Energy-Reliability Limits in Nanoscale Circuits

Avhishek Chatterjee and Lav R. Varshney
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

Abstract—In the nanoscale regime, the behavior of both extant and emerging semiconductor devices are often unreliable. Reliability of such devices often trades-off with their energy consumption, speed, and/or chip area. We study the reliability-energy limits for circuits designed using such devices; examples are drawn from spin electronics. Using the mutual information propagation in logic circuits technique developed by Pippenger, together with optimization, we obtain lower bounds on the energy consumption for computing $n$-input Boolean functions. We observe that the minimum energy requirement is strictly higher in an order sense (as $n$ scales) than that in the case of reliable gates.

Index Terms—Energy-reliability function, fault-tolerant computing, nanoelectronics

I. INTRODUCTION

As area and energy scaling in CMOS technology saturates, the semiconductor industry has been exploring promising new nanoscale devices [1]. A major challenge, however, in using nanoscale devices is that they can be very unreliable, especially when operated at low energy [2]. This has created renewed interest in the study of reliable circuit design using unreliable components [1], [3].

Von Neumann initiated studies of error in logic circuits from a statistical and thermodynamic point of view [4], showing that repeated computations followed by majority logic can offer reliability even when all components are unreliable. Over the last six decades, this result has inspired intermittent bursts of work on statistical models in computation.

Obtaining bounds on the minimum size of a noisy circuit to compute a function with a given reliability has drawn significant interest, e.g. [5]–[7]. There has also been work to obtain an upper bound on gate noise for which a non-trivial reliability can be achieved [8]–[10]. Another interesting thread of work is the study of unreliable circuits preceded and succeeded by reliable encoding and decoding components, respectively. A lower bound on the ratio of encoded input size to true input necessary for reliable computation was obtained in [11]. Use of encoder and decoder modules has also been employed in memory circuits [12]–[14], where LDPC decoders play an important role. Recently, linear function computation for unreliable encoder, decoder and computation circuits was studied in [15].

As von Neumann was evidently influenced by the success of information theory, efforts have been made to connect information theory and noisy computation [16]. Adding encoder and decoder modules to unreliable circuits [11] was an attempt to establish a correspondence between communication channels and computation circuits. A more natural use of information-theoretic techniques to study noisy circuits was developed much later by Pippenger [9], improvements of which led to the lower bound on the size of a circuit with a desired reliability mentioned above [8], [17], [18].

In this work, we study the limits of energy consumption for reliable computation with unreliable components using Pippenger’s information-theoretic technique. It has been observed that energy consumption increases as spin and other nanoscale devices are designed to be more reliable [2]. Thus, there is a fundamental tradeoff between reliability and energy consumption. Our goal is to obtain the information-theoretic limit of energy consumption for achieving a desired reliability in computation. The main result is the minimum computation energy per input bit.

II. SETTING

The goal is to design circuits to compute $n$-input boolean functions using a single type of gate among the set of universal gates, i.e. NAND and NOR, such that each gate has at most $k$ inputs and exactly one output. Gates are interconnected into a circuit to compute the desired function, such that inputs to gates are some of the $n$ inputs to the function, outputs of other gates in the circuit, or constants $\{0,1\}$. The output of the circuit is the output of a certain gate. In general, a boolean function can be realized by several different circuits using the same kind of universal gate. Indeed, elementary digital logic design is concerned with minimal circuits using the same kind of universal gate. We assume that the boolean function $F$ is sensitive to each input, i.e., for each input $i$, there is a configuration of other inputs $x_1 = c_1, x_2 = c_2, \ldots, x_{i-1} = c_{i-1}, x_{i+1} = c_{i+1}, \ldots, x_n = c_n$, such that

$$F(c_1, \ldots, c_{i-1}, 0, c_{i+1}, c_n) \neq F(c_1, \ldots, c_{i-1}, 1, c_{i+1}, c_n).$$

A boolean function of $n$ inputs which is not sensitive to one of its inputs is equivalent to a function of $n-1$ inputs. In combinatorial circuits there is no feedback, i.e., the connections between the gates must form a directed acyclic graph, where gates are vertices and the connections between gates are the edges. The input to a gate is considered the head of a directed edge. A special class of combinatorial circuits of interest in this work are formulas, where the graph is a directed tree.

