Energy Aware Computing Through Probabilistic Switching: A Study of Limits

Krishna V. Palem
Center for Research on Embedded Systems and Technology,
School of Electrical and Computer Engineering,
Georgia Institute of Technology,
777 Atlantic Drive NW,
Atlanta, GA 30332-0250.
http://www.crest.gatech.edu

Abstract

The mathematical technique of *randomization* yielding *probabilistic* algorithms is shown, for the first time, through a physical interpretation based on statistical thermodynamics, to be a basis for energy savings in computing. Concretely, at the fundamental limit, it is shown that the energy needed to compute a single probabilistic bit or PBIT is proportional to the probability p of computing a PBIT accurately. This result is established through the introduction of an idealized *switch*, for computing a PBIT, using which a *network* of switches can be constructed. Interesting examples of such networks including AND, OR and NOT gates (or as functions, boolean conjunction, disjunction and negation respectively), are constructed and the potential for energy savings through randomization is established. To quantify these savings, novel measures of "technology independent" *energy complexity* are introduced—these parallel conventional machine-independent measures of computational complexity such as the algorithm's running time. Networks of switches can be shown to be equivalent to Turing machines and to boolean *circuits*, both of which are widely-known and well-understood models of computation. These savings are realized using a novel way of representing a PBIT in the physical domain through a group of classical *microstates*. A *measurement* and thus detection of a microstate yields the value of the PBIT. While the eventual goal of this work is to lead to the physical realization of these theoretical constructs through the innovation of randomized (CMOS based) devices, the current goal is to rigorously establish the potential for energy savings through probabilistic computing at a fundamental physical level, based on the canonical thermodynamic models of idealized monoatomic gases developed by Boltzmann, Gibbs and Planck.

I. INTRODUCTION

Concerns of *power* (or *energy*) consumption have become increasingly significant in the context of the design as well as the use of embedded and high-performance computing systems. To paraphrase Trevor Mudge, "Power (and energy) are first-class citizens in current considerations of computer system design." While devices, computer architecture and the layers of software that reside and execute at higher levels of abstraction (such as operating systems, run-time, compilers and programming languages) all afford opportunities for being *energy-aware*, the most fundamental limits are truly rooted in the physics of energy consumption–specifically in *thermodynamics*. Based on this premise, this paper embodies the innovation of models of computing for energy-aware algorithm design and analysis, establishing, for the first time, the following thesis central to this work: *the computational technique referred to as randomization, yielding probabilistic algorithms, now ubiquitous to the mathematical theory of probabilistic algorithm design and analysis, when interpreted as a physical phenomenon through classical statistical thermodynamics, leads to energy savings that are proportional to the probability p with which each primitive computational step is guaranteed to be correct (or equivalently to the probability of error, $(1-p)$).*

Historically, randomization was viewed as a mathematically very promising approach to algorithmic design and analysis elegantly stated by Schwartz [1]: "The startling success of the Rabin-Strassen-Solovay (see Rabin [2]) algorithm, together with the intriguing foundational possibility that axioms of randomness may constitute a useful fundamental source of mathematical truth independent of, but supplementary to, the standard axiomatic structure of mathematics (see Chaitin and Schwartz [3]), suggests that probabilistic algorithms ought to be sought vigorously." Since this prediction, randomization has proliferated in a range of areas primarily centered around the theoretical foundations of computer science.

At the heart of the new development in this work is the definition of an abstract *energy-aware switch* (in Section V-A), which is the *first* contribution of this work. A switch is a device for realizing computations that are functions of a single bit. While switches are provably building blocks for constructing boolean gates as well as for describing algorithms, in our context, they serve as idealizations for modeling energy consumption limited solely by classical thermodynamics. Our basic idealizations of a "switch" and "switching" are viewed as transformations to the *state* of a physical device capable of altering its Boltzmann (physical) entropy, with a well-defined accompanying expenditure of energy consistent with the laws of thermodynamics. We note that in realizing computations, switching will be used to alter the current bit or value, say 0, to some other value, say

This work is supported in part by DARPA under seedling contract #F30602-02-2-0124

1. In keeping with traditional idealizations, our switches are dissipationless and switching is always performed at thermal equilibrium, as detailed in Section VII. All computation will be viewed as being a composition of such elemental one-bit switching changes.

In this work and for the first time, the characterization of switching is based on (classical) statistical thermodynamics, referred to as *microphysics* (see Balian [4] for example) in literature. This statistical foundation is essential to proving our fundamental theorems (in Section VIII-B) which, through the *second* contribution of this work, show that *the energy consumed by deterministic switching is never less than* $(-\kappa t \ln 2)$ *Joules, referred to often as the "fundamental limit", whereas the energy consumed by a probabilistic switch with an associated probability of error of (1-p) can be as low as* $(-\kappa t \ln 2p)$ *Joules at the idealized limit*. Here, k is the well-known Boltzmann's constant, T is the *temperature* of the thermodynamic system, and $\ln$ is the natural logarithm. Thus, we show that randomized computing offers the potential for energy savings of $\kappa t \ln \left(\frac{1}{p}\right)$ Joules per primitive switching step. By basing our development on statistical thermodynamics rather than on the more familiar deterministic models of energy consumption known previously (see Section II below), our switches are naturally randomized; they do not need an explicit random source, crucial to the development of the theory of probabilistic algorithms (see Vazirani and Vazirani [5] for example).

All of the energy bounds developed here are based on a novel definition of *value*—the manner in which a single "bit" 0 and 1 is stored in a switch (as detailed by this author in [6] and [7]). In this work, as detailed in Section VI, a bit is represented in the in a switch through a classical *microstate*, and the output value of switching is determined through an instantaneous (classical) "measurement", that detects the existence of a microstate. An example of an instantaneous measurement is the detection of the position of a molecule of gas in a cylinder containing it, as described by Szilard [8]. This novel approach to representing the value represents a *probabilistic bit* or (PBIT) for short), constitutes a *third* contribution of this work. This novel representation is central to determining the gains of randomization at the fundamental limit, and will be contrasted with the traditional approach to representing PBIT s as voltages and measured as averages as characterized by Stein [9] and by Meindl [10]. The relationships between these two approaches will be outlined in Section XIV, and are the subject of the work developed by this author in collaboration with Cheemalavagu and Korkmaz [11].

Given a switch sw, we show (in Section IX) a systematic method for constructing *networks* of switches, to realize computations. Each switch sw has an *input* value which is either 0 or 1 and an explicit *enabling signal* that determines whether a switch is "active" or "inactive". In turn, output of sw is a value from the set $\{0, 1\}$ and possibly, an output enabling signal to a successor switch. A network $\mathcal{N}$ is central to defining a technology independent *energy complexity* of a switch, and constitutes a *fourth* contribution; these complexity measures are introduced in Section IX-C.

In Section X, and moving to the *fifth* contribution, we prove that a standard two-input AND function requires at least two "energy consuming" switches in the deterministic case and hence its *energy complexity* is bounded below by 2. Similarly, we also show (Section XI-B.1) that the same function can be realized using probabilistic switches such that the *expected energy complexity* is $(1 + p) \cdot \kappa t \ln(2p)$, with an associated error probability of $(1 - p)$. Through a straightforward construction, we show in Section VII that logical negation or the NOT function can be realized through a single switch, and therefore, its energy behavior in the deterministic and probabilistic cases is identical to those claimed earlier on.

Reminiscent of combinational logic networks whose power has been systematically studied relative to models of computing such as Turing machines (see early examples due to Pippenger and Fischer [12], Pippenger [13]), the model of a network $\mathcal{N}$ used in this paper uses an enabling signal in the absence of which a switch is not enabled. In keeping with current convention, throughout this paper, we will use the term *circuit* to denote the more traditional combinational logic "network" wherein measures such as *size*, *depth* and *width* were of concern, whereas energy was not; by contrast, the term *network* will be used to denote a structure built using our switches with enabling signals. In Section XIII, we sketch a relationship between networks of switches and circuits, as well as models of energy-aware algorithm design and analysis introduced by this author in an earlier work [6]–the randomized $b$it-level $r$andom $a$ccess $m$achine (or $RaBRAM$).

II. Historical Remarks, Comparisons and Perspectives

The energy behavior of switching has its roots in thermodynamics, whose history traces back to the works of Carnot [14] and Clausius [15] in the early nineteenth century, leading to significant developments in the early part of the twentieth century. Influenced significantly by Maxwell (see [16] and [17] for example), the current statistical interpretation of thermodynamics was first introduced by Boltzmann [18] and was later developed by Gibbs [19] and Planck [20]. In the context of realizing randomization physically, the work of Boltzmann leading to the definition of *thermodynamic entropy* is especially relevant to us.

The notion of a value such as 0 or 1 being modeled in a physical system with a single molecule dates back to 1929 and is attributed to Szilard [8]. In particular, his work and that of several subsequent physicists was motivated by a need to explain the celebrated *Maxwell's demon* [8] paradox which purported to, through a thought experiment, violate the inviolate *second law* of thermodynamics. Several subsequent authors credit Szilard with having invented the modern notion of a "bit" and a machine with two "states". While other celebrated researchers including von-Neumann [21] are credited with having observed that the minimum energy needed to compute a bit is $-\kappa t \ln 2$ Joules, it was Landauer [22] who took a very big step towards

clarifying the Maxwell's demon paradox in his widely known work. In doing so, he also explicitly laid the foundations for the (more) modern field of the thermodynamics of computation. At the heart of Landauer's work is the characterization of *logical irreversibility*, which is also a property of switching in this work.

Bennett [23] had important breakthroughs wherein, by taking advantage of the notion of *thermodynamic reversibility*, constructed conceptually viable and idealized models for reversible computing, wherein computation can proceed through energy recovery. This is accomplished through logically reversible computations. Subsequently, Fredkin and Toffoli [24] demonstrated logical gates that exhibit the same property. By contrast, in all of our work, we model switching as being based on *non-recovering* modes of execution—energy once expended by a switching step is not recovered, even if such recovery is possible.

Moving closer to realizations of switches based on familiar constructs such as transistors within the context of studying the inherent energy needed by deterministic switching, Meindl [10] established fundamental limits and derived energy lower-bounds. The significance of the Meindl's work, which continues the philosophical tradition set by Szilard, von-Neumann and Landauer, among others, is the ability for the first time to model a switch in an idealized manner—without dissipation for example, much as we do—while, at the same time, reconciling the delicate and pragmatic balance needed to model the realities of modern semiconductor devices. In doing so, novel techniques based on an inductive inverter-chain based argument were developed, to model distinguishability of a value of 0 from a value of 1. This technique is a crucial step in bounding the energy consumed from below. While previous work does so implicitly, Meindl [10] is also the first in this series to provide an *explicit* physical construction of a switch as an inverter, within the context of proving a bound on the minimum energy needed, thus making the bounds rather concrete.

The inherent energy bounds for deterministic switching developed in this work (in Section VIII), while having the same conceptual goal as Meindl's approach, are distinct in a three-fold manner. First, our idealizations and hence energy limits are based on the more fundamental energy behavior of idealized monoatomic gases [18], wherein the energy cost of computing a deterministic bit is a special case of that computing a PBIT with some associated probability p. As a result of this first difference, our representation of a bit and hence the concomitant energy bounds support randomization naturally; a study of the energy characteristics of probabilistic switching is an entirely novel contribution of this work. Furthermore, as outlined in Section XIV, the gas-based idealizations used in this paper yield potentially lower limits to computing PBIT s through switching, than those using the representations using the previously mentioned conventional approaches following Stein [9] and Meindl [10].

A consistent theme in all of the previous work is that all computation and hence the value of a bit being computed is deterministic, and thus its physical instantiation is *macrophysical* [4] and not subject to statistical interpretation. Hitherto, the underlying assumptions preclude the possibility that a switch could be designed to be deliberately erroneous with some probability say q, since computation, starting with Turing, was considered to be essentially a deterministic activity. It is not an exaggeration to say that determinism is deeply ingrained in human intuition in considerations of computing. This is perhaps the best explanation as to why an alternate style of computing—the counterintuitive notion of probabilistic algorithms that compute a value that could be sometimes wrong—was not considered until the 1970s from a mathematical perspective by computer scientists. This despite the readily available statistical interpretation of the value of a bit in Szilard's own construction of a molecule defining value, based on its location in a volume of gas.

In the context of computing, Rabin and Scott's influential paper [25] introduced *non-determinism* and broke with this tradition of determinism definitively. Subsequently, Rabin [26] also took the important step of explicitly introducing probability into the definition of an automaton [26] and studied its expressive power and relationship to a deterministic finite-state machine. Attributable perhaps to the deep seated belief in determinism in computing in general, over a decade elapsed before the role of probability became prevalent in the algorithmic computing domain following the influential work of Karp within the context of average-case analysis [27], Gill's characterizations of probabilistic complexity classes [28] as well as the innovation of the Rabin-Solovay-Strassen algorithm referred to earlier.

Feynman's exposition of the thermodynamics of computing [29] uses the gas-based model for switching. All of the developments in this paper are based on a physical system (and a phase-space) akin to that associated with Szilard's construction and Feynman's expositions, although our modeling of a state and a PBIT that it represents is statistical and hence microphysical, whereas previous approaches interpreted them to be deterministic.

III. Roadmap and reading guide

Some essential background facts from thermodynamics will be presented in Section IV. In Section V, we will introduce the mathematical definition of a switch and introduce the concept of switching. This will be followed in Section VI with a process of modeling a switch using an idealized microphysical device. Thus, the physical framework in this paper represents a switch whose thermodynamics are characterized by an idealized monoatomic gas, wherein switching is dissipationless in that it is achieved by a "quasistatic" process. Continuing, using the physical definition of a switch (from Section VI) in Section VII, we outline the physics, and the associated energy changes in Section VIII. Using these physical characteristics and results, in Section IX, we characterize the construction of networks of switches and hence boolean gates. Then, in Section X, we prove a lowerbound on the energy complexity of any deterministic realization of an *AND* gate. In Section XI, we introduce networks

of probabilistic gates, as well as the associated complexity measure of *expected energy complexity* and demonstrate energy savings for an *AND* gate construction through randomization. A comprehensive example of a switch construction using an ideal gas as the physical realization is shown in Section XII. In Section XIII, we characterize the relationship of our networks to established models of computing such as Turing machines and circuits, as well as traditional complexity measures such as running time and size. The degree to which a monoatomic gas can be viewed as a naturally ergodic system has serious implications to the statistical formulations of statistical thermodynamics. To clarify this issue, for completeness, in Section XV, we briefly review this deeper relationship between the foundations of the theory of idealized gases as outlined in Boltzmann's work for example [30], and two celebrated mathematical results—the recurrence paradox and the ergodicity hypothesis.

IV. Brief review of statistical thermodynamics

This section serves as a succinct introduction to the definitions and concepts of thermodynamics, leading to two main theorems (Theorems 4.5 and 4.6), which establish a relationship between energy, thermodynamic entropy, and the change of state of a thermodynamic system. Using these, the relationships between *energy change* and *entropy* (both defined below) will be used to model the energy changes of switching in Section VIII.

Classical thermodynamics is traditionally presented in a *phenomenological* manner in keeping with the best traditions of reductionism with observable properties such as pressure, volume and temperature being derived from a few underlying laws. In this work and perhaps in all of its connection to computing, our interest is in the amount of energy *expended* in making a certain kind of change to information, with the intention of modeling the change through the notion of physical switching. Thus, a more *transformational* view of statistical thermodynamics is presented below, with the expectation that it will better serve the purpose of using it in the context of characterizing switching.

