Energy Complexity of Optical Computations

Akhilesh Tyagi[†] and John H. Reif[‡]

Abstract

This paper provides lower bounds on the energy consumption and demonstrates an energy-time trade-off in optical computations. All the lower bounds are shown to have the matching upper bounds for a transitive function – shifting. Since the energy consumption in an optical transmission is a non-linear function of the distance, a new set of techniques was required to derive these lower bounds. We also characterize the energy requirements of 3-D VLSI computations.

key words: lower bounds, trade-offs, energy, optical, electro-optical, VLSI, computation.

[†]Department of Electrical and Computer Engineering, and Department of Computer Science, Iowa State University . Email: tyagi@iastate.edu

[‡]Department of Computer Science, Duke University, Durham, NC 27708, USA . Email: reif@cs.duke.edu

1 Introduction

Motivation: The field of electro-optical computing (see for example [LPZ98, RT90a, Kia91, WS98a, WS98b, WS98c, Sah01, A2010]) has considerable potential, comparable to the state of Very Large Scale Integration (VLSI) technology say 30 years ago, and a recent survery of optical computing by Ambs [A2010] indicates the field remains very active. The theoretical underpinnings for optical technology – namely the discovery of efficient algorithms and of resource lower bounds, are crucial to guide its development. , and *Volume V*,*energy E* and *time T* are also the resources of interest in a 3-D VLSI circuit or an optical computing system.

Prior results for 2D VLSI : *Area A*, *energy E* and *time T* are three fundamental resources in a 2D VLSI computation. An *Area-Time tradeoff* is an upperor lower bound on a function of Area A and Time T. Two dimensional VLSI Area-Time tradeoff techniques bounding the function AT^2 were pioneeredby Thompson [Tho79, Tho80] and Brent and Kung [BK1981].

Prior results for Energy Complexity of VLSI: Energy has received increased attention recently because the power consumption largely determines the total cost of a high performance computer due to heat dissipation. The theoretical physicists [BL85] have also considered the viability of characterizing the computational costs entirely in terms of energy. Much of the early research activity in energy complexity was directed at the study of the energy requirements in 2-D VLSI computations by Lengauer and Mehlhorn [LM81], Kissin [Kis82, Kis85, Kis87, Kis91], Aggarwal et al [ACR88], Snyder [ST86] and Tyagi [Tya88, Tya89], using techniques similar to the two dimensional VLSI Area-Time tradeoff techniques of [Tho79, Tho80, BK1981]. More specifically, the first formal result in switching energy was due to Lengauer, Mehlhorn [LM81], which shows that the switching energy of transitive functions – E, is $\Omega\left(\frac{n^2}{P\log(AP^2/n^2)}\right)$, which is $\Omega(n^2)$ for $AP^2 = O(n^2)$, where n is the input size, A is the area, and P is the period of a pipelined computation. Kissin [Kis82], [Kis85] proposed a formal model for switching energy distinguishing between uniswitch and multiswitch models. When a wire is assumed to switch at most once during the course of computation, it is a *uniswitch* circuit. Most of the pipelined computations fall in this class. The more general model, that allows each wire to switch any number of times, is called the *multiswitch* model. Snyder, Tyagi [ST86]and Leo [Leo84] considered variations on Lengauer, Mehlhorn result. The first tight bound on uniswitch and multiswitch energy-period product [$\Omega(n^2)$] for shifting was obtained by Aggarwal et. al. [ACR88]. Tyagi [Tya89] derived a

tight bound on multiswitch energy, $\Omega(n^{3/2})$, and average case uniswitch and multiswitch energy. More recent work on energy-time trade-offs includes the work included the following: Martin [MNP2002, M2001] introduced novel metrics for time and energy efficiency of computations. Bingham and Greenstreet [BG2008, BG2012] introduce a novel model for analyzing energy-time trade-offs of CMOS circuits, and proving lower and upper bounds.

Prior results for 3D VLSI: *Volume V* (the volume of the 3D region of the device), *energy E* and *time T* are the resources of interest in a 3-D VLSI circuit. A 3-D VLSI model and its volume-time trade-offs were first studied by Rosenberg [Ros81], and refined by Preparata [Pre83], and Leighton, Rosenberg [LR86]. Barakat and Reif gave some *volume V, time T* upper bounds [BR87b] and lower bounds [BR87a] for 3D optical computations.

Prior results for Parallel and Distributed Architectures: The above prior work (and our paper) assumes that energy is dissipated mostly because of long wires or beams; however some others have assumed that energy is mostly dissipated in the gates or CPU cores, rather than in their communication channels. Korthikanti and Agha analyzed the energy consumption of parallel algorithms on multicore architectures [KA2009], shared memory architectures [KA2010], and their parallel algorithms [KAG2011]. Khude et al [KKK2008] gave time and energy bounds for for distributed computation in wireless sensor networks. Ebergen et al [EGC2004] discuss how to control the speed and energy consumption of a Boolean circuit.

A similar theoretical analysis of energy bounds in VLSIO (O stands for optics, described in Section 2) computations has been missing.

Our results: In this paper, we analyze the energy requirements in 3-D VLSI and VLSIO systems. *Volume V*,*energy E* and *time T* are also the resources of interest both to optical computing system (as well as to a 3-D VLSI circuit). We propose two models for the energy consumption in an optical computer which are consistent with the VLSIO model described in [BR87a]. Our paper assumes that energy is dissipated mostly because of long wires and beams. An energy-time product ET bound can be used to succinctly express a natural trade-off between two key parameters of a computation: the energy E for the computation and the computation time T; where an decrease in the time requires a corresponding increase in energy consumption. Within these models, we demonstrate tight bounds on both energy E and energy-time product ET for the optical computation of several functions. We derive the lower bounds, shown in Table 1. The matching upper bounds are established for a transitive function: *barrel-shifting*, which is to circularly shift the elements of a vector.

	Known 3-D VLSI (linear energy cost)	Known 2-D VLSI (nonlinear energy function $f(x)$)	Our VLSIO (nonlinear energy function $f(x)$)
Uniswitch ET product	$\Omega(n^{3/2})$	$\Omega(n\,f(n)) \quad x < f(x) < x^2$ $\Omega(n^3) \quad f(x) \geq x^2$	$\Omega(nf(\sqrt{n})) \quad x < f(x) < x^2$ $\Omega(n^2) \quad f(x) \geq x^2$
Multiswitch ET product	$\Omega(n^{3/2})$	$\Omega(n^2)$	$\Omega(nf(\sqrt{n})) \quad x < f(x) < x^{4/3}$ $\Omega(n^{5/3}) \quad f(x) \geq x^{4/3}$
Uniswitch Energy	$\Omega(n^{3/2})$	$\Omega(n^2)$	$\Omega(n^{3/2})$
Multiswitch Energy	$\Omega(n^{4/3})$	$\Omega(n^{3/2})$	$\Omega(n^{4/3})$

Table 1: Energy-Time Lower Bounds: Known and Our Results

This paper is organized as follows. Section 2 describes the model. Section 3 deals with the ET lower bounds for the 3-dimensional VLSI, 2-dimensional VLSI with nonlinear energy cost and VLSIO. The upper bounds are shown in Section 4. For the average case energy consumption, the same lower bounds hold if we replace n by the function entropy H_f. We outline this extension in Section 5.