A gate may fail to produce a correct output for a given input with probability $\epsilon$. When a gate fails, its output is flipped.

This work was supported in part by Systems on Nanoscale Information fabrics (SONIC), one of the six SRC STARnet Centers, sponsored by MARCO and DARPA.
This means that for a given input, the output is correct with probability \(1 - \epsilon\) and is flipped with probability \(\epsilon\). Each gate in a circuit or formula fails independently of each other. This model of gate failure is known as the \(\epsilon\)-noisy model [4].

The noise probability \(\epsilon\) of a gate depends on its electrical characteristics such as bias voltage, as well as physical and material properties. These characteristics also determine the energy consumption of the gate. For a given device technology, there is a relationship between energy consumption of the gate and its probability of failure. This function depends on the fundamental nature of the device, whether CMOS or emerging beyond-CMOS technologies such as spin electronics or carbon nanotubes.

**Definition 1:** For a gate with probability of failure \(\epsilon\) and energy consumption \(e_g\), let us define the energy-reliability function of the gate as \(\epsilon = \chi(e_g)\).

When we construct circuits from noisy gates, for any given input, there is a probability that the output of a circuit is incorrect.

**Definition 2:** We say a circuit for the \(n\)-input boolean function \(F\) is \(\delta\)-reliable if for any input configuration \(\{x_1, x_2, \ldots, x_n\} \in \{0, 1\}^n\), the output of the circuit \(y\) satisfies the following:

\[
\Pr(y = F(x_1, x_2, \ldots, x_n)) \geq 1 - \delta,
\]

where probability is over all the failure patterns of the gates in the circuit.

This work aims to answer the following general question: What is the minimum energy expenditure needed to realize an \(n\)-input boolean function using a \(\delta\)-reliable formula? Herein, we restrict attention to circuits (and/or technologies) where each gate in the circuit must have the same electrical characteristics and hence consume the same energy. We want to obtain the minimum energy per device for reliable computation. Some emerging technologies like spin electronics allow different devices in the circuit to have different electrical characteristics [20]. The energy-reliability limit with such variable-quality gates would be tighter; we describe characterization of such settings elsewhere.

### III. Preliminaries

#### A. Circuit Graphs

An input boolean function \(F_n\) can be realized using various different formula structures of (universal) logic gates of a given kind. For a realization we have a given directed graph \(G_g = (V_g, E_g)\), where vertices \(V_g\) are the set of gates and \(E_g\) are the edges corresponding to connections between gates. Each edge is directed towards the gate to which it is an input. We call \(G_g\) the *gate graph* of the realization of the formula. In this work we only consider formulas, and hence, the corresponding graphs are trees.

Another graph that we use subsequently is the *bit graph* \(G_b = (V_b, E_b)\). Here, \(V_b\) corresponds to interconnects/wires in the circuit. In a circuit there are interconnects that run from output of a gate to input of another gate, from inputs to the circuit to inputs of gates, and from fixed sources (corresponding to permanent 0 or 1) to input of gates. Two nodes in \(V_b\) share an edge if the corresponding interconnects are incident to a common gate.

As the circuit is a tree, interconnects between gates have a one-to-one mapping to gates (the gates to which they are outputs). From here it can be seen that \(G_g\) is a subgraph of \(G_b\). In addition \(G_b\) has leaf nodes that correspond to either inputs or to fixed sources. This relation will be useful later.

#### B. Pippenger’s Method

Pippenger developed the technique of information flow in logic circuits to bound the size of a \(\delta\)-reliable circuit constructed using \(\epsilon\)-noisy gates [9], which we review here. For a boolean function \(F\) of \(n\)-inputs, for any input \(i\) there exists a values \(c_i \in \{0, 1\}^{n-1}\) such that \(F_i(x) := F(c_i, x) = x\) or \(\bar{x}\). Hence, for any random input \(X, F_i(X)\) is a one-to-one mapping. When \(F\) is realized using noisy gates, the output is no longer \(F_i(X)\), but a random variable \(Y\), depending on the gate noise and \(G_g\).

