our method could not use any of their fan-outs. On the other hand, when the delay is not a problem our method was able to reduce the energy for these circuits by 50% and 80%, respectively, at the cost of high delay degradation.

V. Conclusions

We proposed in this work a simple, but practicable approach to estimate energy dissipation by reversible gates in a complex circuit. The previous exact method comes at high computational cost with exponential time and space complexity, turning it ineligible for large designs. Then, we developed a new method for design partially reversible circuits in order to reduce the fundamental energy limits of Field-Coupled Nanocomputing based architectures. Simulation results for an established benchmark suite indicate that the presented technique can reduce the fundamental energy limits on average by 17% without any delay penalty. When delay is not a concern, energy can be decreased by up to 80% and on average by 51%, at an average performance penalty of 17.34×.

Acknowledgement

The authors would like to thank CAPES, CNPq and Fapemig for supporting this work.

References

[1] E. P. DeBenedictis, "The boolean logic tax," *Computer*, vol. 49, no. 4, pp. 79–82, 2016.

[2] D. E. Nikonov and I. A. Young, "Overview of beyond-CMOS devices and a uniform methodology for their benchmarking," *Proceedings of the IEEE*, vol. 101, no. 12, pp. 2498–2533, 2013.

[3] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM journal of research and development*, vol. 5, no. 3, pp. 183–191, 1961.

[4] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, "Experimental verification of landauer's principle linking information and thermodynamics," *Nature*, vol. 483, no. 7388, pp. 187–189, 2012.

[5] A. O. Orlov, C. S. Lent, C. C. Thorpe, G. P. Boechler, and G. L. Snider, "Experimental test of landauer's principle at the sub-kbt level," *Japanese Journal of Applied Physics*, vol. 51, no. 6S, p. 06FE10, 2012.

[6] J. Hong, B. Lambson, S. Dhuey, and J. Bokor, "Experimental test of landauer's principle in single-bit operations on nanomagnetic memory bits," *Science Advances*, vol. 2, no. 3, p. e1501492, 2016.

[7] I. Neri and M. López-Suárez, "Heat production and error probability relation in landauer reset at effective temperature," *Scientific reports*, vol. 6, p. 34039, 2016.

[8] S. Agarwal, J. Cook, E. DeBenedictis, M. P. Frank, G. Cauwenberghs, S. Srikanth, B. Deng, E. R. Hein, P. G. Rabbat, and T. M. Conte, "Energy efficiency limits of logic and memory," in *Rebooting Computing (ICRC), IEEE International Conference on*. IEEE, 2016, pp. 1–8.

[9] J. Timler and C. S. Lent, "Maxwell's demon and quantum-dot cellular automata," *Journal of Applied Physics*, vol. 94, no. 2, pp. 1050–1060, 2003.

[10] L. Amarù, P.-E. Gaillardon, and G. De Micheli, "The EPFL Combinational Benchmark Suite," in *Proceedings of the 24th International Workshop on Logic & Synthesis (IWLS)*, 2015.

[11] C. S. Lent and P. D. Tougaw, "A device architecture for computing with quantum dots," *Proceedings of the IEEE*, vol. 85, no. 4, pp. 541–557, 1997.

[12] C. S. Lent, M. Liu, and Y. Lu, "Bennett clocking of quantum-dot cellular automata and the limits to binary logic scaling," *Nanotechnology*, vol. 17, no. 16, p. 4240, 2006.

[13] M. B. Haider, J. L. Pitters, G. A. DiLabio, L. Livadaru, J. Y. Mutus, and R. A. Wolkow, "Controlled coupling and occupation of silicon atomic quantum dots at room temperature," *Physical review letters*, vol. 102, no. 4, p. 046805, 2009.

[14] J. L. Pitters, L. Livadaru, M. B. Haider, and R. A. Wolkow, "Tunnel coupled dangling bond structures on hydrogen terminated silicon surfaces," *The Journal of chemical physics*, vol. 134, no. 6, p. 064712, 2011.