2 Models

Thompson's $2-D$ VLSI model [Tho79]: In Thompson's $2-D$ *VLSI* model a computation is abstracted as a communication graph, which is very much like a flow graph with the primitives being some basic operators that are realizable as electrical devices. A chip layout can be viewed as a convex embedding of the communication graph in a multilayered Cartesian grid. The number of layers is limited to some constant γ, so both the fanin and fanout are bounded by $4\,\gamma$. Each layer of the chip is modeled as a 2D rectangular area, where: Horizontal and vertical tracks are evenly spaced between consecutive tracks, and where certain tracks are occupied by wires. Furthermore, vias connections may be used to connect horizontal and vertical tracks, and certain track intersections have positioned on them active devices. Wires have unit width and bandwidth and processors have unit area. The operation of the chip in Thompson's $2-D$ VLSI model is by synchronous time steps, where in parallel on each synchronous time step: (a) on each wire, at each time step, a signal can propagate across the wire independent of length, and (b) each device can make at most a

single operation per time step. The initial data values are localized to some constant area, to preclude an encoding of the results. The input words are read at the designated nodes called input ports. The input is synchronous and each input bit is available only once. The input and output conventions are where-determinate but need not be when-determinate. We build on Thompson's $2-D$ VLSI model.

Extensions to 3-D VLSI: A number of researchers [Ros81], [LR86] have extended this model to 3-D VLSI, where the only additional assumption is that the grid can be of arbitrary size in each dimension.

Barakat and Reif's VLSIO model: This *VLSIO model* for 3-D optical (with the O standing for optics) computation is described in [BR87a], which combines the use of optics and electronic components in ways that models devices currently feasible. In this model, the fundamental building block is the optical box, consisting of a rectilinear parallelepiped (a 3D figure formed by six parallelograms) whose surface consists of electronic devices modeled by the 2-D VLSI model and whose interior consists of optical devices. Communication from the surface is assumed to be done via electrical-optical transducers on the surface. Given specified inputs on the surface of the optical box, it is assumed that the output to the surface is produced in 1 time unit. For lower bound purposes, we need only to assume that at most 1 bit can be transmitted in a unit time across any cross-section of unit area within the optical box. This assumption has considerable physical justification; for example Gabor established in [Gab61] that an optical beam of bounded cross section has a finite upper limit with respect to information rates. This upper limit is a function of the wave-length of light, the smallest effective optical beam cross-section area, and solid angle of divergence. A VLSIO device consists of a convex volume with a packing of optical boxes whose interiors do not intersect, but may be connected by wires between their surfaces. This allows for communication between two optical boxes. Note that the VLSIO model encompasses the 3-D VLSI model as a subcase: the particular subcase where each optical box is just a 2-D surface with no volume.

Modeling Energy Consumption: It is shown in [Kis82] and [Tya88] that when a wire of length k switches, it consumes energy $\Theta(k)$. In our $f(d)-$*energy model* we will quantize energy consumption as a function $f(d)$ of the distance d that a beam has to travel, with assumed certain reasonable characteristics: where $f(d) = 0$ when $b = 0$ (since a zero length link should consume zero energy) and $f(d)$ has continuous first and second derivatives, as justified below. The energy model depends on how the beam is modu-

lated. Let us first consider amplitude modulation. When an optical beam travels through a medium, it loses energy due to scattering and dissipation in the form of heat. We assume that there is some attenuation due to these effects, no matter however small it is. The net power loss per unit volume of space is proportional to the intensity of the beam, *i.e.*, $P = \alpha I$. The absorption constant α depends on the frequency of light. This gives rise to the Lambert's law: the intensity of the beam is

$$I(x) = I_0 e^{-\alpha x}$$

where I_0 is the initial intensity of the beam and x is the distance traveled by the beam. To determine the energy loss for a beam traversing a distance d, we evaluate the integral giving the energy loss as $f(d) = I_0 \int_0^d \alpha e^{\alpha(d-x)} dx = I_0(e^{\alpha d} - 1)$. Note that this function $f(d)$ is 0 when $b = 0$ and has continuous first and second derivatives which are $I_0 e^{\alpha d}$. This function $f(d)$ can be approximated by a polynomial (since the finite section of the Taylor expansion of $e^{-\alpha d}$ about $d = 0$ up to k terms is a well-behaved polynomial $P(d)$ of degree $k - 1$). For $\alpha d < 1$, only the linear components of the exponent's Taylor expansion are relevant and we can assume that the energy consumed is proportional to the distance traveled by a beam; this will be called the *linear energy model*. In the case of phase encoding, the amplitude remains constant, but the energy loss due to switching is a function of the same form.

From now on, we will treat a wire and a beam in a uniform way. This is because our non-linear energy function lower bounds apply only to a pure optical computation, *i.e.*, the electrical wires are not used to communicate a large part of the information. Otherwise, any VLSI system is also a VLSIO system without any optical components. Energy bounds for that extreme are already known. Our objective is to determine the bounds for the other extreme – an optical computation. Hence, the term *link* refers to either a wire or a beam in the following discussion.

Note that we have assumed certain reasonable characteristics of these energy functions.

We will further distinguish between *average case* and *worst case* energy. For a lower bound, we will only count the energy consumption due to switching of links. Let us assume that a link is carrying only one bit of information, just as in the case of wires in the VLSI model. Thus every link l can be assigned a binary state from $\{0, 1\}$. $f(d)$ units of energy are consumed when a link carries a 1 for distance d. Let $C = (V, W)$ be a VLSIO circuit, where V is the set of nodes in a communication graph and

W is the set of links connecting these nodes. Note that these nodes are unit area electro-optical devices. Let $\mathcal{I}$ be the set of input values to C. When input vector $\vec{x}_i \in \mathcal{I}$ is asserted, let the link $l_j \in W$ of length d_j carry $k_{i,j}$ 1's during the course of this computation. The energy consumption of the link in this computation, $E(C, \vec{x}_i)$, is given by $\sum_{l_j \in W} f(d_j) k_{i,j}$. Then the average energy $E_a(C)$ is its energy consumption averaged over all input vectors, which is $\frac{\sum_{\vec{x}_i \in \mathcal{I}} E(C, \vec{x}_i)}{|\mathcal{I}|}$. We can similarly define the worst case energy consumption $E_w(C)$ as $max_{\vec{x} \in \mathcal{I}} E(C, \vec{x})$, where the maximum is taken over all the input vectors.

Another distinction that we make concerns the number of times a link switches. If we can guarantee that in a given VLSIO circuit, every link will switch at most once, then the circuit is said to follow the Uniswitch model (USM). The pipelined computations with $O(1)$ period are USM computations. The more general model, where every link could switch $k \geq 0$ times is known as Multiswitch model (MSM). Note that our multiswitch model does not preclude cyclic communication graphs $C = (V, W)$ unlike the models used in [Kis82].