Note \(\delta\) is the upper bound on probability of error, \(P_e(c_i)\) for input configuration \(c_i\), and the possible number of values of \(Y\) is \(M = 2\). Then using Fano’s inequality [21]:

\[
I(X; Y) \geq H(X) - h(P_e(c_i)) - P_e(c_i) \frac{\ln(M - 1)}{\ln 2},
\]

and the fact that \(M = 2\), as the output is binary, it follows that:

\[
1 - h(\delta) \leq I(X; Y),
\]

where \(h(\cdot)\) is the binary entropy function [21]. Note that unlike in communication where \(h(P_c)\) is bounded by 1 to obtain a bound on \(P_e\), here the bound is obtained in terms of \(h(P_c)\).

Finally to involve gate failures into this bound, two important observations are made. First, for a perfect gate \(g\) with input random variables \(U_1, U_2, \ldots, U_k\) and output random variable \(U_0\), for any random variable \(Z\):

\[
I(Z; U_0) \leq I(Z; U_1, U_2, \ldots, U_k) \\
\leq \sum_{i=1}^{k} I(Z; U_i).
\]

by the data processing inequality and the distributive rule of mutual information [9], [21]. Second, if the gate is \(\epsilon\)-noisy, then the output is \(\bar{U}_0 = U_0 + N_0\), where \(N_0\) is a Bernoulli(\(\epsilon\)) random variable independent of \(U_0\). It has been shown that [9], [17], [18]:

\[
I(Z; \bar{U}_0) \leq (1 - 2\epsilon)^2 I(Z; U_0).
\]

Based on these two information-theoretic inequalities and some combinatorial arguments, a lower bound on the depth of a formula was obtained in terms of \(\delta\) and \(\epsilon\) in [9], [17], [18]. There, the gate noise was fixed and the energy consumption was not considered. In this work, we build on the technique proposed by Pippenger to study circuits where energy (and hence, noise) of a gate can be tuned and derive bounds on total energy consumption to realize a boolean function.
IV. Computation Energy Per Input Bit

In this section we consider the case where each gate in the circuit must have the same electrical characteristic. This requirement may have to be imposed due to limitations of the design and fabrication technologies of the electronic devices being used. Some emerging technologies do allow variable gate characteristics [20], but here we do not discuss energy-reliability limits in that setting.

For an n-input formula F, consider a given realization of the formula using logic gates and hence a given directed gate tree \( G_g = (V_g, E_g) \). For each input bit \( i \),

\[
1 - h(\delta) \leq I(F_i(X); X) \leq (1 - 2\epsilon)^{2|P_i|},
\]

where \( P_i \) is the path in \( G_g \) from the gate to which \( x_i \) is input to the output gate. This follows by inductively using (3) and (4) along the depth of \( G_g \) from root to the terminal gate into which \( x_i \) is an input [9], [17], [18].

Thus, for any \( \delta \)-reliable circuit realization of \( F \) with total energy consumption \( E \) the following condition must be satisfied.

\[
\begin{align*}
C1. & \quad 1 - h(\delta) \leq (1 - 2\epsilon)^{2|P_i|}, \text{ for } 1 \leq i \leq n, \quad (5) \\
C2. & \quad \chi \left( \frac{E}{|V_g|} \right) = \epsilon.
\end{align*}
\]

Condition C1 follows from the definition of \( \delta \)-reliability of the circuit since for any input bit \( i \) and any configuration \( c_i \), \( Pr(Y = F_i(X)) \geq 1 - \delta \). Condition C2 relating total energy \( E \) and \( \epsilon \) follows because gates with the same electrical characteristics consume the same energy. Hence, energy consumption per gate is \( e_g = \frac{E}{|V_g|} \), which along with \( \epsilon = \chi(e_g) \) implies the condition.