A. Roadmap of this section

In Section IV-B below, several terms are introduced which define the thermodynamic universe of discourse and its attributes. In Section IV-C, *thermodynamic transformations* in this universe are defined. The theorems in Section IV-D establish the relationship between the change in entropy and energy for a single transformation. In Section IV-E, we generalize this development to be applicable to entire sequences of transformations. The treatment of this classical material is necessarily brief and is included here merely for completeness. While an interested reader is invited to peruse all the details, one who is primarily interested in the computational aspects in the remainder of this paper, can proceed by using Theorem 4.5 and its re-interpretation as Theorem 8.1 in Section VIII-A. For a more elaborate treatment of the subject, the reader is referred to standard books by Balian [4], Joos [31], Kittel [32], Boltzmann [30], Hill [33], and Mihalas [34].

B. A Phase Space and Its Microstates

In classical thermodynamics, a *thermodynamic system*, $\mathcal{T}$, is a set of N identical elements. Attributes of these elements are called the *classical observables*. For a single element, these observables are defined to be its *position* and *momentum*. The position of an element can be uniquely represented by its three co-ordinates p_x, p_y, p_z in the standard vector space $\mathbb{R}^3$. The variables q_x, q_y, q_z characterize the momenta along these three dimensions, and are represented by the three dimensional vector space Ω. Concisely, all the observables of an element can be represented as a point in the six-dimensional vector space $\mathbb{R}^3 \times \Omega$, where each point in this vector space is a six tuple $< p_{x_i}, p_{y_i}, p_{z_i}, q_{x_i}, q_{y_i}, q_{z_i} >: 1 \leq i \leq N$ and yields a particular value for each of the six variables associated with element i.

As an example [30] consider a box filled with an *ideal gas* which forms a thermodynamic system $\mathcal{T}$. An ideal gas is defined to be any gas that obeys the laws of the kinetic theory. The individual molecules of the gas form the identical elements constituting the thermodynamic system $\mathcal{T}$. (At any time instant τ, the position and momentum of each of the molecules form the classical observables.)

The entire system $\mathcal{T}$ of N elements is canonically represented by a $6N$-dimensional vector space called the *phase space*, $\mathcal{P}$ of $\mathcal{T}$. A point in the phase space, and thus the current positions and momenta of all the N particles at any time instant τ, can be uniquely represented by a $6N$ tuple of variables denoted by

$$< p_{x_1}, p_{y_1}, p_{z_1}, q_{x_1}, q_{y_1}, q_{z_1}, \ldots, p_{x_N}, p_{y_N}, p_{z_N}, q_{x_N}, q_{y_N}, q_{z_N} >$$

For notational simplicity, let us represent this $6N$ tuple as $\mathbf{A} =< p_1, p_2, p_3, \ldots, p_{3N}, q_1, q_2, q_3, \ldots, q_{3N} >$ where $p_{i'=3i-2} = p_{x_i},\ p_{i''=3i-1} = p_{y_i},\ p_{i'''=3i} = p_{z_i}$; the q_i's are similarly defined. The phase space $\mathcal{P}$, where each of the $6N$ dimensions is finitely bounded by the physical configuration of $\mathcal{T}$, is the *accessible phase space*. By abuse of notation, we will use the phrase phase space, and the symbol $\mathcal{P}$ to denote the accessible phase space.

Referring back to the earlier example, a single point in the phase space $\mathcal{P}$ corresponding to the (ideal gas) thermodynamic system $\mathcal{T}$ outlined above, completely characterizes the positions and momenta of each of the constituent molecules. The number of points in the phase space is uncountable, since the value of each of the variables in the $6N$ tuple characterizing a point in the phase space can vary continuously within a range along the position and momentum dimensions. However, following

Boltzmann [30], from the perspective of the statistical analysis of the physics of phase spaces, this vector space is decomposed to a discrete set of "volume elements".

At any instant of time τ, the classical observables of a thermodynamic system $\mathcal{T}$ are uniquely characterized by a single volume element Δv_i. Let $\hat{p}_{i'} = (p_{i'}, p_{i'} + \Delta p_{i'})$, $\hat{p}_{i''} = (p_{i''}, p_{i''} + \Delta p_{i''})$, $\hat{p}_{i'''} = (p_{i'''}, p_{i'''} + \Delta p_{i'''})$; the intervals $\hat{q}_{i'}$, $\hat{q}_{i''}$ and $\hat{q}_{i'''}$ are defined analogously, and the $6N$ intervals defined above form the sides of the $6N$ dimensional hypercube Δv_i. Consider a partitioning of the phase space into identical, non-overlapping volume elements. Such a partitioning will be referred to as a *Boltzmann partitioning*, consisting of finitely many volume elements from an accessible phase space. (Boltzmann considered $\Delta v \rightarrow 0$ to be very small, to overcome the problem of infinite entropies given a continuous representation of the dimensions of the phase space.) At time τ, let the system be at a point $\mathbf{A} =< p_1, p_2, \ldots, p_{3N}, q_1, q_2, \ldots, q_{3N} >$ as before. A volume element Δv is *occupied* at time τ if and only if for $1 \leq i \leq N$, p_{3i-2}, p_{3i-1} and p_{3i} are respectively in the intervals $\hat{p}_{i'}$, $\hat{p}_{i''}$, and $\hat{p}_{i'''}$; similarly for elements q. The system is said to occupy this volume element Δv_i or equivalently, the system is said to be in a classical *microstate* μ_i at time τ if and only if it occupies Δv_i at τ. In the sequel, we will use the word microstate to denote a classical microstate. The set $\mathcal{S}$ of all microstates of the accessible phase space $\mathcal{P}$ is the set of *feasible* microstates.

Physical Fact 4.1: Given a system $\mathcal{T}$ with at least one element, the accessible phase space $\mathcal{P}$ has at least one feasible microstate.

Physical Fact 4.2: Given a system $\mathcal{T}$ with state $\mathcal{S}$, its $entropy$ is $\kappa \ln(|\mathcal{S}|)$, where κ is the Boltzmann constant, and $|\mathcal{S}|$ is the *multiplicity* of $\mathcal{T}$.

The above definition of entropy is based on the canonical representation of a thermodynamical system $\mathcal{T}$ in contact with a thermal reservoir of infinite capacity which is then approximable, at thermal equilibrium, by the multiplicity $W \approx W_{max}$ as $N \rightarrow \infty$, where W_{max} is the multiplicity of the "equilibrium macro state" (for $N = 1$, $W = W_{max}$); see Mihalas [34] for example. If $\mathcal{T}$ is in isolation, then there can be no energy exchange possible between the system and the rest of the universe in which case W_{max} is the exact multiplicity of $\mathcal{T}$, determined by the energy of $\mathcal{T}$ at the instant of time when it is first isolated from $\mathcal{U}$. The reader can proceed to follow the developments in this paper without delving into this detail.

A *universe* of thermodynamic systems or simply a universe $\mathcal{U}$ is defined to be a set of M thermodynamic systems $\mathcal{T}_i \in \mathcal{U}, 1 \leq i \leq M$. The laws of thermodynamics have been used to characterize the interactions of various attributes between the elements of the universe. This characterization has traditionally been rooted in the operational model of Carnot based on an idealized heat engine, with a view of determining its efficiency. To model the thermodynamic characteristics of computing, abstractly presented as switching in Section V, we will now provide a more convenient transformational interpretation of the laws of thermodynamics below.

C. A Transformational View of Classical Thermodynamics

Let $\mathbf{T} = \mathbb{R}^+ \cup 0$ denote time and $\tau \in \mathbf{T}$ and $\hat{\tau}$ denotes an (open) interval $(\tau, \tau') \in \mathbf{T}$. Also q is reused to mean heat hereafter. The universe $\mathcal{U}$ is associated with five *state vectors* $\mathbf{w} =< w_1, w_2, w_3, \ldots, w_M >$, $\mathbf{q} =< q_1, q_2, q_3, \ldots, q_M >$, $\mathbf{u} =< u_1, u_2, u_3, \ldots, u_M >$, $\mathbf{s} =< s_1, s_2, s_3, \ldots, s_M >$ and $\mathbf{temp} =< t_1, t_2, t_3, \ldots, t_M >$, such that w_i, q_i, u_i, s_i, t_i are identified with system $\mathcal{T}_i \in \mathcal{U}$. A *thermodynamic process* or transformation is represented by a function TRANSFORMATION : $(\mathcal{U} \times \mathbf{T}) \rightarrow < W \times Q \times U \times S \times T \times \mathbf{T} >$, where, W, Q, U, S, T are the set of all state vectors, respectively of **w**, **q**, **u**, **s** and **temp**. Let TRANSFORMATION$(\mathcal{T}_i, \tau_i) = (\hat{\mathbf{w}}_i, \hat{\mathbf{q}}_i, \hat{\mathbf{u}}_i, \hat{\mathbf{s}}_i, \hat{\mathbf{temp}}_i, \tau_f)$. A transformation takes each element of the universe at some state at time instant τ_i, alters its entropy and temperature for some work performed and heat transferred by time τ_f; this transformation occurs in the interval $\hat{\tau} = (\tau_i, \tau_f)$. The variable $\hat{w}_i$ corresponds to the *work performed* on (or by) $\mathcal{T}_i$ during $\hat{\tau}$. In the domain of gases, the work performed results in a change to its *volume* V due to a change in the applied *pressure* P, measured in pascals (the volume V, the pressure P and the temperature T are referred to as *properties* or *observable properties*). In an idealized gas, $\hat{w}_i = \int_{V_i}^{V_f} P_i dV_i$, where V_i and V_f represent the volume of the gas at times τ_i and τ_f respectively. The variable $\hat{q}_i$ corresponds to the *heat transferred* into or out of the system. To be technically precise, in classical thermodynamics, heat flow is determined as the line integral $\oint \delta q$. For our purposes, it suffices to know that $\hat{q}$ represents this quantity over the interval $\hat{\tau}$. The variable $\hat{u}_i$ denotes the energy of $\mathcal{T}_i$. Similarly $\hat{s}_i$ is the entropy of $\mathcal{T}_i$, and t_i is its temperature (noting that the temperature t_i is proportional to the energy u_i).

We will now use the laws of thermodynamics and derive a set of "physically feasible" transformations. Any transformation is said to be *physically feasible* whenever the following conditions hold for $1 \leq i \leq M$:

$$u_i, \hat{u}_i \geq 0 \tag{1}$$

$$s_i, \hat{s}_i \geq 0 \tag{2}$$

$$\tau_f > \tau_i \tag{3}$$

$$\hat{u}_i - u_i = \hat{w}_i + \hat{q}_i \tag{4}$$

$$\sum_{i=1}^{M} \hat{s}_i - s_i \geq 0 \tag{5}$$

The interpretation of constraint(3) which states that $\tau_f > \tau_i$ is that time flows forward and is totally ordered. Hawkins [35], [36] and others [37] discuss this issue and its physical validity. Constraint(4) is referred to as the *first law* of thermodynamics. Similarly, constraint(5) is the *second law* of thermodynamics.

Any physically feasible transformation is defined to be *dissipationless* if $\tau_f - \tau_i \geq R$ where R is known as the *relaxation time*, and is a *fixed* attribute of the particular physical system. If the duration of any thermodynamic transformation is at least as large as the relaxation time, the system is always in an *equilibrium state*. Such a transformation is referred to as *quasistatic* transformation. Our models for computing and switching involve changing the state of a system $\mathcal{T}$ from one equilibrium state to another. Thus throughout, the constraint $\tau_f - \tau_i \geq R$ will be true.

D. Quantifying the Energy Behavior of Feasible Transformations

Let us consider a thermodynamic system $\mathcal{T}_i$ from the universe $\mathcal{U}$. We will now quantify the energy changes associated with $\mathcal{T}_i$ affected by specific transformations. A transformation *decreases* a thermodynamic attribute of $\mathcal{T}_i$, say its entropy s_i, if and only if $\hat{s}_i < s_i$. The decrease or increase of other attributes can be similarly represented. To continue, we adopt terminology dating back to Carnot [14], albeit in transformational terms. Let us fix a thermodynamic system $\mathcal{T}_i$ and continue to characterize attributes of transformations with respect to $\mathcal{T}_i$. A transformation is *isothermal* if and only if $t_{i'} = t_{j'}$ for any $\tau_i \leq \tau_{i'} \leq \tau_{j'} \leq \tau_f$. A transformation is *adiabatic* if and only if $\hat{q}_i = 0$. The following useful properties of these transformations will be used to deduce associated energy changes. We omit the units such as Joules and Kelvin whenever convenient.

Lemma 4.1: In any isothermal transformation $\hat{u}_i = u_i$

Proof: This follows from the definition of temperature $t = c \cdot u$ for some positive constant c. ■

Theorem 4.3: In any isothermal transformation, $-\hat{w}_i \leq t_i(\hat{s}_i - s_i)$ where t_i is the constant temperature of the isothermal system $\mathcal{T}_i$.

Proof: Following Clausius' definition of entropy which is identified at equilibrium with the Boltzmann form of entropy that we use in this paper, and the Second Law of thermodynamics (5), we have the Clausius inequality

$$\frac{\delta q}{t} \leq ds \tag{6}$$

Also, for any isothermal transformation, from (4) we have

$$\delta q + \delta w = 0 \tag{7}$$

and using the definition $\delta w = PdV$ in (7),

$$\delta q = -PdV \tag{8}$$

and therefore from (6),

$$-\int_{V_i}^{\hat{V}_i} PdV \leq \int_{s_i}^{\hat{s}_i} tds \tag{9}$$

where V_i and $\hat{V}_i$, s_i and $\hat{s}_i$ are respectively the volumes and entropies of the system at time instants τ_i and τ_f. ■

Theorem 4.4: In any adiabatic transformation, $\hat{s}_i - s_i \geq 0$

Proof: Since by definition, $\frac{\delta q}{t} \leq ds$, and in any adiabatic transformation, $\delta q = 0$, $ds \geq 0$. ■

We are now ready to state our main theorem. Let $\mathcal{S}$ and $\hat{\mathcal{S}}$ denote the states of $\mathcal{T}$ before and after a transformation.

Theorem 4.5: In any isothermal transformation with initial and final states $\mathcal{S}$ and $\hat{\mathcal{S}}$ respectively, where $|\hat{\mathcal{S}}| = \epsilon \cdot |\mathcal{S}|$ for $0 < \epsilon \leq 1$, $w_i \geq -\kappa t \ln \epsilon$ Joules.

Proof: From Physical Fact 4.2, it follows that

$$\kappa t[\ln(\epsilon \cdot |\mathcal{S}|) - \ln(|\mathcal{S}|)] \geq -\hat{w}_i \tag{10}$$

$$\kappa t \ln(\epsilon) \geq -\hat{w}_i \tag{11}$$

or equivalently

$$\hat{w}_i \geq -\kappa t \ln(\epsilon) \tag{12}$$

concluding the proof. ■

Theorem 4.5 characterizes the fact that in any isothermal transformation, any change in entropy of $\mathcal{T}_i$ is necessarily accompanied by the consumption of mechanical energy from the *rest of the universe*.