3 ET Trade-off

For a function f of n bits let the *Information complexity* $IC(n)$ be the number of bits that are needed to compute $f(x)$ by two distinct processes, each holding at most $1 + n/2$ the input bits, as defined in [Yao79]. Let a function f have information complexity . To reduce the clutter of notation, let IC refer to the information complexity of a function f, $IC_f(n)$. We first prove a cutting lemma that is a 3-dimensional variant of a lemma shown in [Tya89]. In this section, we use the term *box* to refer to a rectilinear parallelepiped rather than an optical box.

Lemma 1 *Let $C = (V, W)$ be a VLSIO circuit that computes f as described above, where V is the set of nodes in a communication graph and W is the set of links connecting these nodes. Let us assume that no input port reads more than $n/486$ input bits. There exist two sets of input ports I_1 and I_2 such that I_1 and I_2 each read at least $n/486$ input bits. Let C_{I_1} and C_{I_2} denote the smallest boxes containing I_1 and I_2 respectively. Let s_1 (s_2) be the surface area of C_{I_1} (C_{I_2}). Let d_x, d_y and d_z be the distance between C_{I_1} and C_{I_2} along x, y and z axes respectively. Then one of the following statements holds true.*

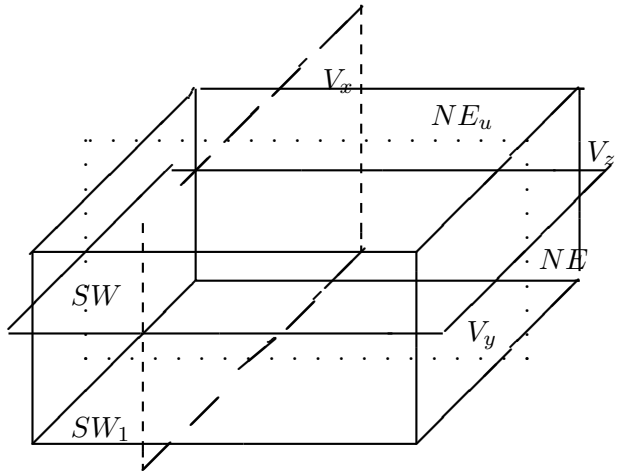

Figure 1: An Illustration of the Cutting Argument

1. $d_x \geq \min(\sqrt{s_1/6}, \sqrt{s_2/6})$.
2. $d_y \geq \min(\sqrt{s_1/6}, \sqrt{s_2/6})$.
3. $d_z \geq \min(\sqrt{s_1/6}, \sqrt{s_2/6})$.

PROOF: Since no input port reads more than $n/486$ input bits, we can find a vertical plane orthogonal to $x-$axis V_x bisecting the chip into two halves such that each half reads at least $n/3$ input bits. The argument for it is similar to the argument in Ullman [[Ull84], page 49]. Let us concentrate on the right-hand side of V_x. Once again, we can find a vertical plane orthogonal to the $y-$axis V_y that cuts the right-hand side of V_x into two parts such that each part reads at least $n/9$ input bits. The left-hand side of V_x is also cut into two parts by the plane V_y. At least one of these parts reads $n/6$ input bits. Without loss of generality (WLOG), let the south-west quadrant SW read at least $n/6$ input bits as shown in Figure 1. Let us further cut this south-west part with a plane orthogonal to $z-$axis H_z such that both the halves of SW read at least $n/18$ input bits. This plane also cuts the the north-east part NE into two parts. At least one of these parts reads at least $n/18$ input bits. WLOG, let the upper part NE_u read at least $n/18$ input bits. In the following, we will consider NE_u and SW_l, which is the lower part of SW. Let C_{I_1} and C_{I_2} be the smallest boxes containing the input ports in the SW_l and NE_u quadrants respectively, as shown in Figure 2.

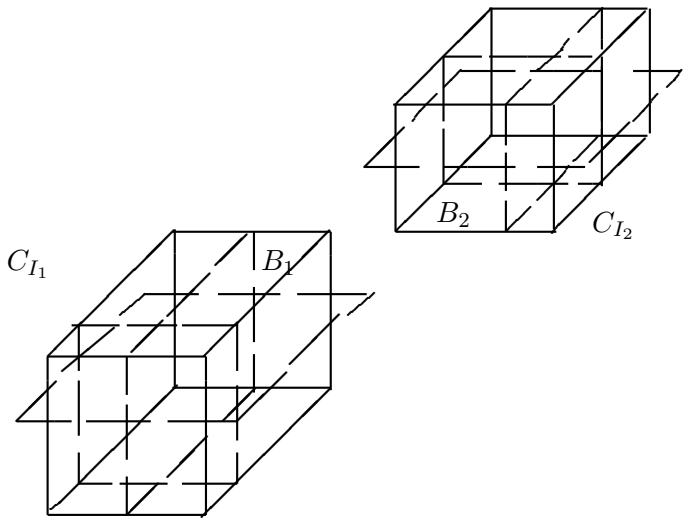

Figure 2: Further Cutting in the Cutting Lemma, where B_1 is a box inside C_{I_1} and B_2 is a box inside C_{I_2}.

Now cut both C_{I_1} and C_{I_2} as follows to create 8 boxes each reading at least $n/486$ input bits. First use a vertical plane V'_x to cut C_{I_1} into two parts each reading at least $n/54$ input bits. Then cut each of these parts with a plane orthogonal to y axis to get four parts each reading at least $n/162$ input bits. Now use a plane orthogonal to the $z-$axis to cut each of these four parts into two each where each of the resultant 8 parts reads at least $n/486$ input bits. Now we have 16 boxes, with C_{I_1} containing 8 boxes and C_{I_2} containing the other 8 boxes. Let us pick a box with the least surface area of all the 16 boxes. Without loss of generality, let a box in C_{I_1} have the smallest surface area. Let this box be B_1 with the smallest surface area s as shown in Figure 2. Since we will be separating this box from one of the boxes in C_{I_2}, any other box in C_{I_1} can only have a larger separation. Consider the box B_2. It has surface area $s_2 \geq s$. Then either its length, or width or height is at least as large as $\sqrt{s/6}$. If the width of B_2 is at least $\sqrt{s/6}$ then the box to the right of B_2 is horizontally separated from B_1 by at least $\sqrt{s/6}$ where the surface area of B_1 is s. Otherwise if the length of B_2 is at least $\sqrt{s/6}$ then the box abutting B_2 along the $y-$axis is horizontally separated from B_1 by at least $\sqrt{s/6}$. Otherwise B_1 is vertically separated from the box above B_2 by at least $\sqrt{s/6}$.

All the other cases follow the same line of reasoning. **QED.**

We will use the following information-theoretic lemma proved in Tyagi [Tya88], [Tya89]. The lemma says that to transmit k information bits using $k' > k$ bits, at least $\Omega\left(\frac{k}{\log(k'/k)}\right)$ switching must be incurred.