Our goal is to find a lower bound on the total energy consumption \( E \) in a circuit realization of \( F \). This implies that a set of conditions weaker than C1 and C2 would give a lower bound on \( E \). Hence, we develop the following weaker conditions.

\[
\begin{align*}
C3. & \quad \frac{1}{4} \ln \frac{1}{1 - h(\delta)} \geq |P_i| \epsilon, \text{ for } 1 \leq i \leq n, \quad (6) \\
C4. & \quad \chi \left( \frac{E}{|V_g|} \right) = \epsilon.
\end{align*}
\]

Note that conditions C2 and C4 are identical, but we give them different names for convenience. Due to linear dependence on graph structure, the conditions C3 and C4 are more tractable than C1 and C2. As we discuss later, these condition lead to a closed-form expression for the lower bound on energy consumption.

The following lemma formally relates the conditions C3 and C4 to the conditions C1 and C2.

**Lemma 1:** For a given formula \( F \), a realization gate tree \( G_g = (V_g, E_g) \), and required reliability \( \delta \), if \( E \) satisfies conditions C1 and C2, then it also satisfies conditions C3 and C4.

**Proof:** We need only consider the relation between C1 and C3. The proof follows by noting that

\[
1 - x \leq \exp(-x),
\]

which implies that if (5) is satisfied for an \( E \), then for that \( E \):

\[
1 - h(\delta) \leq \exp(-4|P_i| \epsilon), \text{ for } 1 \leq i \leq n.
\]

By taking natural logarithm of both sides and noting that as \( \ln(\cdot) \) is a monotonic function, the inequality is not disturbed. Thus we obtain the desired result by multiplying both sides by \(-1\) and flipping the inequality.

It is apparent from the necessary conditions we have just established that the lower bound on energy consumption will depend on the dependence between probability of failure of a gate and its energy consumption. Intuitively, when probability of failure of a gate decays faster with its energy consumption, the total energy consumption of the circuit should also be lower. In this work, we consider a few special cases of the function \( \chi \) that are relevant to practical technologies. Before that, though, we obtain a generic lower bound applicable to a broad class of energy-failure functions \( \chi \), as defined here.

**Definition 3:** Let the class of physical energy-failure functions be ones that satisfy \( \chi : (0, \infty) \to (0, a), 0 < a \leq 1 \), that is strictly decreasing and differentiable with \( \lim_{e_g \to 0} \chi(e_g) = a \) and \( \lim_{e_g \to \infty} \chi(e_g) = 0 \).

Before stating and proving this main result, we give several lemmas that are useful.

**Lemma 2:** For physical energy-failure functions \( \chi \) and for a given n-input boolean function \( F \) with realization graph \( G_g = (V_g, E_g) \),

\[
E \geq |V_g| \chi^{-1} \left( \frac{1}{4 \max |P_i|} \ln \frac{1}{1 - h(\delta)} \right),
\]

for \( \delta < \frac{1}{2} \).

**Proof:** For a given n-input formula \( F \) and its corresponding directed tree \( G = (V, E) \), it follows from (6) that

\[
\epsilon \leq \min \frac{1}{4|P_i|} \ln \frac{1}{1 - h(\delta)} = \frac{1}{4 \max |P_i|} \ln \frac{1}{1 - h(\delta)}.
\]

Next, we establish that \( \chi \) has an inverse that is one-to-one and strictly decreasing. The proof is standard in real analysis.

Since \( \chi \) is strictly decreasing, it is one-to-one. Otherwise there would be an \( e_1 < e_2 \), such that \( \chi(e_1) = \chi(e_2) \). From differentiability and therefore continuity of \( \chi \), it also follows that \( \chi \) is onto.