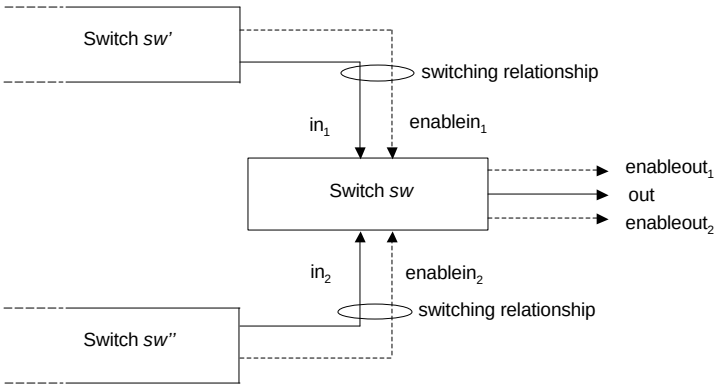


Fig. 1. The inputs to a switch

E. Sequences of Transformations

We now consider a finite sequence of transformations TRANS applied to specific thermodynamic system $\mathcal{T}_i \in \mathcal{U}$. Let $\text{TRANS}_K = <\text{TRANSFORMATION}_1 \circ \text{TRANSFORMATION}_2 \circ \ldots, \circ \text{TRANSFORMATION}_K>$, where each transformation is either isothermal or adiabatic, and let the symbol $\circ$ denote the composition of a pair of transformations. We can now claim the following general theorem about the energy consumed by a sequence of transformations. The *mechanical energy consumed* by a $\text{TRANSFORMATION}_j \in \text{TRANS}_K$ for $1 \leq j \leq K$ is

$$\mathbf{w_j} = \begin{cases} \hat{w}_j & \text{whenever } \hat{w}_j > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Theorem 4.6: Given any sequence TRANS_K, $\mathbf{W} = \sum_{1 \leq j \leq K} \mathbf{w_j} \geq -\kappa t \ln(\mathbf{e})$, where $\mathbf{e} = \prod_{1 \leq j \leq K} \epsilon_j$, and $\epsilon_j = min(\frac{|\hat{\mathcal{S}}_j|}{|\mathcal{S}_j|}, 1)$, the ratio of multiplicities associated with $\text{TRANSFORMATION}_j \in \text{TRANS}_K$.

Proof: If $\mathbf{e} = 1$, we are done. Now let $\mathbf{e} < 1$ and consider TRANS_K. Use ${\text{TRANS}_K}^-$ to denote the set of transformations such that $\text{TRANSFORMATION}_{j'} \in {\text{TRANS}_K}^-$ if and only if $\epsilon_{j'} < 1$. From the definition, $\mathbf{e}$ is the product of values $\epsilon_{j'}$ associated with all $\text{TRANSFORMATION}_{j'} \in {\text{TRANS}_K}^-$. Let $\mathbf{W}^-$ equal the sum of $\hat{w}_{j'}$ associated with all $\text{TRANSFORMATION}_{j'} \in {\text{TRANS}_K}^-$. From Theorem 4.4, every $\text{TRANSFORMATION}_{j'} \in {\text{TRANS}_K}^-$ is isothermal. Therefore, from the definition of $\mathbf{W}$ and Theorem 4.5, $\mathbf{W} \geq \mathbf{W}^- \geq -\kappa t \ln(\mathbf{e})$. ■

To interpret this theorem consider a sequence of λ transformations. Some of them are isothermal whereas the others are adiabatic. Now, the change in entropy is derived entirely from the isothermal transformations. Therefore, the condition stated in Theorem 4.3 must be true. Intuitively the goal of Theorem 4.6 is to help characterize the associated energy change in the context where the energy once expended to compute cannot be recovered, which is the term $\mathbf{W}$. Thus, all the results in this paper corresponds to the class of non-recovering computations based on computing devices in use today, wherein energy is consumed by a computer, but not recovered from the computer back to the energy source. The theorem above states that given a sequence of thermodynamic processes achieving a reduction of multiplicity such that $\mathbf{e} < 1$, the mechanical energy consumed is at least $\kappa t \ln(\mathbf{e})$ Joules.

F. Statistics of a phase space

The following physical fact will be very useful for characterizing the statistical behavior of microstates. Let $Pr[\mu]$ be the probability that a microstate μ exists at time τ.

Physical Fact 4.7: In any phase space of an ideal monoatomic gas, $Pr[\mu_i] = Pr[\mu_j]$ for any pair of microstates μ_i and μ_j from $\mathcal{P}$ at any time instant τ.

V. A SWITCH AND SWITCHING

In Section V-A below, we will define the basic elements of a single switch and its behavior. Following this, the notion of switching will be introduced in Section V-B. This definition will be entirely in "logical" terms without recourse to the physical interpretation, the latter being the topic of Section VI. This formalism is reminiscent of earlier developments in the classical switching theory field described in Kohavi [38].

Input	output
0	0
1	1

Identity Function

Input	output
0	1
1	0

Complement Function

Input	output
0	0
1	0

Constant Function

Input	output
0	1
1	1

Constant Function

Fig. 2. The four choices for switching functions of which sw realizes one choice

Signal	Symbol
enablein	w
in	x
enableout	y
out	z

Fig. 3. Symbols for inputs and outputs of a switch

A. Defining a switch

As shown in Figure 1, each switch sw has (up to) two alternate choices for "input values" as well as "enabling signals". Each input value and enabling signal of sw is in turn the output of a distinct switch (from the set of all switches, **SW**), sw' and sw'' in the example. The outputs of switch sw' are identified with the input value in_1 and the input enabling signal $enablein_1$, whereas the outputs of sw'' are identified with in_2 and $enablein_2$. Any switch sw in turn has two possible (mutually exclusive) enabling signals as output denoted by $enableout_1$ and $enableout_2$, as well a single output value out.

During the entire lifetime of a switch sw, each of its enabling signal $enablein_i, i \in \{1, 2\}$, is "associated with" exactly one $in_j, j \in \{1, 2\}$. Subsequently, these associations will be formalized as "switching relationships" as shown in the figure, $enablein_1$ is associated with in_1 and similarly, $enablein_2$ is associated with in_2; all four possible associations are allowed. In any legal *switching* of sw, *exactly one* of its two enabling signals will be "active", indicated by associating with it the value 1. Thus, in this example, switch sw produces an output value as a function of in_1 whenever $enablein_1$ is active whereas it produces an output value as a function of in_2 whenever $enablein_2$ is active.

To further understand switching, let us suppose that $enablein_1 = 1$. In this case, the output value out is determined by some fixed function f of in_1. Recall from Figure 1 that the association or switching relationship between $enablein_1$ and in_1 implies that whenever $enablein_1 = 1$, out is determined by in_1. In this case, sw now *switches* and produces an output using f, where the function f is one of the four possible choices shown in Figure 2: *identity, complement* and the two *constant* functions.

For convenience, we will use the notation in Figure 3 to denote the various inputs and outputs. Formally, let $I_{sw,1}$ and $I_{sw,2}$ be the two *inputs* to sw wherein $I_{sw,i}$ is the ordered pair $< w, x >$, where $w \in \{0, 1\}$ is an input enabling signal and $x \in \{0, 1, \Gamma\}$ corresponds to the input value, with Γ denoting an undefined value. Thus, each switch sw is associated with a pair of *switching relationships* of the form $< w_{sw,i}, x_{sw,1} >$ and $< w_{sw,i'}, x_{sw,2} >$ where $i, i' \in \{1, 2\}$: intuitively, these ordered pairs capture the association between an input value and an enabling signal—as shown in the example in Figure 1, each of the two distinct enabling signals are in a switching relationship with each of the two inputs. To reiterate, when an input enabling signal, say $w_{sw,i} = 1$, the corresponding value $x_{sw,1}$ is used as the input.

The *outputs* of a switch sw are defined to be no more than two ordered pairs $O_{sw,1}$ and $O_{sw,2}$, where $O_{sw,j}$ is the pair $< y_{sw,j}, z_{sw} >$, $j \in \{1, 2\}$. By definition, while a switch can produce two mutually exclusive output enabling signals with different values as outputs to two possible successor switches, it must always have the same output value z to all successors. In what follows, the subscript sw of w, x, y, z will be omitted whenever the use of the symbols is unambiguous.

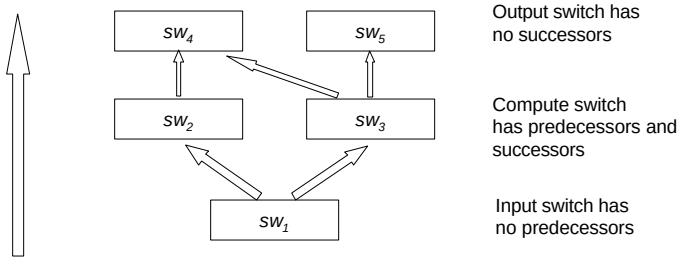


Fig. 4. Three types of switches

B. The process of switching

Recall from Figure 2 that a switch computes a fixed function f which is either the identity, complement or one of the two constant functions. Given a switch sw with an associated function f, a *switching* is defined as follows:

1) $z_{sw} = \Gamma$ and $y_{sw,1} = y_{sw,2} = 0$ whenever both its input enabling signals $w_{sw,1}$ and $w_{sw,2}$ have an identical value.
2) Whenever exactly one input enabling signal say $w_{sw,i} = 1$ and $x_{sw,j}$ is in a switching relationship$< w_{sw,i}, x_{sw,j} >$
   a) if $x_{sw,j} = 0$ then $z_{sw} = f(0)$, $y_{sw,1} = z_{sw}$ and $y_{sw,2} = \bar{z}_{sw}$
   b) if $x_{sw,j} = 1$ then $z_{sw} = f(1)$, $y_{sw,1} = z_{sw}$ and $y_{sw,2} = \bar{z}_{sw}$

Let $f(x) = z$ $(f(\bar{x}) = \bar{z})$ be the *deterministic* switching realized by sw as above. A *probabilistic switching* with a probability parameter $p \geq \frac{1}{2}$ is defined to be $f(x) = z$ with a probability p and $f(x) = \bar{z}$ with probability $(1 - p)$.

Consider a switch sw with its inputs defined by the switching relationships $< w_i, x_1 >$ and $< w_{i'}, x_2 >$ where $i, i' \in \{1, 2\}$. Consider another switch sw' with outputs y'_j, $y'_{j'}$, and z'. Switch sw' is *composed* to the switch sw, denoted by $sw' \prec sw$ if and only if for all $\tau > 0$ at least one of the following conditions are true:

1) one of the values x_1 or x_2 equals z'
2) one of the input enabling signals w_1 or w_2 of sw is identical to one of y'_j or $y'_{j'}$ the output enabling signals of sw'.

Whenever $sw' \prec sw$, switch sw' is said to *drive* switch sw, or equivalently sw is said to be *driven by* switch sw'. Let $sw' \prec sw$. Whenever $x_i = z'$, the ordered pair $< z', x_i >$ is referred to as a *wire*, and x_i is said to be *connected* to z' through the wire $< z', x_i >$. Similarly whenever $w_i = y'_j$, we again say that the wire $< y'_j, w_i >$ connects sw' to sw. Whenever a switch sw' drives sw, sw' is defined to be the *predecessor* of sw and sw is a *successor* of sw'. A switch sw is *strongly connected* if and only if both the elements of at least one of its input switching relationships are connected to a predecessor switch through wires.

To develop structures meant to realize entire computations, we will identify three types of switches: INPUT-SWITCH, OUTPUT-SWITCH and COMPUTE-SWITCH as shown in Figure 4. An INPUT-SWITCH sw_1 has no predecessors and drives at least one switch of type COMPUTE-SWITCH or OUTPUT-SWITCH. A switch such as sw_2 in our example, which is a COMPUTE-SWITCH, is driven by a switch sw_1 which is either an INPUT-SWITCH or a COMPUTE-SWITCH. In turn, COMPUTE-SWITCH sw_2 drives sw_4 which is an OUTPUT-SWITCH; an output switch has no successors and is driven by at least one switch of type INPUT-SWITCH or COMPUTE-SWITCH. Also, when convenient, we use the terminology input-switch, output-switch and compute-switch to refer to switches of type INPUT-SWITCH, OUTPUT-SWITCH and COMPUTE-SWITCH respectively.

VI. Modeling a switch physically

The essential steps of modeling a switch in a physical (thermodynamic) context are outlined in Figure 5. At the most basic level, the values 0 and 1, essential to switching, ought to be represented in a physically consistent and meaningful manner. This will be the topic of Section VI-A. A crucial aspect of this physical realization of a switch involves being able to detect the output z of a switch and to drive an enabling signal w from this value. As characterized in Section VI-B, this is accomplished through a *classical measurement*, in a manner consistent with the laws of thermodynamics. The classical physics constructs from Section IV are used throughout this section to achieve the physical modeling. *Given the goal of characterizing probabilistic switching, and its energy advantages, the contents of Sections VI-A and VI-B are inherently based on the statistical form of classical thermodynamics.* As shown in Figure 5, these models will then form a basis to physically realizing the mathematical notion of switching (from Section V) culminating in the calculations of energy changes associated with switching. These latter developments constitute the contents of Sections VII and VIII respectively.

Briefly, at a given instant of time τ, the current *physical state* (or state for short) of switch sw denoted by the symbols $\mathbf{S}_{sw} \subseteq \mathcal{P}_{sw}$, is the set of all microstates that are accessible or that can exist at τ where $\mathcal{P}_{sw}$ is its phase space. In the sequel and when it leads to unambiguous representation, we will omit the subscript sw and merely use the symbols $\mathbf{S}$ and $\mathcal{P}$ to

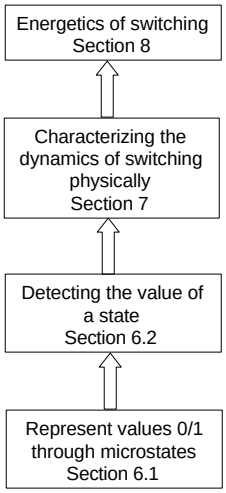


Fig. 5. Roadmap for the steps in realizing a switch and its switching behavior through a physical realization

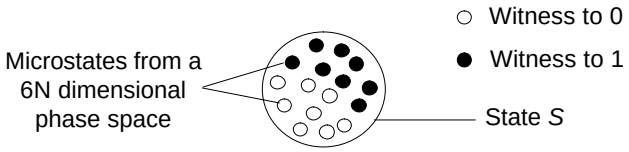


Fig. 6. The *current-state* and representing values through witnesses.

respectively denote the set of all legal states, and the phase-space of any switch $sw \in \mathbf{SW}$. Also, $\mathcal{S} \in \mathbf{S}$ will denote the current state at time τ, of a particular switch sw. Please refer to Section IV for some of the basic background definitions.

A. Representing 0 and 1 physically

A switch is a thermodynamic system with a *current state* $\mathcal{S}$ at time τ characterized by the set of all microstates that can exist in the physical system at that instant of time. Consider the set of all feasible states $\mathbf{S}$ of the switch sw. A distinguished state $\mathcal{S}_\Gamma \in \mathbf{S}$ of the switch denotes the state with maximum entropy across all feasible states. That is given a switch sw at any time $\tau \in \mathbf{T}$, there is no state $\mathcal{S}$ whose entropy is greater than the state $\mathcal{S}_\Gamma$.