Lemma 2 *Let k' and k be two positive integers such that $k' \geq k$. The average number of alternations in transmitting a k' bit encoding of k information bits, $A(k', k) \geq c(k/4\log[(k'/k) + 2])$, for some constant $c > 0$.*

3.1 3-D VLSI with Linear Energy Cost

We will first show a lower bound of $I^{3/2}$ on the ET product of a 3-D VLSI system with linear energy cost. Since VLSIO model is a generalization of the 3-D VLSI, these lower bounds also apply to VLSIO with linear energy cost. Note that these lower bounds hold both in the uniswitch as well as the multiswitch model.

Theorem 1 *The ET product of a 3-D VLSI system computing a function $f(n)$ with information complexity $IC(n)$ has a lower bound of $IC(n)^{3/2}$.*

PROOF: Let us assume that no input port reads more than $n/486$ input bits. If there exists such an input port then the energy-time product ET is $\Omega(n^2)$ giving us the stated lower bound trivially. Then using Lemma 1 we can find two boxes C_{I_1} and C_{I_2} such that they both read at least $n/486$ input bits. Without loss of generality let C_{I_1} have the smaller surface area of two, s. Then their separation is at least $\sqrt{s/6}$.

We build $\sqrt{s/6}$ bisections of the input bits read in C_{I_1} in the following way. Let $L(0)$ be the set of all the unit length link segments (wire segments in pure VLSI model) incident on the surface of C_{I_1}. There are at most $3s$ link segments in $L(0)$. Let $L(i)$ be defined recursively, for $i > 0$, as follows: $L(i)$ is the set of unit length link segments adjacent to the links in $L(i-1)$ and furthermore, not in $L(0), \ldots, L(i-1)$ and not inside C_{I_1}. These are the dotted cubes around C_{I_1} as shown in Figure 3. Since $L(0)$ contain at most $3s$ links and the surface area grows as r^2 for radius r, the number of links in $L(\sqrt{s/6})$, $|L(\sqrt{s/6})|$ is at most $12s$. This is also an upper bound on the number of links in $L(0)$, $L(1), \ldots L(\sqrt{s/6}-1)$. If the computation takes time T, then the number of bits crossing any of these shells cannot exceed $12sT$.

If the majority of the output bits are generated inside a ring $L(i)$, then the information about the input bits read in C_{I_2} needs to come into the

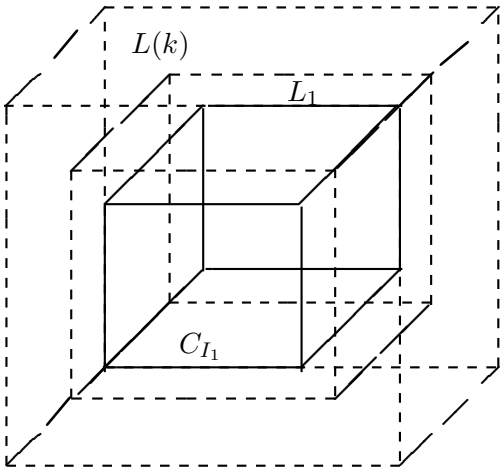

Figure 3: The Cubes Around C_{I_1} indicated by dotted lines.

shell. Otherwise, the information about the input bits read inside the shell needs to go out. In either case, since C_{I_1} and C_{I_2} read at least $n/486$ input bits, the number of information bits required is $\Omega(I)$. Hence the energy consumption at a shell $L(i)$ is $\Omega\left(\frac{I}{\log(12sT/I)}\right)$ by Lemma 2. Adding this consumption over $\Omega(\sqrt{s})$ shells gives a lower bound of $\Omega\left(\frac{I\sqrt{s}}{\log(12sT/I)}\right)$ on the energy consumption. Then the ET product has a lower bound of $\Omega\left(\frac{I\sqrt{s}T}{\log(12sT/I)}\right)$. This is bounded by $\Omega\left(\frac{I\sqrt{sT}}{\log(12sT/I)}\right)$. For the correctness of computation, the number of bits going across a shell sT should be at least $\Omega(I)$. For $sT \geq I$, $\frac{\sqrt{sT}}{\log(12sT/I)}$ attains a minimum at $sT = I$. Then $\Omega(I^{3/2})$ forms a lower bound on ET. **QED.**

We will show in Section 4 that a barrel-shifter (which circularly shifts the elements of a vector) has an ET product of $n^{3/2}\log n$. Note that for barrel-shifting, Theorem 1 gives a lower bound of $n^{3/2}$. This lower bound can be improved if the circuit is known to be systolic in the following sense. If every link segment in the circuit has $O(1)$ length then we call it *systolic*. Note that this implies that an information bit traveling across a distance d takes time $\Omega(d)$. In Section 4 we show a systolic 3-D barrel-shifter with $ET = O(n^{5/3})$.

Lemma 3 *A systolic 3-D VLSI circuit computing a function f with information complexity I has an ET product of $\Omega(I^{5/3})$.*

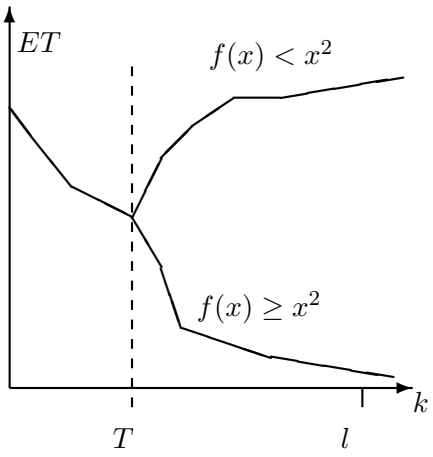

Figure 4: ET versus k

3.2 2-D VLSI with Non-Linear Energy Cost

In the classical 2-D VLSI energy model [Kis82], when a link of length l switches we charge $\Theta(l)$ switching energy. Let us enhance this model to permit non-linear energy consumption. Thus when a link of length l switches, $f(l) \geq l$ switching energy is consumed. We first prove a crucial lemma that considers the minimum value of the ET product when a link of length l is switched. Note that for any $f(l)$, we can always reduce the energy consumption to $\Theta(l)$ by dividing the link of length l into l unit length segments. Then each segment consumes $\Theta(1)$ energy for a total of $\Theta(l)$ energy. But this increases the time to drive the link from $O(1)$ to $\Omega(l)$. Let us state and prove the lemma.

Lemma 4 *Let a link of length l be driven using k stages of intermediate switches. The ET product for this link is minimized with $k = l$ when $f(l) \geq l^2$. For $l < f(l) < l^2$, the ET product is minimum with $k = min\{l, T\}$, where T is the time taken by the whole computation.*

PROOF: Let us reiterate the assumptions about the characteristics of the energy cost functions $f(l)$. These functions should be continuous and should have continuous derivatives. The energy consumption of a zero length link should be zero, $f(0) = 0$.