Let \( 0 < \epsilon < a \) be such that there is no \( e_g > 0 \) with \( \chi(e_g) = \epsilon \) and we show a contradiction. As the function is decreasing, there exists an \( \epsilon \) such that \( \chi(x) < \epsilon \) for all \( x > e_g \) and \( \chi(x) > \epsilon \) for all \( x < e_g \). Consider \( \lim_{x \to e_g} \chi(x) \) and \( \lim_{x \to e_g} \chi(x) \). For any \( x_k \uparrow e_g \), \( \chi(x_k) \) is monotonic and bounded, so \( \lim_{x \to e_g} \chi(x_k) \) exists and is finite. Similarly, this is true for \( x_k \downarrow e_g \) and \( \lim_{x \to e_g} \chi(x_k) \). As \( \chi \) is continuous these two limits match and must equal \( \chi(e_g) \). But as \( \chi(x) < \epsilon < \chi(x') \) for \( x > e_g > x' \), the only possible limit is \( \epsilon \). So, \( \lim_{e_g \to 0} \chi(e_g) = \epsilon \). This implies that \( \chi : (0, \infty) \to (0, a) \) is onto.

So, the inverse exists for one-to-one and onto \( \chi \).

To see that \( \chi^{-1} \) is strictly decreasing, let \( 0 < e_1 < e_2 < a \). Now, there exists \( e_1 \) and \( e_2 \) such that \( e_i = \chi^{-1}(\epsilon_i), i \in \{1, 2\} \).
We show that \( e_1 \leq e_2 \) is not possible. Let \( e_1 \leq e_2 \), then by the strictly decreasing property of \( \chi \), \( \chi(e_1) \geq \chi(e_2) \), implying \( e_1 \geq e_2 \), which is contradictory.

Now as \( \epsilon = \chi(E/|V_g|) \), and \( \chi^{-1} \) is strictly decreasing, \( \epsilon \leq \frac{1}{4 \max_i |P_i|} \ln \frac{1}{1 - h(\delta)} \) implies that for a given formula and \( G_g \),

\[
E \geq |V_g| \chi^{-1} \left( \frac{1}{4 \max_i |P_i|} \ln \frac{1}{1 - h(\delta)} \right).
\]

We also observe some combinatorial properties on trees with a fixed number of leaves.

Lemma 3: Among all directed rooted trees with \( L \) leaves and number of children bounded by \( k \), a \( k \)-ary balanced tree has the minimum number of non-leaf nodes.

Proof: A directed rooted tree can be seen as a structure where directed lines originate from leaves and are eventually merged at the root node. Each non-leaf node merges lines coming from the level below it. Thus there are \( L \) lines to be merged.

As each node can have at most \( k \) children, each non-leaf node can merge at most \( k \) lines. Each non-leaf node can be thought of as a gadget/entity which can take at most \( k \) lines as input and outputs one merged line. Consider any structure of the non-leaves that merges \( L \) lines to 1. In any of these structures if there is a non-leaf node with less than \( k \) inputs, then replacing that with a non-leaf node with \( k \) inputs does not increase the total number of non-leaf nodes. So for any tree there is a \( k \)-ary tree with the same number of leaf nodes, and no more non-leaf nodes. Note that converting a \( k \)-ary tree to a balanced \( k \)-ary tree with the same number of leaf nodes does not increase the total number of nodes and hence, does not increase the number of non-leaf nodes.

Lemma 4: Among all directed rooted trees with \( L \) leaves and number of children bounded by \( k \), a \( k \)-ary balanced tree has the minimum depth for the sub-tree made of non-leaf nodes.

Proof: By similar arguments as in Lemma 3. Note that if there is a tree with certain number of leaves and each node with \( \leq k \) children, then depth does not increase if we change it to a \( k \)-ary tree to same number of leaves. The rest follows by noting that converting a \( k \)-ary tree to a balanced \( k \)-ary tree with the same number of leaves does not increase the depth.

Lemma 5: Among the class of directed rooted trees with bounded children, the minimum depth and minimum size of the sub-tree consisting of non-leaf nodes are monotone in number of leaves.