[15] M. Ottavi, S. Pontarelli, E. P. DeBenedictis, A. Salsano, S. Frost-Murphy, P. M. Kogge, and F. Lombardi, "Partially reversible pipelined QCA circuits: combining low power with high throughput," *Nanotechnology, IEEE Transactions on*, vol. 10, no. 6, pp. 1383–1393, 2011.

[16] M. Kolmer, R. Zuzak, G. Dridi, S. Godlewski, C. Joachim, and M. Szymonski, "Realization of a quantum hamiltonian boolean logic gate on the Si (001): H surface," *Nanoscale*, vol. 7, no. 29, pp. 12 325–12 330, 2015.

[17] D. A. Reis, C. A. T. Campos, T. R. B. Soares, O. P. V. Neto, and F. S. Torres, "A methodology for standard cell design for QCA," in *Circuits and Systems (ISCAS), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 2114–2117.

[18] C. A. T. Campos, A. L. Marciano, O. P. V. Neto, and F. S. Torres, "Use: a universal, scalable, and efficient clocking scheme for QCA," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 35, no. 3, pp. 513–517, 2016.

[19] C. S. Lent, K. W. Henderson, S. A. Kandel, S. A. Corcelli, G. L. Snider, A. O. Orlov, P. M. Kogge, M. T. Niemier, R. C. Brown, J. A. Christie *et al.*, "Molecular cellular networks: a non von Neumann architecture for molecular electronics," in *Rebooting Computing (ICRC), IEEE International Conference on*. IEEE, 2016, pp. 1–7.

[20] J. F. Chaves, M. A. Ribeiro, L. M. Silva, L. M. B. C. de Assis, M. S. Torres, and O. P. Vilela Neto, "Energy efficient QCA circuits design: simulating and analyzing partially reversible pipelines," *Journal of Computational Electronics*, vol. 17, no. 1, pp. 479–489, Mar 2018.

[21] J. Henry and E. P. Blair, "The role of the tunneling matrix element and nuclear reorganization in the design of quantum-dot cellular automata molecules," *Journal of Applied Physics*, vol. 123, no. 6, p. 064302, 2018.

[22] M. Li and P. Vitanyi, "Reversibility and adiabatic computation: trading time and space for energy," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 452, no. 1947. The Royal Society, 1996, pp. 769–789.

[23] I. Ercan and N. G. Anderson, "Heat dissipation in nanocomputing: lower bounds from physical information theory," *Nanotechnology, IEEE Transactions on*, vol. 12, no. 6, pp. 1047–1060, 2013.

[24] N. G. Anderson, "Gate abstractions and reversibility: On the logical-physical link," in *Circuits and Systems (MWSCAS), 2013 IEEE 56th International Midwest Symposium on*. IEEE, 2013, pp. 1059–1062.

[25] E. P. DeBenedictis, M. P. Frank, N. Ganesh, and N. G. Anderson, "A path toward ultra-low-energy computing," in *Rebooting Computing (ICRC), IEEE International Conference on*. IEEE, 2016, pp. 1–8.

[26] M. P. Frank, "Foundations of generalized reversible computing," in *International Conference on Reversible Computation*. Springer, 2017, pp. 19–34.

[27] C. H. Bennett, "Logical reversibility of computation," *IBM journal of Research and Development*, vol. 17, no. 6, pp. 525–532, 1973.

[28] O. Zografos, A. De Meester, E. Testa, M. Soeken, P.-E. Gaillardon, G. De Micheli, L. Amaru, P. Raghavan, F. Catthoor, and R. Lauwereins, "Wave pipelining for majority-based beyond-CMOS technologies," in *2017 Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2017, pp. 1306–1311.

[29] E. Testa, O. Zografos, M. Soeken, A. Vaysset, M. Manfrini, R. Lauwereins, and G. De Micheli, "Inverter propagation and fan-out constraints for beyond-CMOS majority-based technologies," in *VLSI (ISVLSI), 2017 IEEE Computer Society Annual Symposium on*. IEEE, 2017, pp. 164–169.