Clearly, the k segments of the link should all have the same length. Otherwise, the energy cost is higher than the sum of a nonlinear function

of k variables, $\sum_{i=1}^{k} f(l_i)$ with the condition $\sum_{i=1}^{k} l_i = l$ is minimized with $l_i = l/k$. Hence let us consider the scenario when k drivers with each one driving a segment of length l/k are introduced. Then the energy required is given by $k\,f(l/k)$. The time taken is $\mathbf{min}\{k, T\}$, which is $O(k+T)$. We wish to analyze the ET cost for the complete range of values for k and find the value of k that minimizes it. Figure 4 shows the rough shape of this plot.

$k \leq T$: For all functions $f(x) \geq x$, ET is a non-increasing function of k for $k \leq T$. For $k \leq T$, the time is at least T, while the energy is given by $k\,f(l/k)$. Thus the ET product is $g(k) = kT\,f(l/k)$. We claim that $g(k)$ has a derivative that is non-positive for all $1 \leq k \leq l$. $g'(k) = T\left[f\left(\frac{l}{k}\right) + \frac{l}{k}f'\left(\frac{l}{k}\right)\right]$. Since $T \geq 0$, to show that $g'(k)$ is non-positive, we need to show that $\left[f\left(\frac{l}{k}\right) + \frac{l}{k}f'\left(\frac{l}{k}\right)\right]$ is non-positive. Replacing l/k by x, we get $g_1(x) = f(x) + xf'(x)$ for $1 \leq x \leq l$. Since $f(x) = 0$ for $x = 0$, $g_1(0) = 0$. We show that for $f(x) \geq x$ and $x \geq 0$, $\frac{\delta g_1(x)}{\delta k}$ is non-positive, for any $\delta k > 0$. $\frac{\delta g_1(x)}{\delta k}$ is $(-\frac{x}{k})\ [2f'(x) + xf''(x)]$ where $f'(x)$ and $f''(x)$ are the first and second derivatives of $f(x)$ with respect to x. Note that for $f(x) \geq x$, $f'(x), f''(x) \geq 0$. Since $l, k \geq 0$, this derivative is non-positive proving our claim.

$k \geq T$: In this case, ET is a non-increasing function of k for $f(x) \geq x^2$ and non-decreasing otherwise. Here the time is $\Omega(k)$. Thus a lower bound on ET product is $g(k) = k^2 f(l/k)$. Let us consider $g'(k) = 2kf(l/k) - lf'(l/k)$. The sign of $\frac{g'(k)}{l} = \frac{2k}{l}f\left(\frac{l}{k}\right) - f'\left(\frac{l}{k}\right)$ is the same as that of $g'(k)$. Again substituting $x = l/k$, we get $g_1(x) = \frac{2f(x)}{x} - f'(x)$ for $1 \leq x \leq l$. The sign of $g_1(x)$ is the same as that of $g_2(x) = 2f(x) - xf'(x)$. Also $g_2(0) = 0$. Let us analyze the derivative of $g_2(x)$. Note that $\frac{\delta g_2(x)}{\delta k} = (-\frac{x}{k})\ [f'(x) - xf''(x)]$. This derivative is positive for $f(x) < x^2$ and is non-positive for $f(x) \geq x^2$.

These two cases provide us with enough information to prove the lemma. This shows that the ET minima for a function $f(l) \geq l^2$ occurs for the highest value of k which is l. But for a function $f(l) < l^2$, there is a trough in the ET curve at $k = T$. Thus if $T < l$, then $k = T$ provides a minima for ET. Otherwise, since k cannot exceed l, we are in the subcase $k < T$. Then l segments $(k = l)$ make the ET product the least possible.

QED.

Now we can prove a lower bound on the ET product in the $f(l)-$ energy model. Tyagi [Tya89] shows that with linear energy costs $ET = \Omega(I^2)$ for both USM and MSM models. But, surprisingly, the ET product of the uniswitch and multiswitch models differs for the nonlinear energy costs. For the multiswitch case, ET is shown to be $\Omega(I^2)$ for any nonlinear energy function. In the uniswitch case, the ET product increases with the energy function for an energy function $l < f(l) \leq l^2$. It levels off at $\Omega(I^3)$ for the energy cost functions $f(l) > l^2$.

Theorem 2 *Let a VLSI circuit compute a function with the information complexity I. The multiswitch ET product is bounded by $\Omega(I^2)$. In USM, for an energy function $l < f(l) < l^2$, the ET product has a lower bound of $\Omega(I\, f(I))$. For an energy function $f(l) \geq l^2$, the USM lower bound is $\Omega(I^3)$.*

Proof: A 2-dimensional analog of Lemma 1 proven in Tyagi [Tya89] provides two rectangles R_{I_1} and R_{I_2} where each rectangle reads at least $n/71$ input bits. Moreover these rectangles are separated by a distance that is an order of the smaller perimeter of the two rectangles. Without any loss of generality, let p be the smaller perimeter (of R_{I_1}). Notice that I information bits need to go across this channel that is $\Omega(p)$ wide. Let the computation take time T. Also notice that there are at most $O(p)$ links in any ring around R_{I_1}. Then at most pT bits can be sent over the course of computation. By Lemma 2, at least $\Omega\left(\frac{I}{\log(pT/I)}\right)$ bits switch. The path any bit follows through this channel has $\Omega(p)$ length. From Lemma 4 above, the ET product of a switched link and hence of the computation is minimized when energy is $p\, f(1)$ $[T\, f(p/T)]$ for $f(l) \geq l^2$ $[l < f(l) < l^2]$. Consider the MSM lower bound first. Let us first consider the case when $f(l) \geq l^2$. Then the ET product is $\Omega\left(\frac{I\, f(1)\, pT}{\log(pT/I)}\right)$. For $pT \geq I$, as required for the correctness of the computation, $\frac{f(1)pT}{\log(pT/I)}$ has the minimum value at $pT = I$. Then ET is $\Omega(I^2)$. Now we consider the case $l < f(l) < l^2$. Then the ET product is $\Omega\left(\frac{I\, f(p/T)\, T^2}{\log(pT/I)}\right)$. Since $f(p/T)T^2 \geq pT$ for $f(l) > l$, once again $\frac{f(p/T)\, T^2}{\log(pT/I)}$ is minimum when $pT = I$. The expression $f(p/T)T^2$ with $pT = I$ has a minimum value of I with $p = T = \sqrt{I}$. This shows a lower bound of $\Omega(I^2)$ on ET.