Proof: Consider two directed trees with \( L \) and \( L' \geq L \). Then, any graph configuration with \( L' \leq L \) leaves can be transformed to graph with \( L \) leaves without increasing the size (just by dropping \( L' - L \) leaves). Hence, the minimum achievement configuration for \( L' \leq L \) leaves is also a configuration for \( L \) leaves. Hence, the minimum in case of \( L \) leaves is not larger than the minimum in case of \( L' \) leaves.

With these lemmas in hand, we can give the main theorem on minimum energy per input bit.

Theorem 1: For a physical energy-failure function, the minimum energy required to realize any \( n \)-input boolean function using a \( \delta \)-reliable (\( \delta < \frac{1}{2} \)) formula of gates with degree \( k \) (with \( k < n \)), and each with energy-failure function \( \chi \) is

\[
E \geq \frac{n}{k} \chi^{-1} \left( \frac{1}{4 \ln n} \ln \frac{1}{1 - h(\delta)} \right),
\]

where \( \chi^{-1} \) is the inverse of \( \chi \).

Proof: A lower bound on energy consumption over all realization circuits (graphs) and all \( n \)-input boolean functions can be obtained by minimizing the above realization-specific bound over all realizations and \( n \)-input functions:

\[
E \geq \min_{F_n, G_g} |V_g| \chi^{-1} \left( \frac{1}{4 \ln n} \ln \frac{1}{1 - h(\delta)} \right)
\]

where last equality follows because for each formula there is a gate tree and vice-versa.

For a rooted tree \( G = (V, E) \), let \( \bar{\ell}(V) \) and \( \tilde{\ell}(V) \) denote the leaf and non-leaf nodes of a graph \( G \), respectively. Then, it is clear from the discussion on bit graphs and gate graphs that \( \ell(V) = V_g \) and \( \bar{\ell}(V) = E_g \cap (V_g \times V_g) \).

As \( G_g \) has \( n \) inputs, the corresponding \( G_k \) has at least \( n \) leaves. Thus, one can write

\[
\min_{G_k; n \text{ inputs}} |V_g| = \min_{G_k; \geq n \text{ leaves}} \tilde{\ell}(V) = \ell(V).
\]

By Lemma 5 and the fact that adding more constraints increases the minimum we have

\[
\min_{G_k; \geq n \text{ leaves}} \tilde{\ell}(V) = \ell(V).
\]

By Lemma 3, a \( k \)-ary \( G_k \) achieves the minimum. Now as there are \( n \) leaves, in a \( k \)-ary tree there are at most \( \lceil \frac{n}{k} \rceil \) nodes in the level above. In turn, there are at most \( \lceil \frac{n}{\delta k} \rceil \) in the level above that and so on. This continues until we have only one node at the top level. Thus total number of non-leaf nodes are lower bounded by

\[
\left( \frac{1}{k} \right) + \left( \frac{1}{k} \frac{1}{k} \right) + \ldots + 1 = \frac{\left\lfloor \frac{n}{k} \right\rfloor}{k}.
\]

This implies that \( \min_{F_n, G_g; n \text{ inputs}} |V_g| \geq \frac{n}{k} \).

To bound the other term note that \( \chi^{-1} \) is strictly decreasing, so \( \chi^{-1} \left( \frac{1}{4 \max_i |P_i|} \ln \frac{1}{1 - h(\delta)} \right) \) is minimized when \( \frac{1}{4 \max_i |P_i|} \ln \frac{1}{1 - h(\delta)} \) is maximized. This is because \( \ln \frac{1}{1 - h(\delta)} \) is maximized for \( \delta < \frac{1}{2} \). Thus when \( \max_i |P_i| \) is minimized the other term is also minimized.

As \( \max_i |P_i| \) is the depth of \( G_g \), by Lemma 4, this minimum is achieved by a \( k \)-ary tree. Now, by Lemma 5 and the relation
between $G_b$ and $G_g$, the depth of $G_g$ minimized when number of leaves in $G_b$ are minimized. The number of leaves in $G_b$ is $\geq n$, as there are $n$ inputs. For a $k$-ary tree with $n$ leaves, the depth is at least $\left\lceil \frac{\ln n}{\ln k} \right\rceil \geq \frac{\ln n}{\ln k}$.

