ENERGY BOUNDS ON PARITY

by

Stephen Cook* and Gloria Kissin

Technical Report #514

July 1988

* University of Toronto, Dept. of Computer Science, Toronto, Canada
ENERGY BOUNDS ON PARITY

Stephen Cook
Department of Computer Science
University of Toronto
Toronto, Ontario, Canada

Gloria Kissin
Department of Computer Science
Technion
Haifa, Israel

1. INTRODUCTION

Energy consumption is a fundamental property of physical computation [MC80], [Ke87]. The energy consumed by a circuit is dissipated as heat. How well a circuit can dissipate heat determines its operational limitations. In VLSI, energy dissipation is a critical concern, accounting for a large percentage of a typical computer's costs [Me86]. Thus, understanding energy consumption is important to VLSI design.

Kissin [Ki87] developed a theory of energy consumption. She defined two classes of models for measuring the energy consumption of VLSI circuits: In the Uniswitch Model (USM), a wire in an acyclic circuit can switch at most once. The Multiswitch Models (MSM) account for race conditions whereby a wire can switch several times before settling down to a final value. The MSM class is a set of models, which provide for the different ways that race conditions (also called hazards) can occur. MSM can also be used to measure the energy of circuits with feedback and pipelined circuits.

In USM, [Ki82] showed a nontrivial lower bound on the energy consumed by monotone circuits (i.e., circuits with no negations). Both [LM81] and [Ki85] derived lower bounds on the uniswitch energy of multiple output functions. For single output functions, [Ki85] describes circuits and layouts for OR, AND and comparison functions, that use linear worst case uniswitch energy.

This note describes the first nontrivial lower bound on the uniswitch energy of a single output function.
The model definitions are in sections 2 and 4. The bounds on parity are derived in sections 3, 4 and the Appendix. Open problems are given in section 5.

2. MODELS and DEFINITIONS

The standard VLSI circuit model is assumed. That is, a VLSI circuit is a labelled directed acyclic graph embedded in the plane. A circuit is in a legal state if every wire has a value that is consistent with the inputs to the node at the tail of the wire and the function labelling that node. The uniswitch energy of a circuit is defined on a pair of legal states, in the following way.

Suppose VLSI circuit $C$ changes (legal) state from $s_0$ to $s_f$, denoted by $C : s_0 \rightarrow s_f$. Further assume that wire $w$ has initial value $s_0(w) = w_0$ and final value $s_f(w) = w_f$ where $w_0, w_f \in \{0, 1\}$. This change in the value of $w$ is denoted by $w : w_0 \rightarrow w_f$. Then $w$ is switched (switches) iff $w_0 \neq w_f$. A wire of length $L$ that switches accounts for $L/k$ switching energy, where $k > 0$ is the length of wire that accounts for 1 unit of switching energy. If $W = \{w\}$ is the set of wires in circuit $C$, and $X$ is the input set such that $C : s_0 \rightarrow X$, then the wire energy, $E_w$, consumed by $C$ is

$$E_w(C, s_0, X) \triangleq \frac{1}{k} \left( \sum_{w \in W} ||w|| \right), \quad \text{where } ||w|| \text{ is the area of wire } w. \quad E_w(C, s_0, X) \leq \frac{1}{k} \text{area}(C) \quad \text{where } \text{area}(C) \text{ is the total wire area of } C.$$

Uniswitch energy is defined below for the worst case and average case. $E_{\text{worst}}^U$ picks the maximum wire energy expended when all (initial state, input) pairs are considered. The average energy model $E_a$, defined below, averages the wire energy expended by a circuit over all (initial state, input) pairs, assuming all legal states are equally likely.

If $C_n$ is a VLSI circuit computing $f_n : \{0, 1\}^n \rightarrow \{0, 1\}^m$ such that $C_n$ is in state $s_0$ at time $t_0$, and $E_w(C_n, s_0, X')$ is the wire energy consumed by $f_n$ when $X = (x_1, \ldots, x_n)$ is the input to $C_n$ at time $t > t_0$, then $E_{\text{worst}}^U(C_n)$, the worst case uniswitch energy, is given by
\[ E_{\text{worst}}^U(C_n) \triangleq \max_{(s_0, X)} E_w(C_n, s_0, X) \]

and \( E_a(C_n) \), the \textit{average case uniswitch energy} is given by

\[ E_a(C_n) \triangleq \sum_{(s_0, X)} E_w(C_n, s_0, X) / 2^{2n} \]

where \( 2^{2n} \) is the number of \((s_0, X)\) pairs.