Now let us consider the uniswitch case. The subcase $l < f(l) < l^2$ follows from the previous discussion as follows. The ET product $\Omega\left(\frac{I\, f(p/T)\, T^2}{\log(pT/I)}\right)$ is minimum when $pT = I$. Due to uniswitch property, we also know that $p \geq I$. Then the minimum value of the expression $f(p/T)\, T^2$ under the

constraints $pT = I$ and $p \geq I$ is $f(I)$. This proves an ET lower bound of $\Omega(I\, f(I))$. The other situation occurs for $f(l) \geq l^2$. Note that in USM every link can transmit at most one bit for the duration of the entire computation. Hence the total number of bits transmitted across a ring during the computation is $O(p)$. Then the ET product is $\Omega\left(\frac{I\, f(1)\, p^2}{\log(p/I)}\right)$ since the time taken is $\mathbf{max}(p,\, T)$. Since $p \geq I$, this is $\Omega(I^3)$. **QED.**

3.3 ET Bounds in the VLSIO Model

Now we are ready to prove some lower bounds in the VLSIO model. Notice that for the purposes of lower bound proofs, the VLSIO model is essentially 3-D VLSI model with a non-linear energy cost function $f(l)$. We have already developed techniques for these two cases in the previous discussion. We will be combining the techniques used in 3-D VLSI lower bounds and non-linear energy VLSI model lower bounds to get a $I\, f(\sqrt{I})$ lower bound on the ET product for VLSIO multiswitch model for $f(l) < l^{4/3}$. In MSM for $f(l) \geq l^{4/3}$, this lower bound is $\Omega(I^{5/3})$. In the uniswitch case with $f(l) \geq l^2$ $[l < f(l) < l^2]$ the lower bound is $\Omega(I^2)$ $[\Omega(I\, f(\sqrt{I}))]$.

Theorem 3 *Let a VLSIO system compute a function with the information complexity I. In MSM, the ET product of this computation is lower-bounded by $\Omega(I\, f(\sqrt{I}))$ for an energy function $l < f(l) < l^{4/3}$. An energy function $l > l^{4/3}$ leads to a lower bound of $\Omega(I^{5/3})$ for the multiswitch case. The uniswitch circuits have an ET product of $\Omega(I^2)$ for $f(l) \geq l^2$ and $\Omega(I\, f(\sqrt{I}))$ for $l < f(l) < l^2$.*

PROOF: The proof is very similar to the proofs of Theorem 1 and Theorem 2. Let us consider the multiswitch case.

MSM: $l < f(l) < l^{4/3}$: As we saw in Theorem 1, Lemma 1 gives us $\sqrt{s}$ shells around a box C_{I_1} which reads at least $n/471$ input bits. This box is separated from another box C_{I_2} by a distance of at least $\sqrt{s}$ where s is the surface area of C_{I_1} and C_{I_2} also reads at least $n/471$ input bits. Let the computation take time T. Thus we have $\Omega\left(\frac{I}{\log(sT/I)}\right)$ paths of length $\Omega(\sqrt{s})$ that switch. By Lemma 4, the switching energy of these paths is $\Omega\left(\frac{I\, f(\sqrt{s}/T)\, T}{\log(sT/I)}\right)$. Then the ET product is $\Omega\left(\frac{I\, f(\sqrt{s}/T)\, T^2}{\log(sT/I)}\right)$. This product is minimized with $sT = I$. In addition, for $f(l) < l^{4/3}$, it is minimized when $T = O(1)$ and $s = \Omega(I)$. Then the lower bound is $\Omega\left(I\, f(\sqrt{I})\right)$.

MSM: $f(l) \geq l^{4/3}$: Here Lemma 4 dictates two cases: $l^{4/3} \leq f(l) < l^2$ and $f(l) \geq l^2$. In the first case, the lower bound is $\Omega\left(\frac{I\, f(\sqrt{s}/T)\, T^2}{\log(sT/I)}\right)$. This is minimized to $\Omega(I^{5/3})$ with $sT = I$ and with $s = I^{2/3}$. In the second case, the energy-time product lower bound is $\Omega\left(\frac{I\, f(1)\, \sqrt{s}T}{\log(sT/I)}\right)$. As shown in Baudet [Bau81] and Tyagi [Tya89], due to information-theoretic requirements $s \geq I^{2/3}$. This alongwith the fact that $sT = I$ gives a minimum value, shows a lower bound of $\Omega(I^{5/3})$.

USM: $l < f(l) < l^2$: As in the first subcase, the ET lower bound here is $\Omega\left(\frac{I\, f(\sqrt{s}/T)\, T^2}{\log(sT/I)}\right)$. There are two exceptions due to uniswitch property. The maximum number of bits transmitted is $O(s)$ rather than $O(sT)$. Since each link can carry at most one bit, $s \geq I$. These two imply a lower bound of $\Omega\left(I\, f(\sqrt{I})\right)$.

USM: $f(l) \geq l^2$: Again USM implies that $O(s)$ bits are transmitted and that $s \geq I$. Then an ET lower bound is $\Omega\left(\frac{I\, f(1)\, \sqrt{s}\left(T+\sqrt{s}\right)}{\log(s/I)}\right)$. This is $\Omega\left(\frac{I\, f(1)\, s}{\log(s/I)}\right)$. For $s \geq I$, this is minimized for $s = I$ and is given by $\Omega(I^2)$.

QED.

4 Upper Bounds

We show upper bounds on the ET complexity of *barrel-shifting*, which is to circularly shift a vector of n bits. Note that since shifting is a transitive function, its information complexity $IC(n)$ is $\Omega(n)$. Hence by Theorem 2, the multiswitch ET is $\Omega(n^2)$. The USM ET product is $\Omega(n\, f(n))$ for $l < f(l) < l^2$ and $\Omega(n^3)$ for $f(l) \geq l^2$.

A barrel shifter, a butterfly type of network, can be implemented in area $O(n^2)$ and it takes time $O(\log\ n)$ to shift one set of data, as described in Ullman [Ull84]. For each of the $\log n$ bits in the control input (shift value) there is a stage in the barrel shifter. The ith stage shifts the data bits by 2^{i-1} bit positions. Thus, in general, the wires corresponding to the ith stage, for $1 \leq i \leq \log n$, have a length of $\Theta(2^{(i-1)})$. The wires for the stage corresponding to the most significant bit (a shift by $n/2$) have length $\Theta(n/2)$. On average, half of the wires at each stage switch. *This is a uniswitch circuit.* For $f(l) \geq l^2$, every wire can be divided into unit length segments and driven

by a separate device. Then the time is $O(n \log n)$. Then the average energy consumption for this circuit is $\Theta(A)$, which is $O(n^2)$. This gives an upper bound of $O(n^3 \log n)$ on ET product for shifting. Compare it with the lower bound of $\Omega(n^3)$. For $l < f(l) < l^2$, the energy is $O(n\, f(n))$ and the time is $\log n$ giving an ET product of $n\, f(n) \log n$.

For the multiswitch case, consider the square shifter, which is a barrel-shifter loyed out in a 2D area, as described in detail in Ullman [[Ull84], page 69]. The n bits to be shifted are stored along a $\lceil\sqrt{n}\rceil \times \lceil\sqrt{n}\rceil$ array. We number the rows bottom-up from 1 to $\lceil\sqrt{n}\rceil$ and columns left-right from 1 to $\lceil\sqrt{n}\rceil$. The least significant input bit x_1 is stored at $(1,1)$ position, $x_{\lceil\sqrt{n}\rceil}$ at $(\lceil\sqrt{n}\rceil, 1)$, $x_{\lceil\sqrt{n}\rceil+1}$ at $(1,2)$, and x_n at $(\lceil\sqrt{n}\rceil, \lceil\sqrt{n}\rceil)$. The first half control bits $c_{\log n} \ldots c_{(\log n)/2+1}$ specify the horizontal shift amount and the least significant half bits specify the vertical shift amount. It takes area $\sqrt{n} \times \sqrt{n}$ and time $O(\sqrt{n})$. For each horizontal shift through a column, half of $\sqrt{n}$ bits are expected to switch, which switches half the area. The expected shift amount is $\sqrt{n}$, giving rise to $O(n^{3/2})$ average switching energy for both horizontal and vertical shifting. The time taken is $O(\sqrt{n})$. Thus we have an ET product of $O(n^2)$ which matches the lower bound.