Physical Fact 6.1: There exists a unique state $\mathcal{S}_\Gamma$ in the $\mathbf{S}_{sw}$ of a switch sw.

Consider a microstate $\mu_i \in \mathcal{S}_\Gamma$. This is defined to be a *witness* to a unique symbol 0 or 1. Thus $\mathcal{S}_\Gamma$ is partitioned into non-empty subsets $\mathcal{S}_0$ and $\mathcal{S}_1$ where a $\mu_i \in \mathcal{S}_j$ if and only if it is a witness to a value $j \in \{0, 1\}$. If the state of the system is $\mathcal{S}_0$, then we say that its value is 0, since the only witnesses correspond to the value 0; similarly with $\mathcal{S}_1$. In the case of a deterministic switch, if the state of the system is $\mathcal{S}_\Gamma$, by definition, we can say that it represents the undefined value Γ. At some time τ, when the physical realization of switch is in a particular state $\mathcal{S}_i$, its output value and enabling signal are determined through a specific measurement of $\mathcal{S}_i$.

Let us now use Figure 6 to clarify this concept. Suppose, the current state $\mathcal{S}$ of some switch sw in some (open) interval of time $\hat{\tau} = (\tau_1, \tau_2)$ consists of witnesses to 0 and 1 as shown in Figure 6. As stated earlier, each microstate is a witness to exactly one of the values 0 or 1. Its existence at time $\tau \in \hat{\tau}$, detected through a measurement, will determine the current values of the output z and the output enable signal y of switch sw. Formally, we use characteristic functions ξ_0 and ξ_1 to represent the existence of a witness (or microstate) at time τ:

$$\xi_0(\mathcal{S}, \tau) = \begin{cases} 1 & \text{if and only if there exists a microstate } \mu_i \in \mathcal{S}_0 \text{ at time } \tau \\ 0 & \text{otherwise} \end{cases}$$

and similarly,

$$\xi_1(\mathcal{S}, \tau) = \begin{cases} 1 & \text{if and only if there exists a microstate } \mu_i \in \mathcal{S}_1 \text{ at time } \tau \\ 0 & \text{otherwise} \end{cases}$$

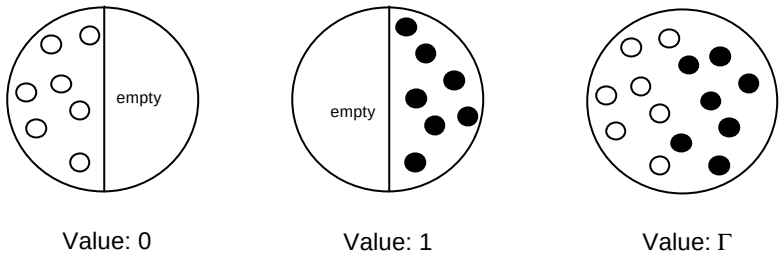


Fig. 7. Illustrating the central notion of value of a state modeled in the physical domain

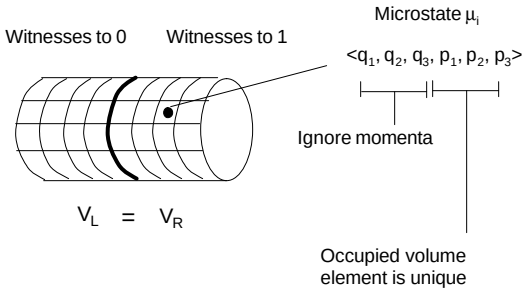


Fig. 8. Example of representing (deterministic) value as occupancy of a molecule in a cylinder of gas

A physical realization of the mathematical notion of switching in Section VII below will involve changing the current physical state of the switch. The act of switching involves the elimination of the "undesirable" witnesses from S. After the switching transformation, if $\mathcal{S}$ consists of witnesses to the value 0 only, the state of the "computing element" represented by $\mathcal{S}$ is determined to be 0. Similarly, if value consists of witnesses to 1 only, then the value is 1. However, if witnesses to both types of values exist, then, depending on the time τ at which the measurement is made, the value can be "erroneously" determined to be either 0 or 1, *thus laying the foundations for a physical interpretation to randomization.*

B. The value of a state and its properties

Using the above notions of a witness and its associated characteristic function, we are able to introduce the central definition of the *"value of a state"* $\mathcal{S}$ at time τ denoted by $\mathcal{V}$

$$\mathcal{V}(\mathcal{S},\tau) = \begin{cases} 1 & \text{iff } \xi_1(\mathcal{S},\tau) = 1 \\ 0 & \text{iff } \xi_1(\mathcal{S},\tau) = 0 \end{cases}$$

(We note that ξ_0 could have been used instead to define $\mathcal{V}$ as well.) We will now interpret this somewhat technical and *deterministic* definition and use the illustration in Figure 7 to help explain it. Let us suppose for purposes of illustration that $\mathcal{S}$ resides in a phase space $\mathcal{P}$. At any instant of time τ, a single microstate μ_i from $\mathcal{S}$ exists. Equivalently, at any instant of time τ, a single witness to a value—either 0 or 1—exists. Returning to Figure 7, the function $\mathcal{V}$, which models measurement, detects the microstate μ_i at time τ. Thus, the switch sw is in state $\mathcal{S}_1$, then the function $\mathcal{V}$ (or the associated detector) always yields a value of 1 and vice-versa.

C. An example of representing values through a physically plausible idealized gaseous system

The formal definition of a state and its value described above using microstates is physically meaningful in the context of a physical system that exhibits this structure. In particular, for this physical behaviour to be demonstrable constructively, the value determined by $\mathcal{V}$ ought to be detected, as mentioned above through a classical measurement. To help establish a "proof-of-concept" of the physical feasibility of switches as introduced here, as shown in Section VII, throughout this paper, we will use an idealized monoatomic gas as a basis for constructing switches and switching.

Briefly, let us consider a cylinder as shown in Figure 8 enclosing a single molecule of an idealized monoatomic gas, such as that studied by Szilard [8]. Let us suppose that occupancy of the left half of the cylinder by the molecule corresponds to the value $\mathcal{V}$ of 0, whereas it corresponds to a value $\mathcal{V}$ of 1 if the right half is occupied at a time instant τ. Now, as the molecule moves around in the cylinder (which it will whenever the temperature $t > 0\ K$), depending on its location at different instances of time, the value can be interpreted to be either 0 or 1. This statement is constructively true whenever a classical

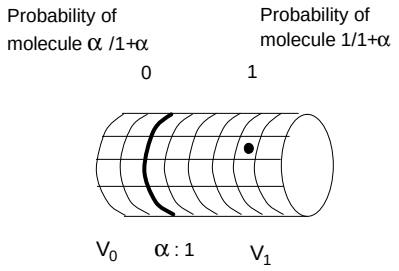


Fig. 9. Representing (probabilistic) value as occupancy of a molecule in a cylinder of gas

measurement can distinguish the position of the molecule between the two halves of the cylinder. Numerous simple detectors can be realized for this unimolecular case, which is one of the motivations for considering examples with a single molecule. Additionally, in the unimolecular case, it is an easy exercise to verify that any given microstate is *uniquely* associated with one of the two geometric halves V_L and V_R of the cylinder as shown in Figure 8. Equivalently, the partitioning of the cylinder into parts V_L and V_R also induces a partitioning of the set of all microstates in $\mathcal{S}_\Gamma$, where $\mathcal{S}_\Gamma$ is determined by the volume of the entire cylinder.

Thus, a physical volume element $< p_x, p_y, p_z >$ is uniquely associated with a microstate. The molecule's presence in a physical volume element in the left half of the cylinder V_L contribute to a value of 0 whereas its presence in a volume element in the right half V_R correspond to the value of 1. To detect the output value of the switch whose state is modeled through the thermodynamic system in Figure 8, a measurement to detect a microstate will involve detecting the presence of the molecule *instantaneously* at time τ, either in a volume element of V_L or in V_R; the former will imply an output value z of 0 whereas the latter will imply an output value of 1.

While it is important to consider the case of *explicit* switch construction using more than one molecule in general, (for reasons discussed in Section XV-B referring to the deeper issues connected to the ergodicity of a unimolecular system, and for the purpose of establishing the results in the sections to follow in contexts beyond switches constructed from a single molecule), for convenience of explanation, we will use examples that are unimolecular for the most part. A clarification of the nuances related to extending these "unimolecular switches" to those constructed using $N > 1$ molecules will be the subject of Section XII.

D. Defining value in the probabilistic case

Considering the probabilistic case, let us suppose that the current state $\mathcal{S} \equiv \mathcal{S}'_0 \cup \mathcal{S}'_1$ where $\mathcal{S}'_0$ and $\mathcal{S}'_1$ are non-empty subsets of $\mathcal{S}_0$ and $\mathcal{S}_1$ respectively. In the probabilistic case some of the witnesses to the value 0 remain in the current state $\mathcal{S}$, in addition to those for the value 1. The probability of detecting each individual value through a measurement depends on the relative proportions of witnesses or precisely, the cardinalities of $\mathcal{S}'_0$ and $\mathcal{S}'_1$.

Returning to the constructive example of a molecule of gas in a volume, a measurement at time τ will yield a value that is either 0 or 1. As shown in Figure 9, let the volume be altered such that the ratio of volumes is now $\alpha : 1$ as shown for $0 \leq \alpha \leq 1$. It is easy to see from this example that this corresponds to a partitioning of the microstates yielding the ratio of the associated set of witnesses to be $\frac{\mathcal{S}'_0}{\mathcal{S}'_1} = \alpha$.

More generally, consider a change to the state of a thermodynamic system representing a switch such that $\frac{\mathcal{S}'_0}{\mathcal{S}'_1} = \alpha$ from an initial state $\mathcal{S}$ such that $\frac{\mathcal{S}_0}{\mathcal{S}_1} = 1$. Now, following Physical Fact 4.7, a straightforward calculation yields the probability of a witness to the value 0 (or in the constructive example, a the molecule being detected in a volume element from V_0) is $\frac{\alpha}{1+\alpha}$ whereas that of finding a witness to the value 1 (or equivalently, the molecule in volume V_1) is $\frac{1}{1+\alpha}$.

Lemma 6.1: Let $\mathcal{S}'_0 = \alpha\mathcal{S}'_1$ for $0 \leq \alpha \leq 1$ and $\mathcal{S}' = \mathcal{S}'_0 \cup \mathcal{S}'_1$. Then, $\mathcal{V}(\mathcal{S}', \tau) = 0$ with probability $\frac{\alpha}{1+\alpha}$ and 1 with probability $\frac{1}{1+\alpha}$.

VII. THE PHYSICS OF SWITCHING

Computation proceeds by switching which involves transforming the current state of sw characterized by $\mathcal{S}$ into a state $\mathcal{S}'$. While several specific physical processes can be invoked to realize these transformations supporting switching, *canonically, these transformations will be characterized by the addition or subtraction of microstates to a current state $\mathcal{S}$* (at time $\tau \geq 0$) in such a way that the result is state $\mathcal{S}'$ at time $\tau' > \tau$. Informally, a single switching transformation takes the set of microstates that correspond to the current state $\mathcal{S}$ of sw and either adds or subtracts microstates from the set $\mathcal{S}_{IN}$ to derive $\mathcal{S}_{OUT}$, the output state. Let $\hat{\tau} = (\tau, \tau')$ be the interval in time during which this transformation is realized. As stated in Section IV,

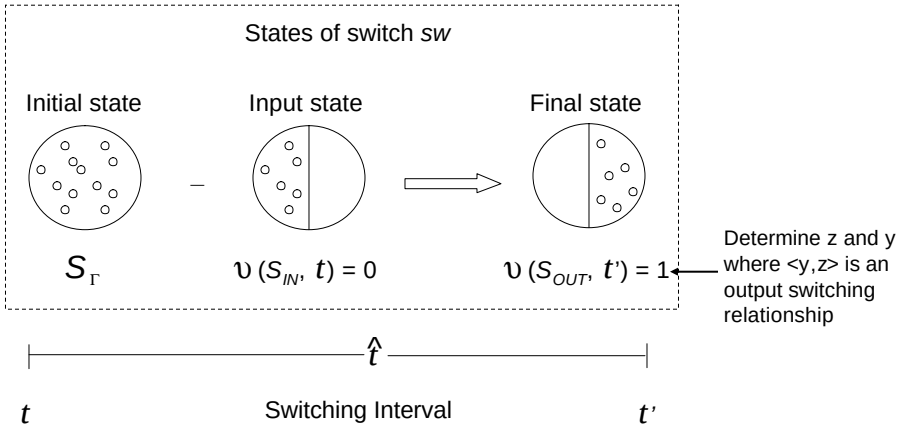


Fig. 10. An example switching function determines an output state $\mathcal{S}_{OUT}$ from $\mathcal{S}$ and $\mathcal{S}_{IN}$ using set difference

throughout this paper, we will be concerned with the case wherein the switching interval $\hat{\tau}$ corresponds to a change that is quasistatic, and is equal for all of the switches.

Let INT denote the set of all intervals of time. Using the terminology from Section VI, and relating the mathematical notion of switching from Section V, all the microstates in $\mathcal{S}_{IN}$ are witnesses to x whenever $w = 1$ and $< w, x >$ is an input switching relationship. Similarly, the microstates in $\mathcal{S}_{OUT}$ are witnesses to a value z and an output enabling signal y in a switching relationship with z. As shown in Figure 10, the output of sw is determined by the current state $\mathcal{S}$ and by the set of microstates that correspond to the input value selected by one of its (input) enabling signals; this selection is ignored in this example. It is easy to see that the function f in this example computes the logical complement of the input.

Formally, the process of *physical switching* (or when there is no ambiguity with its mathematical counterpart, simply switching) is realized as a function *associated with* switch sw and is defined to be SWITCHING : $\mathbf{S} \times \mathbf{S} \rightarrow \mathbf{S} \times INT$ where SWITCHING$(\mathcal{S}, \mathcal{S}_{IN}) = (\mathcal{S}', \hat{\tau})$ and $\mathcal{S}' = \mathcal{S} \otimes \mathcal{S}_{IN}$. For a given switch sw, $\otimes$ is fixed to be the set union ($\cup$) or set difference ($-$) operator from standard set theory. Thus, new states of a switch are derived from the current state $\mathcal{S}$ either through its set-theoretic union or difference with the input value to sw, determined by detection of a witness microstate (as outlined in Section VI-A) in the state of the physical representation of sw' driving sw. For convenience, let us denote $\mathcal{S}_{IN}$ by the symbol $\mathcal{I}$ and $\mathcal{S}_{OUT}$ by the symbol $\mathcal{O}$. Of all points in time, let τ_f denote the smallest value when the switching operation is completed, namely there is no $\tau'' < \tau_f$ for which the current state of sw is $\mathcal{O}$ whereas it is $\mathcal{O}$ at τ_f. To reiterate, for a switch sw, $\hat{\tau} = (\tau_i, \tau_f)$ will be referred to as a switching interval.