3. LOWER BOUNDS in USM

The bounds obtained in this section are in USM. Therefore, \( E_{\text{worst}} \) is understood to be an abbreviation of \( E_{\text{worst}}^U \).

Consider the \textit{parity} function on \( n \) boolean variables (ie. \( x_1 \oplus x_2 \oplus \cdots \oplus x_n \)). In the special case where the circuit basis is \( \{ \oplus, \neg \} \), a superlinear energy lower bound for parity is derived in the following theorem, which was independently obtained by J. Leo [Lc84].

\textbf{Theorem 1:}

To compute parity of \( n \) nonconstant inputs with a VLSI circuit \( C \) of area \( A \) requires \( E_a(C) = \Omega(A) \)

\textit{when the circuit basis is \( \{ \oplus, \neg \} \).}

\textbf{Proof:}

Let \( W = \{ w \} \) be the wires of \( C \).

\[ \forall w \in W, \Pr(w \text{ switches}) = 1/2 \therefore E_a(C) \geq \frac{1}{2} A. \]

The definitions of \( E_a \) and \( E_{\text{worst}} \) yield the following Corollary to Theorem 1.

\textbf{Corollary 1:}

To compute parity of \( n \) nonconstant inputs with a VLSI circuit \( C \) of area \( A \) requires \( E_{\text{worst}}(C) = \Omega(A) \)

\textit{when the circuit basis is \( \{ \oplus, \neg \} \).}

The Appendix contains an alternate proof of the worst case lower bound for parity. The alternate proof yields a deterministic polynomial algorithm for computing a pair of states that induces a lot of energy.
Corollary 2:
To compute the parity function on \( n \) boolean variables with a VLSI circuit \( C_n \) of \( O(\log n) \) depth with the I/O ports on a convex boundary of \( C_n \) requires \( E_a(C_n) = E_{\text{worst}}(C_n) = \Omega(n \log n) \) when the basis for \( C_n \) is \( \{\emptyset, \neg\} \).

Proof:
Let \( A \) be the area of \( C_n \). By Theorem 1 and Corollary 1, \( E_a(C_n) = E_{\text{worst}}(C_n) = \Omega(A) \). Since circuit \( C_n \) must fanin the \( n \) inputs, and since nodes have indegree \( \leq 2 \), \( A \) is at least as large as the area of a binary tree on \( n \) leaves. \([BK80]\) and \([Ya81]\) showed that such a tree requires area \( \Omega(n \log n) \) when the depth is \( O(\log n) \) and the leaves are on a convex boundary.

\[
\therefore E_a(C_n) = E_{\text{worst}}(C_n) \geq \frac{1}{2} A = \Omega(n \log n).
\]

4. PARITY in the MULTISWITCH MODELS

In the following, the definitions of \( E_w \) and \( E_{\text{worst}} \) are modified to account for race conditions that are caused by propagation delays and asynchronous input arrival times.

Suppose \( W = \{w_i\} \) is the set of wires in VLSI circuit \( C \) and \( X \) is the input vector. Let \( \Delta \) be the wire delay function that accounts for propagation delays. Let \( I \) be the input schedule that determines when each input bit arrives at an input port. If wire \( w_i \) switches \( z_i \) times when \( C : s_0 \rightarrow_I X \), then the wire energy, \( E_w(C, s_0, X, \Delta, I) \), is given by

\[
E_w(C, s_0, X, \Delta, I) = \frac{1}{k} \sum_{i=1}^{\|W\|} z_i \circ (\|w_i\|), \quad \text{where } \|W\| \text{ is the cardinality of set } W \text{ and } \|w_i\| \text{ is the area of wire } w_i.
\]