Let us consider the three-dimensional case. By Theorem 1, shifting has a lower bound of $\Omega(n^{3/2})$ in the 3D-VLSI model. For a systolic system, the lower bound is $\Omega(n^{5/3})$. A three-dimensional version of the barrel shifter will have a $\sqrt{n} \times \sqrt{n}$ square for each stage. The links from stage i to stage $i-1$ are $2^{i/2}$. Hence the volume of this shifter is $O(n^{3/2})$. Time taken is still $\log n$ giving an ET product of $n^{3/2} \log n$. A three-dimensional analog of the square shifter will have dimensions $n^{1/3} \times n^{1/3} \times n^{1/3}$. By our definition, it is a systolic shifter. It has an ET product of $n^{5/3}$.

In the VLSIO case, the 3-dimensional shifters described above serve the purpose. For the multiswitch case, in the range $f(l) < l^{4/3}$ the barrel shifter serves the purpose. Its energy consumption is $n\, f(\sqrt{n})$ and the time taken is $\log n$. For the range $l^{4/3} \leq f(l)$, the cubic shifter has an ET product of of $n^{5/3}$. In USM, the 3-D barrel shifter has an ET product of $n\, f(\sqrt{n}) \log n$ for $f(l) < l^2$. For $f(l) \geq l^2$, every link is divided into unit length segments giving $E = n^{3/2}$ and time $\sqrt{n}$. This gives an ET product of n^2.

5 Average Energy Case

¿From power supply design and heat dissipation perspective, it is the average case energy consumption that matters. The lower bounds derived in

Section 3 still hold if we replace I by the function entropy H_f using the techniques in [Tya89]. The function entropy measures the skew of the distribution of the output values with respect to input values.

Definition 1 *For a function f, let n_y be the number of input values that give rise to the output value y. Assuming n-bit input values, the entropy of f H_f is $\sum_{y \in O} \frac{n_y}{2^n} \log\left(\frac{2^n}{n_y}\right)$, where O is the set of output values.*

Then the following corollaries can be stated.

Corollary 1 *The average ET product of a 3-D VLSI computation is $\Omega\left(H_f^{3/2}\right)$.*

Corollary 2 *For a VLSIO system computing a function f, its MSM ET product is $\Omega(H_f\, f(\sqrt{H_f}))$ for an energy function $l < f(l) < l^{4/3}$. An energy function $l > l^{4/3}$ leads to a lower bound of $\Omega(H_f^{5/3})$ for the multiswitch case. The uniswitch circuits have an ET product of $\Omega(H_f^2)$ for $f(l) \geq l^2$ and $\Omega(H_f\, f(\sqrt{H_f}))$ for $l < f(l) < l^2$.*

Tyagi [Tya89] shows a lower bound of $\Omega(n)$ on H_f for shifting, DFT and integer multiplication. $\Omega(n)$ is also a lower bound on the information complexity I of these functions. This implies that an asynchronous computation has no asymptotic advantage over synchronous computation for these functions.

6 Conclusions and Acknowledgments

This paper provided various lower bounds and energy-time trade-off on energy consumption for some some fundamental optical computations. Considerable more work needs to be done to provide upper bounds on energy consumption for other optical computations. A preliminary version of this paper appeared in John H. Reif and A. Tyagi, Energy Complexity of Optical Computations, 2nd IEEE Symposium on Parallel and Distributed Processing (SPDP90), Dallas, TX, pp. 14-21, December 1990.

The research of A. Tyagi was supported by NSF Grant #MIP-8806169, NCBS&T Grant #89SE04 and a Junior Faculty Development Award from UNC, Chapel Hill. J. H. Reif was supported by Supported by National Science Foundation Grants CCF-0432038, CCF-0432047, ITR 0326157, EIA-0218376, EIA-0218359, and EIA-0086015. DARPA/ARO contract DAAL03-88-K-0195, Air Force Contract AFOSR-87-0386, DARPA/ISTO contract N00014-88-K-0458, NASA subcontract 550-63 of primecontract NAS5-30428.

References

[ACR88] A. Aggarwal, A. K. Chandra, and P. Raghavan. Energy Consumption in VLSI Circuits. In *Proceedings of ACM Symposium on Theory of Computing*, pages 205–216. ACM-SIGACT, 1988. New York, NY, 1988. ISBN:0-89791-264-0. doi:10.1145/62212.62230.

[A2010] A.P. Ambs, Optical Computing: A 60-Year Adventure, Advances in Optical Technologies, vol. 2010, Article ID 372652, 15 pages, 2010. doi:10.1155/2010/372652

[Bau81] G. M. Baudet. On the Area Required by VLSI Circuits. In *Proceedings of CMU Conference on VLSI – VLSI Systems and Computations*, pages 100–107. CMU, Computer Science Press, 1981.

[BL85] R. Bennett and R. Landauer. Fundamental Physical Limits of Computation. *Scientific American*, pages 48–58, July 1985.

[BR87a] R. Barakat and J. Reif. Lower Bounds on the Computational Efficiency of Optical Computing Systems. *Applied Optics*, 26:1015–1018, March 1987.

[BR87b] R. Barakat and J. Reif. Polynomial Convolution Algorithm for Matrix Multiplication with Application for Optical Computing. *Applied Optics*, Vol. 26, No. 14, pp. 2707-2711, July 15, 1987.

[BG2008] B.D. Bingham and M.R. Greenstreet, Computation with Energy-Time Trade-Offs: Models, Algorithms and Lower-Bounds, Proc. Int'l Symp. Parallel and Distributed Processing with Applications, pp. 143-152, 2008.

[BG2012] B.D. Bingham and M.R. Greenstreet, Modeling Energy-Time Trade-Offs in VLSI Computation, IEEE Transactions on Computers, pp. 530-547, vol. 61 no. 4, April 2012

[BK1981] R.P. Brent and H.T. Kung, The Area-Time Complexity of Binary Multiplication, J. ACM, vol. 28, no. 3, pp. 521-534, 1981.

[EGC2004] J.C. Ebergen and J. Gainsley and P. Cunningham, Transistor Sizing: How to Control the Speed and Energy Consumption of a Circuit, Proc. 10th Int'l Symp. Asynchronous Circuits and Systems (ASYNC '04), pp. 7-16, Apr. 2004.

[Gab61] D. Gabor. Light and Information. In E. Wolf, editor, *Progress in Optics*, pages 111–153. North-Holland, Amsterdam, Holland., 1961.