Hence,

$$E \geq \frac{n}{k} \chi^{-1} \left( \ln \frac{1}{1 - h(\delta)} \frac{\ln k}{4 \ln n} \right).$$

To understand the implication of Theorem 1 let us consider the simple case where we fix $\delta \in (0, \frac{1}{2})$ and $k = O(1)$. Then, by the theorem to realize $\delta$-reliability using gates of at most $k$ inputs the minimum energy requirement scales with number of inputs $n$ as

$$\Omega \left( n \chi^{-1} \left( c(\delta, k) \frac{\ln n}{\ln k} \right) \right),$$

for some constant $c(\delta, k)$.

Typical nanoscale circuits compute functions of large numbers of inputs. As $n$ is large in these circuits, the above order behavior is of fundamental interest. In most practical scenarios (as also under the condition in Theorem 1), $\chi^{-1}$ is strictly decreasing and $\lim_{\epsilon \to 0} \chi^{-1} = \infty$. Thus, as $n \to \infty$, the term

$$\chi^{-1} \left( c(\delta, k) \frac{\ln n}{\ln k} \right) \to \infty.$$

This implies that the minimum energy requirement per input bit over all boolean functions increases with number of input bits for any $\delta \in (0, \frac{1}{2})$ and $k = O(1)$.

This is in sharp contrast with traditional circuits with perfect gates. In that case there are various $n$-input functions that can be realized using $O(n)$ gates. As each gate requires only a constant amount of energy for its perfect operation, the total energy consumption is $O(n)$ for these circuits. Hence, it seems that as devices and gates become unreliable there is a price to be paid in terms of the energy per input bit.

Next, we build a more quantitative understanding of the scaling of energy per input bit by considering a few relevant classes of energy-failure dependence functions $\chi$.

V. SPIN ELECTRONICS AND OTHER TECHNOLOGIES

For a typical spin device a closed-form expression relating energy and failure has been derived in [2] based on the physics of the device. An approximate form of the functional dependence can be given by

$$\epsilon_0 \exp(-cI),$$

where $I$ is the supply current of the device. Constants $c > 0$ and $\epsilon_0 \in (0, 1)$ depend on the device parameters, like critical current, gate delay, and other switching parameters. As energy consumption scales as square of current, in a spin device reliability $\epsilon$ and energy $\epsilon$ are related as

$$\epsilon_0 \exp(-c'\sqrt{\epsilon}),$$

with $c' > 0$.

A generic way to capture this kind of dependence is through stretched exponentials,

$$\epsilon = \epsilon_0 \exp(-c\epsilon^\beta),$$

where $\epsilon_0, \beta \in (0, 1)$, and $c > 0$.

Polynomial functions form another wide class of energy-failure dependence functions that can be used to approximate different nanoscale devices.

$$\epsilon = \frac{\epsilon_0}{(e+1)^\beta},$$

with $\beta > 0$, $\epsilon_0 \in (0, 1)$.

The following lower bounds for polynomial and stretched exponential energy-failure dependence follow from Theorem 1.

**Corollary 1:** The minimum energy required to realize any $n$-input boolean function using gates with $\leq k$ inputs, $k < n$, and each with $\epsilon = \chi(\epsilon)$ is

$$\frac{n}{k} \left( \ln \frac{4\epsilon_0 \ln n}{\ln k} - \ln \ln \frac{1}{1 - h(\delta)} \right)^\frac{1}{\beta},$$

when $\chi(\epsilon) = \epsilon_0 \exp(-c\epsilon^\beta)$, and

$$\frac{n}{k} \left( \ln \frac{4\epsilon_0 \ln n}{\ln k} \right)^\frac{1}{\beta},$$

when $\chi(\epsilon) = \epsilon_0 \exp(-c\epsilon^\beta)$.