If \( C_n \) is a VLSI circuit computing \( f_n : \{0,1\}^n \rightarrow \{0,1\}^m \Rightarrow C_n \) is in state \( s_0 \) at time \( t_0 \), and \( E_w(C_n, s_0, X, \Delta, I) \) is the wire energy consumed by \( f_n \) when \( X = (x_1, \ldots, x_n) \) is the input to \( C_n \) after time \( t_0 \) according to input schedule \( I \), then \( E_{\text{worst}}(C_n) \), the worst case multiswitch energy, is given by

\[
E_{\text{worst}}(C_n) \triangleq \max_{(s_0, X, \Delta, I)} E_w(C_n, s_0, X, \Delta, I)
\]
The definitions of $E_w^M$ and $E_{\text{worst}}^M$ allow for delays from essentially arbitrary sources. [Ki87] also considers special cases of $E_{\text{worst}}^M$, which are less pessimistic. In particular, multiswitch model $M_1$ is defined to have constant wire delay, i.e. propagation delay is independent of wire length. Further, $M_1$ has constant input delay, which means that all inputs arrive at the input ports at the same time. Hence in $M_1$, hazards can be caused only by circuit asynchrony. In particular, different path lengths to a given node can cause hazards, and variable gate switching speeds can induce hazards.

Consider a VLSI circuit $C$ which is a complete binary tree whose nodes are labelled with $\Theta$-gates. Since all gates of $C$ realize the same function and have minimal area, we can assume that the speed at which a node switches is the same for all nodes of $C$. Further, all paths to a given node in $C$ are equal in length. Therefore, under MSM $M_1$, $C$ exhibits no hazards and therefore has the uniswitch property (i.e. each node switches at most once). \[ E_{\text{worst}}^{M_1}(C) = \Theta(A), \] where $A$ is the area of $C$. The upper bound follows from the uniswitch property and the lower bound follows from Theorem 1.

[Ki87] also defined MSM $M_2$ in which $M_1$ is augmented with wire delays. In particular in $M_2$, the time to transmit a bit along a wire $w$ is a monotonic function of the length of $w$. Consider $C$ to be a standard embedding in which the input wires to a given node have the same length. (A minimal H-tree embedding has this property.) Then $C$ has the uniswitch property in $M_2$, and $E_{\text{worst}}^{M_2}(C) = \Theta(A)$.

Consider the most general MSM and consider a "bad" input schedule $I_b$ in which the primary input bits to the circuit switch sequentially. Let $M_3$ denote the MSM that results from augmenting $M_2$ with input schedule $I_b$. Since $C$ is a parity circuit, $I_b$ will cause the wires near the root of $C$ to switch each time an input switches, inducing $\Theta(n \log n)$ total switching in $C$. Thus, if $C$ is an H-tree embedding, $E_{\text{worst}}^{M_3}(C) = \Theta(n^{3/2})$, obtained by combining the area recurrence and the switching recurrence for $C$.

Consider an arbitrary parity circuit, but restrict the leaves to be on the boundary of the layout. The following derives an energy bound in the most general MSM, which results from the bad input schedule $I_b$. 

-5-
Theorem 2:
To compute parity by a VLSI circuit $P$ whose input nodes are on a convex boundary consumes
$E_{\text{worst}}^M(P) = \Theta(n^2)$.

Proof:
Consider the lower bound. Let $P$ be a CIF circuit (defined and motivated in [Ki87]). In particular, assume that $P$ has a constant number of copies of each input and that all copies of a given input are within a constant distance of each other. Let $z$ be the output node of $P$. Let $N$ be a subset of $n/2$ inputs that are farthest away from $z$. Thus, $z$ is $O(n)$ distance away from any input in $N$.

Apply input schedule $I_b$, i.e. switch each input one at a time. Each time an input $x_i \in N$ switches, $z$ switches. Hence a path from $x_i$ to $z$ must switch. Such a path is $O(n)$ long and $n/2$ such long paths switch, using $\Omega(n^2)$ multiswitch energy.

The upper bound of $O(n^2)$ is exhibited by a circuit in which each node is a $\Theta$-gate. Thus, for an embedding in which the input ports are on a convex boundary, $E_{\text{worst}}^M(C) = \Theta(n^2)$. []

5. OPEN PROBLEMS

The USM lower bounds on parity are derived in the special case where the circuit contains only $\Theta$-gates and negations.

Conjecture:
To compute the parity function on $n$ bits by an $O(\log n)$ depth circuit in which the inputs are on a convex boundary requires $\Omega(n \log n)$ uniswitch energy.

The conjecture above does not restrict the basis of the parity circuit. Note that in order to obtain an $\Omega(\text{area})$ uniswitch lower bound for parity in the general case, a notion of a "minimal" circuit is required. This is because an extraneous circuit that uses $o(\text{area})$ energy can always be "attached" to a parity circuit.

What about the majority function? We believe that majority also requires superlinear uniswitch energy if computed by a shallow depth circuit.
ACKNOWLEDGEMENTS

We thank Leslie Valiant for his contributions to the proof in the Appendix.