To illustrate the function SWITCHING, in the deterministic case, consider the example shown in Figure 11. Let $< w_i, x_j >$ be the switching relationship of interest where w_i is determined by the wire driven by the predecessor switch sw'' whereas x_j is the value determined by the wire driven by sw'. During the switching interval $\hat{\tau} = (\tau, \tau')$, switch sw has $w_i = 1$, determined by the value of (current state) $\mathcal{S}''$ through $\mathcal{V}(\mathcal{S}'', \tau) = 1$ at the beginning of the switching interval. Also $\mathcal{V}(\mathcal{S}', \tau)$ yields the input value x_i. In our example, the current state $\mathcal{S}$ of sw' at time τ is identical to $\mathcal{S}'_0$ and therefore, $\mathcal{V}(\mathcal{S}', \tau) = x_j = 0$. Continuing, let switch sw realize the set difference function namely $\otimes = -$, and the current state of sw be $\mathcal{S}_\Gamma$. Then $\mathcal{S}_{OUT} = \mathcal{S}_\Gamma - \mathcal{S}_0 = \mathcal{S}_1$. Thus at time τ' which is the end of the switching interval $\hat{\tau}$, a measurement of $\mathcal{S}_{OUT}$ through $z = \mathcal{V}(\mathcal{S}_{OUT}, \tau')$ yields a value of 1. The switch sw shown in this example is one way of physically realizing *negation* or the *logical complement* operation.

A. Physical instantiation of deterministic and probabilistic switching

A SWITCHING is *deterministic* if and only if $\mathcal{O} \subseteq \mathcal{S}_0$ or $\mathcal{O} \subseteq \mathcal{S}_1$ whenever $\mathcal{I}_0 \subseteq \mathcal{S}_0$ or $\mathcal{I}_1 \subseteq \mathcal{S}_1$. Let $\mathcal{I}_0 \subseteq \mathcal{S}_0$ and $\mathcal{I}_1 \subseteq \mathcal{S}_1$ denote the two possible inputs representing values 0 and 1, whereas $\mathcal{O}'$ and $\mathcal{O}''$ denote the output states with $\mathcal{I}_0$ and $\mathcal{I}_1$ as the respective inputs. A deterministic switching is *non-oblivious* if and only if $\mathcal{O}' \cap \mathcal{O}'' = \phi$.

Now considering the probabilistic variant, as before, let $\mathcal{O}'$ and $\mathcal{O}''$ denote the output states with $\mathcal{I}_0$ and $\mathcal{I}_1$ as the inputs respectively. Let $\hat{\mathcal{S}}_0 \subseteq \mathcal{S}_\Gamma$ and $\hat{\mathcal{S}}_1 \subseteq \mathcal{S}_\Gamma$ denote the maximal set of witnesses respectively to values 0 and 1. Define

$$\begin{aligned} \omega'_0 &= \mathcal{O}' \cap \hat{\mathcal{S}}_0 \\ \omega'_1 &= \mathcal{O}' \cap \hat{\mathcal{S}}_1 \end{aligned}$$

(ω''_0 and ω''_1 can be defined similarly). A SWITCHING function is randomized with probability parameter $p \geq \frac{1}{2}$ if and only if

1) given input $\mathcal{I}_0$, $\frac{|\omega'_0|}{|\omega'_1|+|\omega'_0|} = p$ whereas with input $\mathcal{I}_1$ $\frac{|\omega''_1|}{|\omega''_1|+|\omega''_0|} = p$ whenever f is an identity function whereas

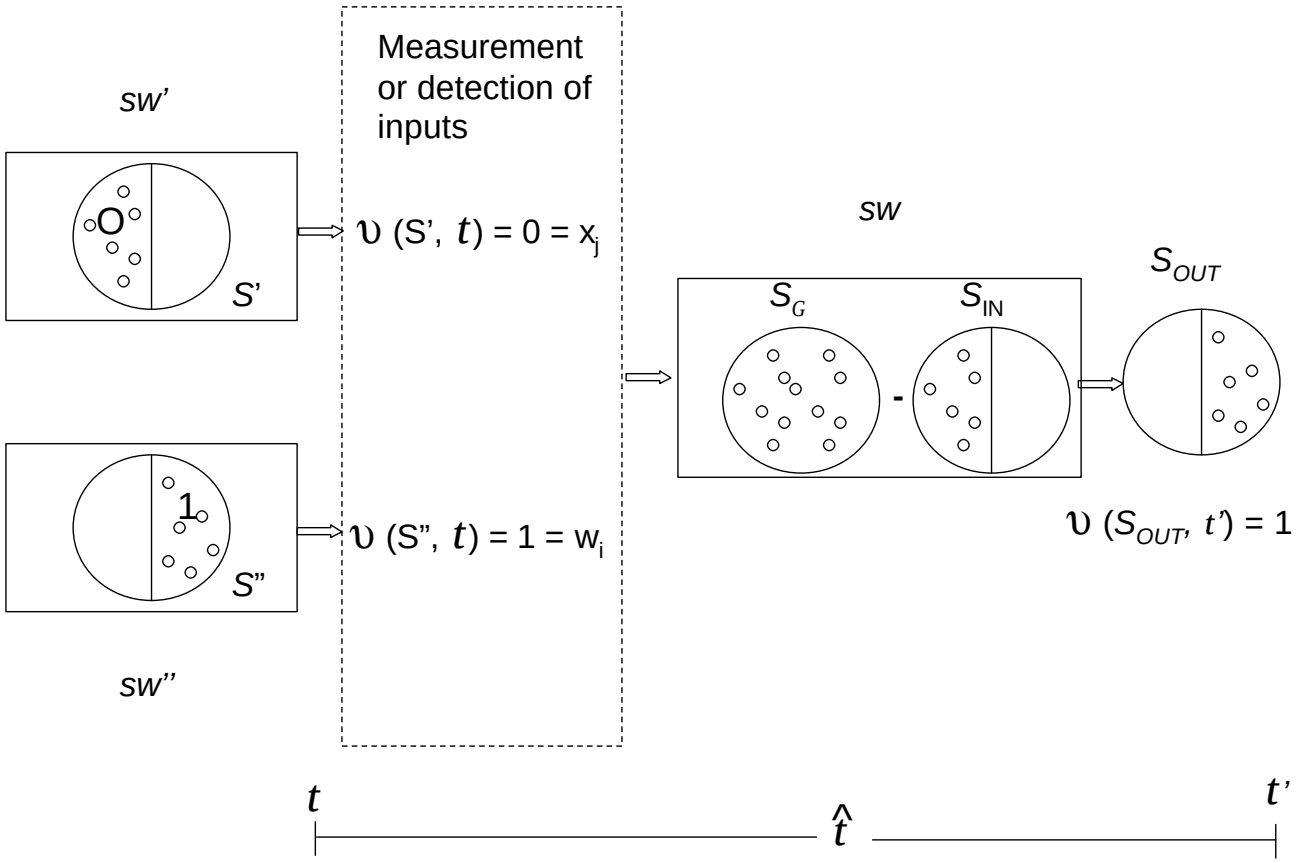


Fig. 11. Switching of sw to realize negation

2) given input $\mathcal{I}_0$, $\frac{|\omega_1'|}{|\omega_1'|+|\omega_0'|} = p$ whereas with input $\mathcal{I}_1$, $\frac{|\omega_0''|}{|\omega_1''|+|\omega_0''|} = p$ whenever f is the complement function.

In what follows, unless otherwise specified, whenever a switch sw is randomized with probability parameter p, we mean one that has the above property. The following observation summarizes the probabilistic aspects of SWITCHING succinctly.

Observation 7.1: Let $x \in \{0, 1\}$ be the output of a SWITCHING function f that is randomized with a probability $p < 1$. In any probabilistic (physical) SWITCHING implementing f, $\mathcal{V}(\mathcal{O}, \tau') = x$ with probability p and $\bar{x}$ with probability $(1 - p)$, whereas in deterministic SWITCHING, $p = 1$.

It is easy to verify that the mathematical definition of a randomized switching function (from Section V-A) with probability parameter p corresponds to SWITCHING realized above physically, so long as the statistical mechanics of the switch as outlined in Physical Fact 4.7 are true.

B. An important digression to detecting the value of a state and the cost of a classical measurement

In the above definition, during a switching interval $\hat{\tau}$, a switch sw was enabled by its input signal and in turn produced an output enabling signal, as well as an output value based on the function $\mathcal{V}$; the particular issue of detecting this value in a switch sw' or sw'' and using it to drive another switch sw was not specified. To understand this issue better, it is convenient to consider the RABRAM model, introduced by this author [6]. This energy-aware model of computing is a variant of a random-access machine whose execution on each *atomic step* is determined by a sequence of *read*, *execute* and *write* sub-steps as shown in Figure 12.

Relating this notion of a step of a RABRAM with a single transition of a switch, namely SWITCHING, the *read* corresponds to determining the value of $\mathcal{V}$ from the current state of sw', which drives switch sw. In our current context, a physical instantiation of *read* will be realized through a *classical measurement*. Considering switch sw', any such measurement detects witnesses from $\mathcal{S}_0'$ and $\mathcal{S}_1'$ from $\mathcal{S}'$ the state of sw' at time τ. (More generally, it can be coarser and be based on detecting groups of witnesses as opposed to single witnesses, so long as the constraints on the energy and the probability of the existence of the group of witnesses are consistent with the corresponding properties of single witnesses.) Continuing with the execution of a RABRAM step, by contrast with the *read* sub-step, an execute sub-step as well as the *write* sub-step correspond to a single switching step as developed in the previous sections, that changes $\mathcal{S}$ to $\mathcal{S}_{OUT}$ as shown in Figure 12.

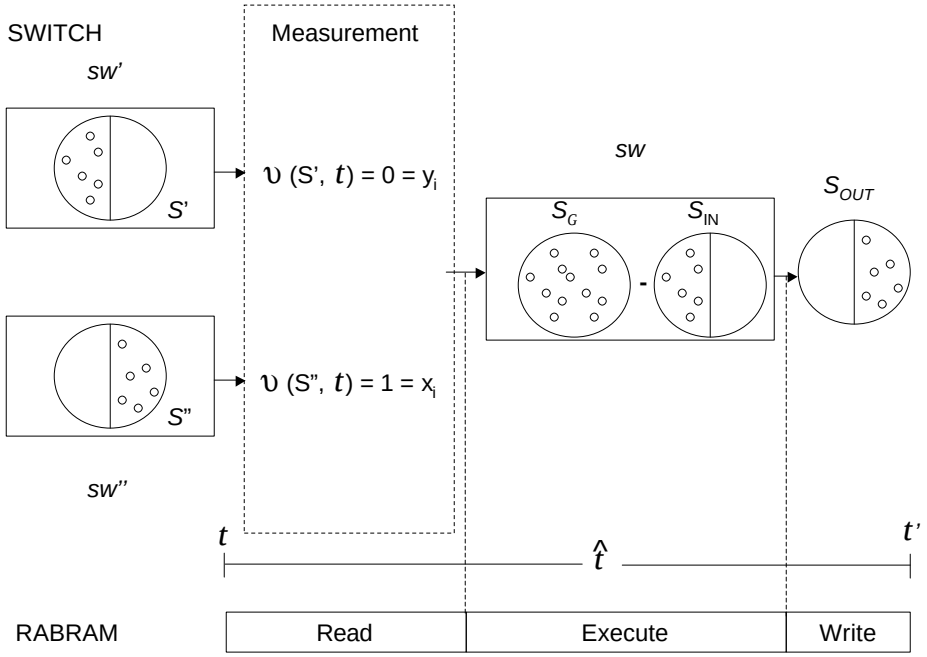


Fig. 12. The execution of a single atomic step of a RABRAM and its relationship to switching

Throughout, the amount of energy consumed by a (classical) measurement through which an *read* is realized is considered to be negligible, when compared with the energy cost of switching. We will now digress from the main theme of this work to discuss this issue of measurement. Landauer, in addition to taking a significant leap towards resolving the Maxwell's demon paradox [22], helped establish the cost of *inherent* energy consumption of a logically irreversible computational step by binding it to a macroscopic (non-statistical) analog of switching as characterized above (or in RABRAM terminology, with the execute and write sub-steps), as opposed to the measurement (or *read*) sub-step. For the reader familiar with the Maxwell's demon paradox, this implies associating the energy cost to the demon's act of "recording" the molecule's position and momentum thus *erasing* its current state of memory, as opposed to making the measurement itself. Several distinguished researchers including Szilard [8] and Brillouin [39] studied this issue without making this subtle yet crucial distinction. Further debate as well as clarification of the cost of measurement assumption continued (see Bennett and Landauer [40]), with Bennett [41] clarifying this issue conclusively.

Let us now interpret these comments from thermodynamics from the perspective of energy analysis within the switching and RABRAM models of interest to us. Using the RABRAM for convenience, (while noting the equivalence shown in Figure 12) the energy consumed by a single step of execution is lumped with the *execute* and *write* sub-steps whereas the *read* sub-step has negligible energy cost, which is ignored. With these clarifications, we will continue with a development of the framework as if a physical measurement process supporting the above conditions and the function $\mathcal{V}$ were to exist, and return to detailing the issue of demonstrating a plausible construction in in Section XII.

VIII. Energy characteristics of switching

Using the physical construction of a switch defined above and the SWITCHING function, we will now characterize the energy consumed in the switching, first in the deterministic case in section VIII-B followed by a consideration of the probabilistic case in Section VIII-C. Before doing this, in Section VIII-A below, we will rephrase Theorem 4.5 from Section IV-D using the terminology of a state $\mathcal{S}$, and of switching, developed in previous section.

A. Characterizing switching through change to state and related properties

Consider a switch $sw \in \mathbf{SW}$ with a particular current state $\mathcal{S}$, a particular input $\mathcal{I}$ and a particular output $\mathcal{O}$, and let $\hat{\tau}$ be the time interval for this switching. Given switch sw, let EF : $(\mathbf{S} \times \mathbf{S}) \rightarrow \mathbb{Q}^+$ be the *energy factor* representing the change in entropy of switch sw during an interval $\hat{\tau}$. Using terminology from Section IV, we recall that this is denoted by the term $\ln(\frac{|\mathcal{S}|}{|\mathcal{O}|})$. As usual, both the index of sw as well as $\hat{\tau}$ are implicit parameters. Recall from Section IV that the *energy consumed* $\mathbf{W}$ by a thermodynamic system $\mathcal{T}_i$ is related to the change in its entropy through Theorem 4.6.

Define the ENERGY : $(\mathbf{S} \times \mathbf{S}) \rightarrow \mathbf{T}$ of a switching to be the energy consumed by sw, identified with $\mathcal{T}_i$ equal to $\mathbf{W}$ Joules, where the *switching* of sw is realized using an *arbitrary sequence of thermodynamic transformations* defined in Section IV-E. Specifically, a single operation SWITCHING$(\mathcal{S}, \mathcal{I}) = (\mathcal{O}, \tau)$ is physically realized as a composition of thermodynamic

transformations TRANS of finite length K where the state at the beginning of (applying the first element of) TRANS is $\mathcal{S}$ and the state after applying TRANS is $\mathcal{O}$. Intuitively, a single atomic SWITCHING can be viewed as being "expanded" into a sequence of thermodynamic transformations of finite length, each of which has a current state determined by its predecessor in the sequence, and is either adiabatic or isothermal, as stated in Section IV-E. Then, the following convenient restatement of Theorem 4.6 is useful throughout the sequel.