**Proof:** Note that both of the energy-failure functions are physical, and so the result follows by substituting the appropriate $\chi^{-1}$ in Theorem 1.

In cases where we have a lower bound on $\chi$ rather than an exact functional form, we can still obtain a lower bound on energy consumption. Towards this the following lemma is in place.

**Lemma 6:** Let energy-failure functions $\chi_1$ and $\chi_2$ be physical and $\chi_1(\epsilon) \leq \chi_2(\epsilon)$ for all $\epsilon$. For gates with $\chi_1$ if there exists no $\delta$-reliable circuit $G_g$ for a formula $F$ with total energy consumption $\leq E$, then the same is true for $\chi_2$ gates.

**Proof:** The result follows by noting that if a per-gate energy $\epsilon_2$ in case of $\chi_2$ achieves $\delta$-reliability, then by the monotonicity of $\chi_1$ and $\chi_2$, the dominance between them and the condition (6), $\epsilon_2$ energy per gate achieves $\delta$-reliability in case of $\chi_1$ gates.

Thus, if we have a lower bound on $\chi$, we can use the above lemma to find the fundamental lower bound on energy consumption. This is useful in scenarios where the physics of the device is not tractable and an exact functional form is hard to derive. Here we use this to make an interesting generic observation about a broad class of energy-failure functions.

Note that the exponential function $\epsilon = \epsilon_0 \exp(-c\epsilon)$ lower bounds polynomial and stretched exponential functions. Thus, by Lemma 6, in fact, the exponential function can be used to obtain a lower bound on energy consumption for all sub-exponential energy-failure functions. Since energy-failure functions fall in the sub-exponential class for most practical devices, we can obtain a generic bound using it.

From Corollary 1 it follows that for $\epsilon = \epsilon_0 \exp(-c\epsilon)$, the energy consumption for computing $n$-input functions is lower bounded by

$$\frac{n}{ck} \left( \ln \ln n - \ln \ln k + \ln(4\epsilon_0) - \ln \ln \frac{1}{1 - h(\delta)} \right).$$

(7)
As $k = O(1)$ and so is $\delta$, for large $n$ the leading term is $\frac{n \ln \ln n}{ck}$. So, in the large $n$ regime, as long as $\delta < \frac{1}{2}$, irrespective of $\delta$ the minimum energy requirement scales at least as $\frac{\ln \ln n}{ck}$. This has interesting implications.

First, if gates are erroneous and have sub-exponential energy-failure functions, then to achieve any non-trivial reliability ($\delta < \frac{1}{2}$), the energy requirement per bit of computation scales at least as $c' \ln \ln n$, and this lower bound on scaling is independent of the reliability requirement (as long as it is non-trivial). This means that the reliability requirement is not the bottleneck in obtaining linear scaling of energy with number of inputs. Rather, the bottleneck is the sub-exponential unreliability of the gates.

Second, even if we allow a decreasing reliability requirement with the increasing input size, it does not help in obtaining a constant energy consumption per input bit. As long as $\delta_n \uparrow \frac{1}{2}$ such that $h(\delta_n) = 1 - \omega(\frac{1}{n})$, the minimum energy requirement per input bit scales at least as $\ln \ln n$, irrespective of the scaling of $\delta_n$. This can be seen by substituting $h(\delta_n)$ for $h(\delta)$ in (7). This further implies that the fundamental bottleneck is the sub-exponential unreliability of the gates, not the reliability requirement.

VI. Conclusion and Future Work

We obtained a lower bound on the minimum energy consumed in computing an $n$-input boolean function using unreliable gates. We observed that a superlinear scaling of energy consumption with the number of input bits is unavoidable, irrespective of the reliability requirement. In this work we only considered circuits with uniform energy consumption across gates. But, with some emerging nanoscale technologies, it is possible to design circuits with non-uniform energy consumption across gates. As ongoing work, we are developing lower bounds on energy consumption and optimal energy allocations for these circuits.

ACKNOWLEDGMENT

Discussions with Ameya Patil and Naresh Shanbhag are appreciated.

REFERENCES