[Kia91] F. Kiamilev, P. Marchand, A. Krishnamoorthy, S. Esener, and S. Lee, Performance comparison between optoelectronic and VLSI multistage interconnection networks, Journal of Light-wave Technology, 9, 12, 1991.

[Kis82] G. Kissin. Measuring Energy Consumption in VLSI Circuits: a Foundation. In *Proceedings of ACM Symposium on Theory of Computing*. ACM-SIGACT, pp. 99 - 104, ACM New York, NY, 1982. ISBN:0-89791-070-2 doi:10.1145/800070.802181.

[Kis85] G. Kissin. Functional Bounds on Switching Theory. In *Chapel Hill Conference on VLSI*. U.N.C., Chapel Hill, Computer Science Press, 1985.

[Kis87] G. Kissin. *Measuring Energy Consumption in VLSI Circuits*. PhD thesis, Department of Computer Science, University of Toronto, Toronto Computer Science Technical Report 202/87, Toronto, 1987.

[Kis91] G. Kissin. Upper and lower bounds on switching energy in VLSI Journal of the ACM (JACM), pp. 222-254, Volume 38 Issue 1, Jan. 1991

[KA2009] V. Korthikanti and G. Agha. Analysis of Parallel Algorithms for Energy Conservation in Scalable Multicore Architectures, Proc. International Conf. Parallel Processing, pp. 212-219, Sept. 2009.

[KA2010] V. Korthikanti and G. Agha. Towards Optimizing Energy Costs of Algorithms for Shared Memory Architectures, Proc. 22nd ACM Symp. Parallel Algorithms and Architectures (SPAA2010), pp. 157-165, June 2010.

[KAG2011] V.A. Korthikanti and G. Agha, and M. Greenstreet. On the Energy Complexity of Parallel Algorithms, International Conference on Parallel Processing (ICPP), pp. 562 - 570, 2011.

[KKK2008] N. Khude and A. Kumar and A. Karnik. Time and Energy Complexity of Distributed Computation of a Class of Functions in Wireless Sensor Networks, IEEE Transactions on Mobile Computing, pp. 617-632, vol. 7 no. 5, May 2008.

[Kun79] H.T. Kung. Let's Design Algorithms for VLSI Systems. In *Proceedings of the Caltech Conference on Advanced Research in VLSI: Architecture, Design, Fabrication*, pages 65–90. Caltech, January 1979.

[Leo84] J. Leo. Energy Complexity in VLSI. Master's thesis, Informatics Department, Nijmegen University, The Netherlands, February 1984.

[LM81] T. Lengauer and K. Mehlhorn. On the Complexity of VLSI Computations. In *Proceedings of CMU Conference on VLSI – VLSI Systems and Computations*, pages 89–99. CMU, Computer Science Press, 1981.

[LR86] F. T. Leighton and A. L. Rosenberg. Three-Dimensional Circuit Layouts. *SIAM Journal on Computing*, 15(3):793–813, March 1986.

[M2001] A.J. Martin, Towards an Energy Complexity of Computation, Information Processing Letters, vol. 77, pp. 181-187, 2001.

[MNP2002] A.J. Martin and M. Nystrm and P.I. Pnzes, Et^2 : A Metric for Time and Energy Efficiency of Computation, Power Aware Computing, R. Melhem and R. Graybill, eds., Kluwer, 2002.

[LPZ98] K. Li; Y. Pan; S. Zheng (Eds.), Parallel Computing Using Optical Interconnections, Series: The Springer International Series in Engineering and Computer Science, Vol. 468, 1998.

[Pre83] F. P. Preparata. Optimal Three-Dimensional VLSI Layouts. *Math. Systems Theory*, 16:1–8, 1983.

[Ros81] A. L. Rosenberg. Three Dimensional Integrated Circuits. In *Proceedings of CMU Conference on VLSI*, pages 69–79. CMU, Computer Science Press, 1981.

[RT90b] J. Reif and A. Tyagi. a preliminary version of this paper appeared as Energy Complexity of Optical Computations, 2nd IEEE Symposium on Parallel and Distributed Processing (SPDP90), Dallas, TX, pp. 14-21, December 1990.

[RT90a] J. Reif and A. Tyagi. Efficient Parallel Algorithms for Optical Computing with the DFT Primitive. In *Proceedings of the Tenth Conference on Foundations of Software Technology & Theoretical Computer Science*. Lecture Notes in Computer Science, Springer-Verlag, December 1990. Published as Efficient Algorithms for Optical Computing with the discrete Fourier transform (DFT) primitive, Journal of Applied Optics, Vol. 36, pp. 7327-7340, 1997.

[Sah01] S. Sahni, Models and Algorithms for Optical and Optoelectronic Parallel Computers, 1999 International Symposium on Parallel Architec-

tures, Algorithms, and Networks (I-SPAN99) held in Perth/Fremantle, Australia, 2001

[ST86] L. Snyder and A. Tyagi. The Energy Complexity of Transitive Functions. Proceedings of the Twenty-fourth Annual Allerton Conference on Communication, Control and Computing, Allerton House, Monticello, University of Illinois, Illinois, 1986.

[Sto83] Q. Stout, Mesh connected computers with broadcasting, IEEE Transactions on Computers, 1983, 826-830.

[Tho79] C.D. Thompson. Area-Time Complexity for VLSI. In *Proceedings of ACM Symposium on Theory of Computing.* ACM-SIGACT, 1979.

[Tho80] C.D. Thompson. A Complexity Theory for VLSI, PhD dissertation, Carnegie Mellon Univ., 1980.

[Tya88] A. Tyagi. *The Role of Energy in VLSI Computations.* PhD thesis, Department of Computer Science, University of Washington, Seattle, 1988. *UWCS Technical Report Number 88-06-05.*

[Tya89] A. Tyagi. Energy-Time Trade-Offs in VLSI Computations. In *Proceedings of the Ninth Conference on Foundations of Software Technology & Theoretical Computer Science (FSTTCS89).* Lecture Notes in Computer Science #405, Springer-Verlag, pp. 301-311, 1989.

[Ull84] J.D. Ullman. *Computational Aspects of VLSI.* Computer Science Press, Rockville, Md., 1984.

[WS98a] C. Wang and S.Sahni, OTIS optoelectronic computers, in Parallel Computing Using Optical Interconnections, Kluwer, 1998, 99-116. Ed. Li and Zhenq.

[WS98b] C. Wang and S. Sahni, Basic operations on the OTIS-Mesh optoelectronic computer, IEEE Trans. on Parallel and Distributed Systems, 9, 12, 1998, 1226-1236.

[WS98c] C. Wang and S. Sahni, Basic operations on the OTIS-Mesh optoelectronic computer, IEEE Trans. on Parallel and Distributed Systems, 9, 12, 1998, 1226-1236.

[Yao79] A.C. Yao. Some Complexity Questions Related to Distributed Computing. In *Proceedings of ACM Symposium on Theory of Computing.* ACM-SIGACT, 1979.