Theorem 8.1: $\text{ENERGY}(\mathcal{S},\mathcal{I}) = -\kappa t \text{EF}(\mathcal{S},\mathcal{I})$ Joules, whenever $\frac{|\mathcal{S}|}{|\mathcal{O}|} > 1$ and is zero otherwise.

Proof: Restatement of Theorem 4.6 with $\text{ENERGY}(\mathcal{S},\mathcal{I}) = \mathbf{W}$. ■

The *switching energy* of a switch sw with initial state $\mathcal{S}$ is $max\{\text{ENERGY}(\mathcal{S},\mathcal{I}_0), \text{ENERGY}(\mathcal{S},\mathcal{I}_1)\}$.

B. Lowerbounds on deterministic switching

We consider the deterministic setting here, and state and prove the following theorem about the current state of a switch sw, leading to a lowerbound on the energy consumed by *any* deterministic switching in Theorem 8.3 below.

Lemma 8.1: In any non-oblivious deterministic switching, $\mathcal{O} \subset \mathcal{S}$.

Proof: Suppose not, then $\mathcal{S} - \mathcal{O} \neq \phi$. If $\otimes = -$ then we are done since $\mathcal{O} = \mathcal{S} - \mathcal{I}$. Now if $\otimes = \cup$, since SWITCHING is non-oblivious and deterministic, there exist non-empty sets $\mathcal{I}$ and $\bar{\mathcal{I}}$ where $\mathcal{I} \cap \bar{\mathcal{I}} = \phi$, $\text{SWITCHING}(\mathcal{S},\mathcal{I}) = \mathcal{O}$ and $\text{SWITCHING}(\mathcal{S},\bar{\mathcal{I}}) = \mathcal{O}'$, where

$$\begin{aligned} \mathcal{O} &= \mathcal{S} \cup \mathcal{I} \text{ implies } \mathcal{S} \subset \mathcal{O} \\ \text{and } \mathcal{O}' &= \mathcal{S} \cup \bar{\mathcal{I}} \text{ implies } \mathcal{S} \subset \mathcal{O}' \end{aligned}$$

Combined with the fact that by definition $\mathcal{O} \cap \mathcal{O}' = \phi$, it follows that $\mathcal{S} = \phi$, which contradicts Physical Fact 4.1. ■

It immediately follows from this proof that

Corollary 8.2: $\otimes = -$, in any deterministic non-oblivious switching.

Using these facts, we are now able to bound the energy change in any deterministic non-oblivious switching from below. As before, let $\mathcal{I}$ and $\bar{\mathcal{I}}$ denote the input sets corresponding to the two possible input values 0 and 1.

Theorem 8.3: In any non-oblivious switching, the switching energy of sw with any initial state $\geq -\kappa t \ln 2$ Joules and is a minimum when $|\mathcal{S}_0| = |\mathcal{S}_1|$.

Proof: That energy of $\text{SWITCHING}(sw) = -\kappa t \ln 2$ Joules when $|\mathcal{S}_0| = |\mathcal{S}_1|$ follows immediately from Theorem 8.1. Now, if the theorem is false, there exists a switch sw with state $\hat{\mathcal{S}}$ such that $|\hat{\mathcal{S}}_0| = \alpha \cdot |\hat{\mathcal{S}}|$ and $|\hat{\mathcal{S}}_1| = (1-\alpha) \cdot |\hat{\mathcal{S}}|$ for $\alpha \neq \frac{1}{2}$, such that its switching energy is less than $-\kappa t \ln 2$ Joules.

$\alpha < \frac{1}{2}$: From the definition of a deterministic non-oblivious switch, there exist inputs $\mathcal{I}$, $\bar{\mathcal{I}}$ to sw such that the respective outputs are $\mathcal{O}$ and $\mathcal{O}'$. Also from the definition of a deterministic switching, and Lemma 8.1 it follows that $\hat{\mathcal{S}}_0 \supset \mathcal{O}$ and $\hat{\mathcal{S}}_1 \supset \mathcal{O}'$. Now since $\alpha < \frac{1}{2}$ it follows from Theorem 8.1 and Corollary 8.2 that $\text{ENERGY}(\hat{\mathcal{S}},\mathcal{I}) = -\kappa t \ln \frac{|\hat{\mathcal{S}}|}{\alpha|\hat{\mathcal{S}}_0|} > -\kappa t \ln 2$ Joules, a contradiction.

$\alpha > \frac{1}{2}$: A similar argument completes the proof with $\hat{\mathcal{S}}_1$ replacing $\hat{\mathcal{S}}_0$. ■

C. Energetics of probabilistic switching

Given a probability parameter $p \geq \frac{1}{2}$, the following theorem characterizes the energy consumed by probabilistic switching.

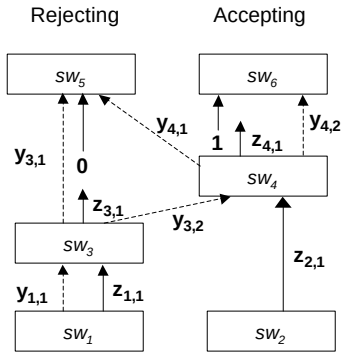
Theorem 8.4: Given any implementation of the function $\mathcal{V}$ and for any non-oblivious probabilistic switch with probability parameter p, $\text{ENERGY}(\mathcal{S},\mathcal{I})$ can be as low as $-\kappa t \ln 2p$ Joules. Therefore the switching energy of a probabilistic switch with initial state $\mathcal{S}$ and probability parameter p is $-\kappa t \ln 2p$ Joules.

Proof: (SKETCH): Given a realization of $\mathcal{V}$, that the switching energy of a non-oblivious probabilistic switch is $-\kappa t \ln 2p$ Joules is immediate from the definition of value of a state $\mathcal{V}(\mathcal{O},\tau')$, Observation 7.1 characterizing the probability of value detected by $\mathcal{V}$ and Theorem 8.1. The potential energy savings follow trivially from Theorem 8.3. ■

From this, we can immediately deduce

Corollary 8.5: The difference in switching energy between deterministic and probabilistic non-oblivious switches is $\kappa t \ln p$ Joules, which is a potential for saving of $\kappa t \ln \frac{1}{p}$ Joules.

1) Comments on the definition of switching: In all of the above physical characterizations of switching, the output state $\mathcal{O}$ is related to the current state $\mathcal{S}$ under the $\subseteq$ relationship. (Thus, either $\mathcal{O} \subseteq \mathcal{S}$ or vice-versa.) One can consider a definition of switching wherein $\mathcal{S}$ is incomparable to $\mathcal{O}$ for example, by the application of both the "-" as well as the "$\cup$" operators during a single switching step. This raises an interesting issue concerning the atomicity of a single switching operation wherein, multiple operations from set-algebra can be used to characterize a switching step. An implication of allowing these potentially more general definitions for switching, wherein the output and the current states are not necessarily restrictions of each other— contrary to the definitions used in this work— is that the set of witnesses in $\mathcal{O}$ need not necessarily be those in the current state $\mathcal{S}$. However, it is possible to extend the basic definitions as well as bounds on the energy derived above, to be valid in the context of a broad range of more general definitions of switching, so long as the witnesses in a given (non-empty) state



Switch	(Input) switching relationship	Input	Output enabling signals	
sw_3	$< y_{1,1}, z_{1,1} >$	$z_{1,1}$	$y_{3,1} = z_{3,1}$	$y_{3,2} = \bar{z}_{3,1}$
		0	1	0
		1	0	1
sw_4	$< y_{3,2}, z_{2,1} >$	$z_{2,1}$	$y_{4,1} = z_{4,1}$	$y_{4,2} = \bar{z}_{4,1}$
		0	1	0
		1	0	1

(b)

Fig. 13. A deterministic 2-canonical network resolving the AND function where the COMPUTE-SWITCHes sw_3 and sw_4 drive the *accepting* switch sw_6 and *rejecting switch* sw_5

$\mathcal{S}$, $\mathcal{I}$ or $\mathcal{O}$ have a fixed meaning—any witness (or microstate) from any of these sets of a given switch sw are permanently associated with the value 0 or with the value 1 throughout the life-time of sw. Furthermore, $\mathcal{O}$ ought to be determined as a function of $\mathcal{S}$ and $\mathcal{I}$, within the context of non-recovering computations as characterized in Section IV. The details of such extensions are beyond the scope of this work.

IX. NETWORKS FOR DETERMINISTIC COMPUTATIONS

In this section, we will first introduce a *network* of switches for deterministic computation (Section IX-A) and define their energy complexity (in Section IX-C). In Section X, we will prove a non-trivial lowerbound on the energy consumed by any deterministic network that can compute the logical AND function.

A. Defining a Network

A *Network* of switches is a connected directed acyclic graph $\mathcal{N} = (\mathbf{SW}, \text{WIRES})$ such that the vertices are switches, the edges are wires and the switches that are of the type OUTPUT-SWITCH as well as those of type COMPUTE-SWITCH are all strongly connected. Each switch has no more than two predecessors and no more than two successors. This assumption of bounded (two) degree entails no loss of generality in terms of the energy complexity introduced in Section IX-C below. A switch is defined to be *extrinsic* if and only if at least one of its inputs is driven by an input switch. It is defined to be *intrinsic* otherwise. A network is said to be *k-canonical* for $k \geq 1$ whenever it has

1) exactly k input switches.
2) two output switches, and
3) exactly one COMPUTE-SWITCH with one of its input enabling signals $w_j \equiv y_k$ where y_k is the output enabling signal of an input switch. For convenience we will refer to this switch as the START switch.

From a computer science and automata theory perspective, when the network is used as a basis for (formal) language recognition, it is convenient to view the two output switches as being either an *accepting* switch or a *rejecting* switch. Also, given an input which is a binary string determined by the settings of its input switches at time $\tau = 0$, the START switch has input enabling signal $w_j = 1$ thus "triggering" the computation.

In Figure 13(a), a network computing the logical AND function is sketched. Switches sw_1, sw_2 are of type INPUT-SWITCH, sw_3, sw_4 are of type COMPUTE-SWITCH (implementing the complement function) whereas switches sw_5 and sw_6 are of type OUTPUT-SWITCH; sw_5 is the rejecting switch and sw_6 is the accepting switch. Every switch unless it is of the type INPUT-SWITCH is strongly connected, and the network is directed and acyclic. Furthermore this network is 2-canonical, since in addition to the above constraints, exactly one COMPUTE-SWITCH sw_3 has an input enabling signal $y_{1,1}$ from INPUT-SWITCH sw_1. In Figure 13(b), we show the crucial relationships between the input values to sw_3 and sw_4, and their output and enabling signals in a "truth-table-like" structure.

B. Execution of a network

A switch sw is said to have switched by time τ if and only if its switching interval (τ', τ'') is such that $\tau'' \leq \tau$. In what follows, the input switches are all assumed to have switched by time $\tau = 0$ so that their output values can be determined by the function $\mathcal{V}$. Continuing, at time $t = 0$, $y_{sw,1}$ is 1 since the input-switch sw_1 has switched by time $\tau = 0$. (While this implies possibly negative switching times, this is merely a technicality and can be easily changed.) Thus, $\tau = 0$ is interpreted to denote the time when the first compute-switch starts switching. In our example of Figure 13, this switch is sw_3 which is

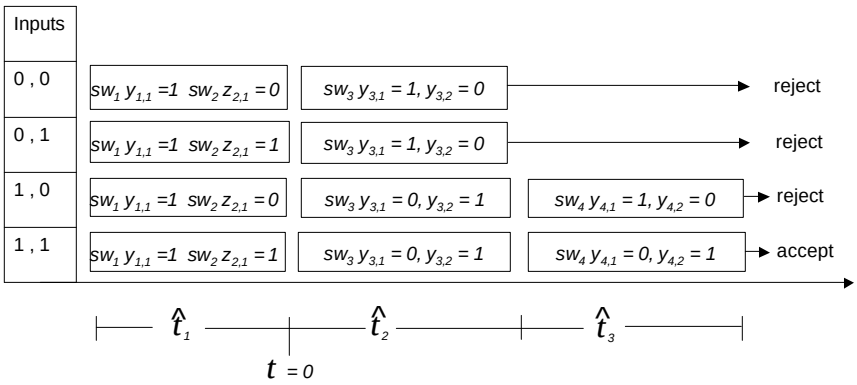


Fig. 14. A trace for the deterministic network resolving the AND function

enabled at $\tau = 0$. During the interval $\hat{\tau}_2$ as shown in Figure 14, it switches and the associated enabling signal values are shown in the figure. In the next interval $\hat{\tau}_3$, sw_4 might undergo a SWITCHING with the associated output and enabling signal values.

Let $\mathcal{N} = (\mathbf{SW}, \text{WIRES})$ be a k-canonical network. An *input binding* or input for short, is a function $I_{\mathcal{N},k} : \mathbf{SW}_{IN} \rightarrow \{0, 1\}$ where $\mathbf{SW}_{IN} \subseteq \mathbf{SW}$ is the set of all switches of type INPUT-SWITCH and $I_{\mathcal{N},k}(sw \in \mathbf{SW}_{IN})$ is the input value for switch sw, which by definition is either 0 or 1. An *execution* of $\mathcal{N}$ determined by or associated with input $I_{\mathcal{N},k}$ is a partial function $\mathcal{E} : \mathbf{SW} \rightarrow INT$, and $\mathcal{E}(sw) = \hat{\tau}$ is defined for some switch sw whenever

1) one of the input enabling signals w, of sw is 1 at time τ, where $\hat{\tau} = (\tau, \tau')$ and
2) τ is the smallest value for which this is true.

Let $\mathbf{SW}_{\mathcal{E}} \subseteq \mathbf{SW}$ be such that $sw \in \mathbf{SW}_{\mathcal{E}}$ if and only if $\mathcal{E}(sw)$ is defined. Given this fact, let $\hat{\tau}_f = (\tau_f, {\tau'}_f)$ be the final interval, that is the interval with the largest starting value of time $\tau = \tau_f$ in an execution of $\mathcal{N}$ (with input $I_{\mathcal{N},k}$) and let $\mathcal{E}(sw_f) = \hat{\tau}_f$.

Observation 9.1: sw_f is either an accepting or a rejecting switch.

Consider a finite sequence of intervals $\mathbf{t} =< \hat{\tau}_1, \hat{\tau}_2, \hat{\tau}_3, \cdots, \hat{\tau}_l >$ such that $\hat{\tau}_1 = (0, \tau' > 0)$ and for any $\hat{\tau}_i$ where $1 < i \leq l$, $\tau'_{i-1} = \tau_i$ where $\hat{\tau}_i = (\tau_i, \tau'_i)$. A sequence $\mathbf{t} =< \hat{\tau}_1, \hat{\tau}_2, \hat{\tau}_3, \cdots, \hat{\tau}_l >$ is said to be a *trace* of a network $\mathcal{N}$ induced by execution $\mathcal{E}$ with input $I_{\mathcal{N},k}$ whenever $\mathcal{E}$ is a bijection from $\mathbf{SW}_{\mathcal{E}}$ to $\mathbf{t}$. In Figure 14, the traces associated with inputs $< 0, 0 >, < 0, 1 >, < 1, 0 >$ and $< 1, 1 >$ to the network in Figure 13 are shown. Traces will be used to characterize the execution of a network, given an input. We note in passing that in the deterministic network, each input binding $I_{\mathcal{N},k}$ is associated with a unique execution $\mathcal{E}$.