REFERENCES


[Me86] Mead, C., private communication


APPENDIX

The following presents an alternative proof of the worst case uniswitch energy lower bound for parity. In the main body of this note, the following result is obtained as a Corollary to the bound on the average energy.

**Theorem A.1:**

To compute parity on \( n \) boolean variables with a VLSI circuit \( C_n \) of area \( A \) requires \( E_{\text{worst}}(C_n) = \Omega(A) \) when the basis for \( C_n \) is \( \{\oplus\} \) and when \( C_n \) contains no constant inputs and no nodes that compute a constant function.
Proof:

Note that when the basis for $C_n$ is $\{\oplus\}$, each node of $C_n$ computes a parity function of a nonempty subset of the inputs. Let $S$ be the set of nodes of $C_n$. Let $X = (x_1, \ldots, x_n)$ be the input nodes (variables) of the circuit (function).

Definition:

Let $p \in S$ and let $f_p$ be the parity function computed at node $p$. Let $X_p \subseteq X \Rightarrow f_p$ is the parity function of inputs $X_p$. Let $w_{p1}, w_{p2}$ denote the output edges from $p$. (Recall that $C_n$ has fanout $\leq 2$). Let $\text{weight}(p) \triangleq \text{area}(w_{p1}) + \text{area}(w_{p2})$. When $p$ has fanout 1, $\text{area}(w_{p2}) = 0$. When $p$ has fanout 0, $\text{weight}(p) = 0$.

Lemma 1:

There exists an assignment $B$ of boolean values to $x_1, \ldots, x_n$ such that when $C_n$ is in state $B$, then

$$\sum_{p \in S \land (f_p(X_p) = 1)} \text{weight}(p) \geq \frac{1}{2} A$$

Proof of Lemma 1:

The following construction sequentially defines an assignment $B$ of values to the inputs $B(x_1), \ldots, B(x_n)$ — that will cause at least half the area of $C_n$ to be "1". In the following, $S_k$ is the subset of nodes of $C_n$ that depends only on the inputs $x_1, \ldots, x_k$. $A_k$ is the area of the out edges of nodes in $S_k$.

More formally,

let $S_k = \{ p \in S : x_k \in X_p \text{ and } \forall i > k, x_i \notin X_p \}$ and

let $A_k = \sum_{p \in S_k} \text{weight}(p)$

Basis of assignment: $B(x_1) = 1$

$B(x_1) = 1 \Rightarrow \sum_{p \in S_1 \land (f_p(x_1) = 1)} \text{weight}(p) = A_1$

In general, suppose $B(x_1), \ldots, B(x_{k-1})$ have been determined. To determine $B(x_k)$:

There are two choices for $B(x_k)$:
Suppose $B(x_k) = 0$.

Let $W_0 = \sum_{p \in S_1 \land (f_p(X_p) = 0)} \text{weight}(p)$

Let $W_1 = \sum_{p \in S_1 \land (f_p(X_p) = 1)} \text{weight}(p)$

**case 1:**

If $W_1 \geq \frac{1}{2} A_k$ then done. i.e. $B(x_k) = 0$

**case 2:**

If $W_1 < \frac{1}{2} A_k$ then set $B(x_k) = 1$. Since $\forall p \in S_1$, $f_p$ is a parity function and $x_k \in X_p$; then it follows that changing $B(x_k)$ from 0 to 1 changes $f_p(X_p)$ from 0 to 1.

Note that setting $x_k$ does not affect the functions realized by nodes in $S_i$ for $1 \leq i < k$.

[\[ \sum_{p \in S_1 \land (f_p(X_p) = 1)} \text{weight}(p) \geq \frac{1}{2} A_k \]

[\[ \sum_{p \in S \land (f_p(X_p) = 1)} \text{weight}(p) \geq A_1 + \frac{1}{2} \sum_{k=2}^{n} A_k \geq \frac{1}{2} A \]

[] (end of Lemma)

Since $C_n$ consists of $\oplus$ nodes only, $X = 0^n \Rightarrow$ all wires in $C_n$ have value 0. Let $B(X)$ be the value of $X$ determined by Lemma 1. If $C_n$ is switched such that $X : 0^n \rightarrow B(X)$, then $E_{\text{worst}}(C_n) \geq \frac{1}{2} A$. 

[]