C. Energy complexity of deterministic networks

We will now introduce the definition of the *energy complexity* of a network. An interval $\hat{\tau}$ is oblivious in a trace $\mathbf{t}$ whenever switch sw in this interval, determined by $\mathcal{E}^{-1}(\hat{\tau})$, is oblivious. It is non-oblivious otherwise. We define the *energy characteristic* of $\mathcal{N}$ in an execution represented by the function $EC : (\mathbf{SW} \times \mathbf{T}) \rightarrow \{0, 1\}$, where

$$EC(sw, \hat{\tau}) = \begin{cases} 1 & \text{whenever } sw \text{ is a COMPUTE-SWITCH and } \mathcal{E}(sw) = \hat{\tau} \text{ is a non-oblivious interval} \\ 0 & \text{otherwise} \end{cases}$$

Let $\mathbf{t} =< \hat{\tau}_1, \hat{\tau}_2, \hat{\tau}_3, \cdots, \hat{\tau}_l >$ be a trace induced by execution $\mathcal{E}$ of a network $\mathcal{N}$, corresponding to input $I_{\mathcal{N},k}$. Then the *effective energy* of $\mathbf{t}$ is

$$EE(\mathbf{t}) = \sum_{1 \leq j \leq l} \sum_{\forall sw \in \mathbf{SW}} EC(sw, \hat{\tau}_j)$$

The *deterministic energy complexity* $\mathbf{E}$ of a network with respect to a family of traces $\bar{T}$ induced by input bindings $\bar{I'}_{\mathcal{N},k}, \bar{I''}_{\mathcal{N},k}, \cdots$ is

$$\mathbf{E}(\mathcal{N}) = \max_{\forall \mathbf{t} \in \bar{T}} EE(\mathbf{t})$$

Consider a network $\mathcal{N}$ to be homogeneous if all of its switches of type COMPUTE-SWITCH have the same switching energy, say β. Then

Observation 9.2: There exists an input binding $I_{\mathcal{N},k}$ and execution $\mathcal{E}$ that induces a trace $\mathbf{t}$ such that, the energy expended by the switches of the type COMPUTE-SWITCH in $\mathcal{N}$ during the trace $\mathbf{t}$ is $\mathbf{E}(\mathcal{N}) \cdot \beta$ Joules.

Network Input <b_1, b_2 >	Bindings	
	sw_1	sw_2
0,0	0 (b_1)	0 (b_2)
0,1	0 (b_1)	1 (b_2)
1,0	1 (b_1)	0 (b_2)
1,1	1 (b_1)	1 (b_2)

(a)

Network Input <b_1, b_2 >	Bindings	
	sw_1	sw_2
0,0	0 (b_1)	0 (b_2)
0,1	0 (b_1)	1 (b_2)
1,0	0 (b_2)	1 (b_1)
1,1	1 (b_1)	1 (b_2)

(b)

Fig. 15. Consistent and inconsistent bindings

X. Lower bound for the energy complexity of an AND network

In the sections above, the notion of a network and the energy complexity of such a network have been defined. In this section, we will prove a non-trivial lower bound on the energy complexity of any network that resolves the logical AND function.

To generalize the previous example of a network realizing the AND boolean function, formally, we define $\mathbf{B}$ to be a set of k-vectors from $\{0,1\}^k$. Consider a fixed vector $\mathbf{b} \in \mathbf{B}$, such that $\mathbf{b} =< b_1, b_2, b_3, \cdots, b_k >$ where $b_i \in \{0,1\}$. Let $\mathcal{B}$ be a k-ary boolean function with $\mathbf{B}$ as its domain. Let $sw'_1, sw'_2, \cdots, sw'_k \in \mathbf{SW}_{IN}$ be the input switches of a k-canonical network $\mathcal{N}$. $I_{\mathcal{N},k}$ binds $\mathcal{N}$ to $\mathbf{b}$ if and only if $I_{\mathcal{N},k}(sw'_j) = b_{j'}$ for $1 \leq j, j' \leq k$, and it is an onto function. $\mathcal{N}$ *resolves* $\mathbf{b}$ *with respect to* $\mathcal{B}$ whenever, given a suitable $I_{\mathcal{N},k}$ and $\mathcal{E}$,sw_f in (the corresponding trace $\mathbf{t}$) is an accepting (rejecting) switch, $\mathcal{B}(\mathbf{b}) = 1$ (0 respectively). Consider a family of input bindings $\mathbf{I} = \bar{I}_{\mathcal{N},k}, \bar{I'}_{\mathcal{N},k}, \bar{I''}_{\mathcal{N},k}, \cdots$. A family of bindings $\mathbf{I}$ bind $\mathcal{N}$ to $\mathbf{B}$ if and only if given any $\mathbf{b} \in \mathbf{B}$, there exists a unique $I_{\mathcal{N},k} \in \mathbf{I}$ such that $I_{\mathcal{N},k}$ binds $\mathcal{N}$ to $\mathbf{b}$. A network $\mathcal{N}$ *resolves a boolean function* $\mathcal{B}$ with an input family of bindings $\mathbf{I}$ (that binds $\mathbf{B}$ to it) provided given any $\mathbf{b} \in \mathbf{B}$, $\mathcal{N}$ resolves $\mathbf{b}$ with respect to $\mathcal{B}$. We define $\mathbf{I}$ to be *consistent* if and only if given any member pair of bindings $\bar{I}'_{\mathcal{N},k}$ and $\bar{I}''_{\mathcal{N},k}$ and corresponding vectors, $\mathbf{b}$ and $\mathbf{b}'$ such that $\bar{I}'_{\mathcal{N},k}(sw) = b_j \in \mathbf{b}$, then $\bar{I}''_{\mathcal{N},k}(sw) = b'_j \in \mathbf{b}'$. Informally all bindings associate the same index from the input (assignment) $\mathbf{b}$ with the same INPUT-SWITCH switch sw. In what follows, we will only be concerned with families of consistent bindings $\mathbf{I}$. For example the bindings in Figure 15(a) are consistent whereas those in Figure 15(b) are not, since for example, switch sw_1 has b_1 and b_2 assigned to it.

Given a trace $\mathbf{t} =< \hat{\tau}_1, \hat{\tau}_2, \hat{\tau}_3, \cdots, \hat{\tau}_l >$, a maximally oblivious subsequence $< \hat{\tau}_j, \hat{\tau}_{j+1}, \cdots, \hat{\tau}_{j'} >$ is any subsequence of intervals such that every interval in the subsequence is oblivious and furthermore whenever $j > 0$ $\hat{\tau}_{j-1}$ and, whenever $j' < l$, $\hat{\tau}_{j'+1}$ are non-oblivious intervals. A boolean function $\mathcal{B}$ is non-trivial if and only if there exist $\mathbf{b}, \mathbf{b}' \in \mathbf{B}$ such that $\mathcal{B}(\mathbf{b}) \neq \mathcal{B}(\mathbf{b}')$.

Lemma 10.1: Given $\mathcal{N}$, let $\mathbf{t}$ be a trace and $\bar{\mathbf{t}} =< \hat{\tau}_j, \hat{\tau}_{j+1}, \cdots, \hat{\tau}_{j'} >$ where $j' = j + \lambda$ for $1 \leq j \leq j' \leq l$ be any maximally oblivious subsequence of $\mathbf{t}$. Then, $\mathcal{N}$ cannot resolve a non-trivial boolean function if $\lambda = l - 1$ for any trace $\mathbf{t}$. Also let $\mathcal{E}^{-1}(\hat{\tau}_j) = sw$ and $\mathcal{E}^{-1}(\hat{\tau}_{j'}) = sw'$ and z, z' be their respective outputs in $\mathbf{t}$. Whenever $\mathcal{E}^{-1}(\hat{\tau}_j) \equiv \mathcal{E}^{-1}(\hat{\tau'}_j) \equiv \mathcal{E}^{-1}(\hat{\tau''}_j) \cdots$ and z is constant in traces **t,t',t"**, $\cdots$ associated with all of the input bindings $\bar{I}_{\mathcal{N},k}, \bar{I'}_{\mathcal{N},k}, \bar{I''}_{\mathcal{N},k}, \cdots$ from $\mathbf{I}$, $\mathcal{E}^{-1}(\hat{\tau}_{j'}) \equiv \mathcal{E}^{-1}(\hat{\tau}'_{j'}) \equiv \mathcal{E}^{-1}(\hat{\tau}''_{j'}) \cdots$ and z' is a constant.

Proof: The first part of the claim is a straight forward induction on the length of the sequence and the definition of an oblivious function. The second part of the claim additionally follows the definitions of an input, and an output to a switch and the definition of switching. ■

Theorem 10.1: $\mathbf{E}(\mathcal{N}) \geq 2$ for any 2-canonical network $\mathcal{N}$ that resolves the AND boolean function.

Proof: (SKETCH) If the theorem is false for a network $\mathcal{N}$ then, in any trace $\mathbf{t}$ induced by an execution $\mathcal{E}$ with input binding $I_{\mathcal{N},k}$, there can be no more than one non-oblivious interval. Without loss of generality, let this interval be τ_j for $1 \leq j \leq l$ where $\mathbf{t} =< \tau_1, \tau_2, \tau_3, \cdots, \tau_l >$. If there is no such interval, $\mathcal{N}$ cannot resolve a non-trivial boolean function from Lemma 10.1. Let $\mathcal{E}^{-1}(\tau_j) = sw$. sw must be extrinsic or else, from Lemma 10.1, once again, $\mathcal{N}$ computes a trivial boolean function. Let INPUT-SWITCH sw' drive sw and consider inputs $\mathbf{b_1} \equiv< 1,1 >, \mathbf{b_2} \equiv< 0,1 >$ and $\mathbf{b_3} \equiv< 1,0 >$ to $\mathcal{B}$, which by hypothesis is an AND function. From pigeon-holing, there exist two inputs in any consistent family of input bindings, $\mathbf{b_1}$ and $\mathbf{b_2}$ without loss of generality, such that the input value to sw derived from sw', $x = 1$, and $\mathcal{B}(\mathbf{b_1}) \neq \mathcal{B}(\mathbf{b_2})$. Then, from Lemma 10.1 and the fact that $\mathcal{N}$ is 2-canonical, we know that in both cases the corresponding traces have the same accepting or rejecting switch sw, in their final interval, whereas $\mathcal{B}(\mathbf{b_1}) \neq \mathcal{B}(\mathbf{b_2})$. ■

From this theorem and Theorem 8.3, it follows that

Observation 10.2: There exist inputs to any 2-canonical network that resolves the AND boolean function such that the energy consumed is at least $-2\kappa t \ln 2$ Joules.

XI. NETWORKS FOR RANDOMIZED COMPUTATIONS

In Section IX above, the notion of a *deterministic* network that can resolve a boolean function was introduced, and its energy complexity defined. Implicit in the definition is the fact that every switch sw in network $\mathcal{N}$ is a deterministic switch. In Section V-B the notion of a probabilistic switch was introduced. We will now use this definition to construct probabilistic networks.

A. Probabilistic networks and their execution

A k-canonical probabilistic network $\mathcal{R}$ is any network with the property that a COMPUTE-SWITCH can either be deterministic or probabilistic. A *randomized execution* is $\hat{\mathcal{E}} : \mathbf{SW} \rightarrow \mathbf{T}$ as before, where the individual switches are randomized with some probability parameter p. Given a k-canonical probabilistic network R and an input binding $I_{\mathcal{R},k}$, by contrast with the deterministic case, we now have a family of *randomized executions* $\hat{\mathcal{E}}_1, \hat{\mathcal{E}}_2, \hat{\mathcal{E}}_3, \cdots, \hat{\mathcal{E}}_l$ each inducing traces $\rho_1, \rho_2, \rho_3, \cdots, \rho_l$, with respective probabilities $r_1, r_2, r_3, \cdots, r_l$. Let $\hat{\mathcal{E}}_j(sw) = \hat{\tau}$ as before and let sw' be the switch in interval $\hat{\tau}'$ immediately preceding $\hat{\tau}$ in ρ_i, that is $\hat{\tau}'$ and $\hat{\tau}$ are of the form (τ'', τ) and (τ, τ') respectively. Also, let x' be its input value at time τ''. (We recall that in any legal execution, the input value to a switch is defined, and one of its input enabling signals is asserted.) Now, the conditional probability q_i associated with $\hat{\tau}$, is the probability that the output enabling signal from sw' driving sw equals one at time τ, given the input to sw' is x'. Informally, each member of this family of traces is generated due to the probabilistic or randomized nature of the output of the probabilistic switches $\mathcal{R}$ in the network. This is in contrast to a deterministic network $\mathcal{N}$ which has a unique trace given an input binding $I_{\mathcal{N},k}$.

Observation 11.1: By definition, $r_i = \prod_{1 \leq j \leq l'} q_j$ where $\rho_i =< \hat{\tau}_1, \hat{\tau}_2, \hat{\tau}_3, \cdots, \hat{\tau}_{l'} >$. Furthermore, $\sum_{1 \leq j' \leq l} r_{j'} = 1$.

Given a single input, a probabilistic network can, depending on the execution, invoke different switches with varying probability parameters each leading to a distinct trace. Thus, whereas a fixed input is associated with a unique trace in the deterministic case, it is associated with a family of traces whose relative probabilities are as stated in Observation 11.1.

B. Energy complexity of probabilistic networks

We will now introduce the notion of the *energy complexity* of a probabilistic network. Let $\rho =< \hat{\tau}_1, \hat{\tau}_2, \hat{\tau}_3, \cdots, \hat{\tau}_{l'} >$ be a trace from the *family* of traces $\mathcal{F}$ induced by execution $\mathcal{E}$. Then the *expected effective energy* of $\mathcal{F}$ is

$$REE(\mathcal{F}) = \sum_{\rho_i \in \mathcal{F}} r_i \cdot EE(\rho_i)$$

where r_i is the probability of trace ρ_i. An example of an AND network will help illustrate this notion in Section XI-B.1 below. These definitions are an adaptation of the measure of *logical work* introduced by Palem [6] in the context of a RABRAM. The *randomized energy complexity* $\mathbf{RE}$ of a network is

$$\mathbf{RE}(\mathcal{R}) = \max_{\forall \mathcal{F} \in \mathbf{F}} REE(\mathcal{F})$$

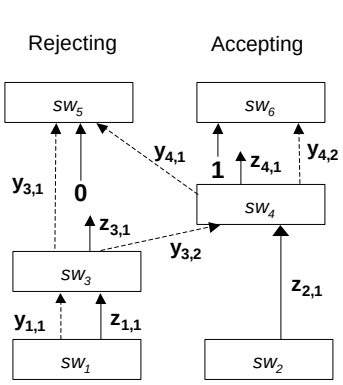
where $\mathbf{F}$ is the *set of all trace families* induced by executions associated with all the consistent input bindings $\bar{I}_{\mathcal{R},k}, \bar{I'}_{\mathcal{R},k}, \bar{I''}_{\mathcal{R},k}, \cdots$. Let a probabilistic network $\mathcal{R}$ be r-homogeneous if and only if it is homogeneous and all switches of the type COMPUTE-SWITCH are randomized. Then,

Observation 11.2: There exists a binding $I_{\mathcal{R},k}$ such that the energy in Joules consumed by the COMPUTE- SWITCHes of $\mathcal{R}$, averaged over all the traces induced by associated executions $\hat{\mathcal{E}}_1, \hat{\mathcal{E}}_2, \cdots$, is $\mathbf{RE}(\mathcal{R}) \cdot \text{ENERGY}(sw)$, where sw is any probabilistic COMPUTE-SWITCH in $\mathcal{R}$.

1) Upperbound on the energy complexity of the AND network: In Figure 16, we show a probabilistic network $\mathcal{R}$ that can resolve the logical AND function. The details of its construction are identical to those in the deterministic case from Figure 13. Switches sw_3 and sw_4 are of type COMPUTE-SWITCH and are randomized, with a probability parameter p, they compute complement function. The output signals of these switches as a function of the inputs are shown in the table in Figure 16. Recall from Theorem 8.4 that in any switching, a probabilistic switch with probability parameter p consumes $-\kappa t \ln 2p$ Joules. Thus, using this fact and $\mathbf{RE} = (1 + p)$, from Observation 11.2, we have the expected energy consumed by switches of type COMPUTE-SWITCH in the example probabilistic network to be $(1 + p)\kappa t \ln(2p)$ Joules. This is less than its deterministic counterpart, which is a minimum of $2\kappa t \ln(2)$ Joules, following the lowerbound established in Theorem 10.1 and Observation 10.2.

XII. A PLAUSIBLE CONSTRUCTION OF A SWITCH AND SWITCHING

A *mechanical* description of a switch sw as well as its realization in a manner consistent with the principles of statistical thermodynamics was presented in Section VI-A and Section VII above. We will now provide an interpretation of the abstract switch construction (from Section VII) through a mechanical-pneumatic device which serves as a plausible physical realization of the switch. For convenience of explanation and to establish the physical validity of the theoretical framework and results derived in the previous sections, following Szilard [8], we will first consider a cylinder with a single molecule in it.



Switch	(Input) switching relationship	Input	Output enabling signals	
SW_3	$< y_{1,1}, z_{1,1} >$	$z_{1,1}$	$y_{3,1}$	$y_{3,2}$
		0	$1 = z_{3,1}$ (p) $0 = \bar{z}_{3,1}$ (1-p)	$0 = \bar{z}_{3,1}$ (p) $1 = z_{3,1}$ (1-p)
		1	$0 = z_{3,1}$ (p) $1 = \bar{z}_{3,1}$ (1-p)	$1 = \bar{z}_{3,1}$ (p) $0 = z_{3,1}$ (1-p)
SW_4	$< y_{3,2}, z_{2,1} >$	$z_{2,1}$	$y_{4,1}$	$y_{4,2}$
		0	$1 = z_{4,1}$ (p) $0 = \bar{z}_{4,1}$ (1-p)	$0 = \bar{z}_{4,1}$ (p) $1 = z_{4,1}$ (1-p)
		1	$0 = z_{4,1}$ (p) $1 = \bar{z}_{4,1}$ (1-p)	$1 = \bar{z}_{4,1}$ (p) $0 = z_{4,1}$ (1-p)

Fig. 16. A probabilistic 2-canonical network resolving the AND function

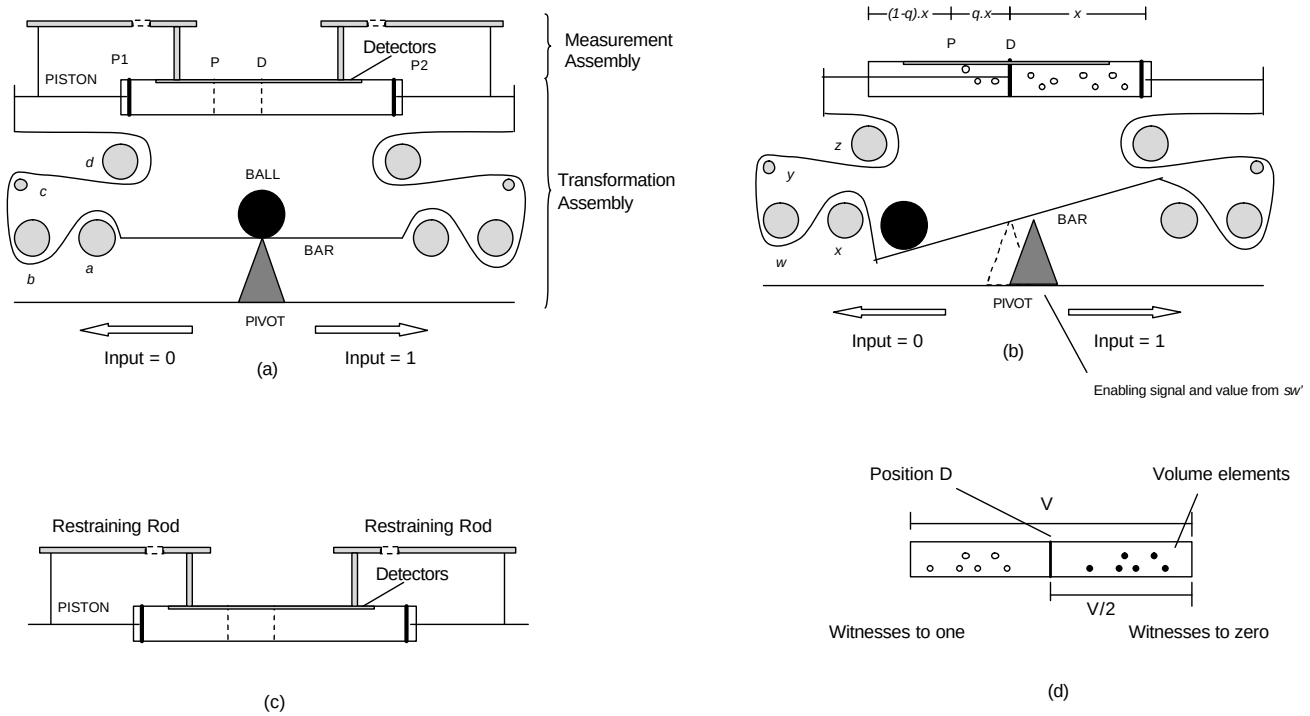


Fig. 17. A plausible pneumatic device that realizes switching

A. Device Construction

As shown in Figure 17(a), the device consists of two *assemblies*, denoted as the *measurement assembly* and the *transformation assembly*. Recall (from Section VII-B) that a switching operation consists of read step followed by an execute-write step in that order. The measurement assembly performs the physical equivalent of a "read" whereas the transformation assembly is equivalent to an integrated "execute-write" portion of the execution of a switch. In the sequel, we will describe these two assemblies and their interactions to realize switching, starting with the transformation assembly. The nature of this description will be analogous to a thought experiment in physics.

B. The transformation assembly

The transformation assembly of a switch sw consists of a rigid massless BAR, balanced over a massless PIVOT. In practice, these could be physical components of negligible mass. The PIVOT rests on a horizontal frictionless surface and is free to move as shown in Figure 17(b). A (perfectly) spherical BALL with mass M is balanced on the pivot, so that the entire assembly is static and balanced at physical equilibrium. Let switch sw' be a predecessor to switch sw such that an enabling signal from

sw' initiates sw into switching. (In this example, switch sw realizes the negation operation. A similar construction can be used to illustrate the realization of the identity function.) When sw' enables sw — an action resulting from measuring the state of sw', which, in this example, yields a value 1 as detailed in Section XII-C below in the context of sw. As a result, the PIVOT moves to the right as shown in Figure 17(b). This in turn (eventually) causes the BAR to tilt as shown, caused by the BALL rolling to the left. The following elementary fact from classical mechanics helps understand the dynamics of this assembly once the PIVOT is displaced by a small distance δ in either direction.

Observation 12.1: For any $\delta > 0$, once the pivot has moved and the bar has tilted, subsequent measurements cannot move the pivot in the opposite direction since the horizontal component of the force generated by the bar on the pivot prevents it from doing so.

Returning to the example of Figure 17, the movement of the bar in turn moves a piston through the three frictionless pulleys labeled *a,b,c* and *d*. The piston is in a cylinder enclosing, for purpose of example, a single molecule of an ideal mono-atomic gas, and the cylinder has perfectly conducting frictionless walls , and is in an isothermal bath (not shown). Thus, the movement of the piston (isothermally) compresses the gas. In the deterministic case, the piston is pushed to the position marked D (as shown in Figure 17(d)) in the context of a deterministic switch, which is at a point corresponding to half of the cylinder's volume V as shown. For simplicity, let us assume that that the underlying gaseous system is ergodic–the unimolecular assumption has important implications to the ergodicity of the gaseous system and influences the measurement to yield a particular probability p; the discussion in Section XV-B below will address this point in greater detail. Given this assumption, the volume elements and hence microstates leading to the definition of $\mathcal{V}$ (modeling measurement) in Section VI-B, are associated with the Euclidean coordinates of the volume enclosed within the cylinder.

Observation 12.2: In the case of the uncompressed gas occupying the whole of the cylinder when pistons P1 and P2 are fully retracted, the probability of the gas molecule occupying the left or the right half is equal representing the state $\mathcal{S}_\Gamma$ before switching. Also, the probability of the gas molecule existing in one of the halves is unity when either pistons P1 of P2 are at point marked D, while the other piston is fully retracted, respectively denoting the states $\mathcal{S}_1$ or $\mathcal{S}_0$.

In the general probabilistic case, the piston stops at the point marked P in Figure 17(a) with an associated probability parameter p. It is easily seen (Figure 17(b)) that with $x = \frac{V}{2}$, and the occupancy of all physical (Euclidean) volume elements in the cylinder being of equal probability, $q = \left(\frac{1}{p} - 1\right)$ for $\frac{1}{2} \leq p \leq 1$. In each case, the witnesses to the values 0 and 1 are as shown in Figure 17(d). Now, the following observation coupled with Theorem 8.1 clarifies the energy changes associated with compressing the gas as characterized by Theorem 4.3.

Observation 12.3: The change in potential energy of the ball equals the mechanical work done on the gas by the piston.

C. The Measurement Assembly

Let switch sw have switch $\hat{sw}$ as one of its successors. Also, let the function being realized by sw be such that if its input value is 0 (activated by an appropriate enabling signal specified as part of a switching relationship), then it enables switch $\hat{sw}$ with an output value of 1, and vice-versa. For convenience, let us additionally consider that the input enabling signals, and the input values to switch $\hat{sw}$ are both derived from switch sw as part of a single switching relationship. Now, the mechanism for switching $\hat{sw}$ based on the PIVOT-BALL construction is identical to that described in the context of switch sw in Section XII-B above.

The process of measurement is a collision detector at the granularity of a molecule, one detector per (Euclidean) surface element such that the ratio of detectors to volume elements is $1:1$. Now, the arrangement in Figure 17(b) will cause a (minute) movement to the pivot corresponding to a successor of switch sw, triggering a switching. In greater detail, when a collision occurs with the sensor, a miniscule amount of work associated with the measurement is done to transmit this and move the "pivot" of successor switch $\hat{sw}$ of sw.

1) The restraining mechanism: Considering the deterministic case, the enabling signal is defined (in Section VII) to be valid only at the end of the switching interval; if not, false values can be detected. In order to ensure that the result of the measurement be transmitted to $\hat{sw}$ from sw after switch sw has fully switched, namely its piston P2 is in position D, we introduce a particular restraining mechanism shown in Figure 17(c). In our example this corresponds to the point in time when piston P2 is completely in position D. To realize this mechanism, the piston, as it moves, also shifts a restraining ROD (Figure17(c)) which prevents the measurement assembly from transmitting the value encoded by sw to $\hat{sw}$ till piston P2 reaches position D. From the time instant τ when the PIVOT of switch sw starts moving, to the instant $\tau' > \tau$ when piston P2 reaches position D forms the switching interval $\hat{\tau}$. This mechanism can be easily extended to the case of a switch with any associated probability parameter p. Thus from Theorem 8.4

Observation 12.4: Not accounting for the cost of measurement, the work done on the gas is exactly $-\kappa t \ln 2$ to reach position D, whereas it is $-\kappa t \ln 2p$ to reach position P, and is derived from the potential energy stored in the ball prior to the switching.

D. Comments on extensions to the multimolecular case

While the preceding discussion is adequate to establish the energetics and energy savings achieved through probabilistic switching at the fundamental limit (from Section VIII), for completeness, we will now outline an approach to extending the

above mechanism to the case of a switch using $N > 1$ molecules. Briefly, in this case, the system in question is designed in exactly the same way as above with a few additional constraints. Thus, in this case, the detection devices along the walls of the cylinder need to satisfy the following additional property: *the first collision in time is detected by one of the detectors after which all further detection is suspended through restraining mechanisms.* Simple extension of the restraining mechanism to the measurement assembly from Section XII-C satisfy this additional condition. Once this condition is satisfied, a straightforward calculation can be used to show that as before, with $x = \frac{V}{2}$, and the occupancy of all physical (Euclidean) volume elements in the cylinder being of equal probability, $q = \left(\frac{1}{p} - 1\right)$ for $\frac{1}{2} \leq p \leq 1$.

Now, the energy needed in the deterministic and probabilistic cases with N molecules need to be multiplied by the factor N compared to those presented in Section VIII, bound from below in the deterministic case and above in the probabilistic case. Thus, in the probabilistic case for example, the energy consumed would be $-N\kappa t \ln 2p$ Joules whereas it would be $-N\kappa t \ln 2$ Joules in the deterministic case—this deterministic switching energy can be shown to be a lowerbound using the techniques from Section VIII-B so long as a consistent definition of witness and measurement is used in the deterministic and probabilistic cases—which warrants the following restatement of the energy savings at the limit stated as Corollary 8.5 and observation 12.4.

Observation 12.5: The difference in switching energy between deterministic and probabilistic non-oblivious switches is $-N\kappa t \ln p$ Joules and thus, the potential for energy savings is $-N\kappa t \ln \frac{1}{p}$ Joules.

1) Relationship between volume elements and microstates when $N > 1$: With $N > 1$ molecules, a detection of a molecule in a physical volume element does not correspond to a "pure" microstate. However, the outline above is meant to demonstrate the possibility of using an approximate form of measurement, whose statistical behavior is identical to that of detecting individual (microstate) witnesses—thus, the probability parameter p is preserved in going from $N = 1$ to $N > 1$ molecules; however, the energy bounds as outlined above are different from those based on detecting "pure" witnesses. We also note in passing that it is possible to construct significantly more complicated measurement and restraining assemblies such that the energy consumed is independent of the number of molecules N and thus is identical to the unimolecular case; in such a way that, the switching energy at the fundamental limit is characterized by observation 12.4 rather than by observation 12.5. Since this section is aimed at establishing the physical viability of energy savings at the fundamental limit as characterized in the previous sections rather than being focussed on the construction of gas-based switches, we will not go into these constructions here.

XIII. Complexity theoretic characterization of the power of networks

As shown earlier, a switch sw can be used to realize AND as well as NOT "gates". Disjunction or OR gates can be similarly realized, whose energy complexity in the deterministic and probabilistic cases are identical to those established for AND gates. It will be useful to extend these foundational constructs and results to the broader scope of realizing entire computations and designing energy-aware algorithms, using these switch constructs as building blocks. To accomplish this goal, we will sketch relationships below, between a network $\mathcal{N}$ and established models of computation such Turing machines and circuits, from the theory of computation. Papadimitriou [42] and Sipser [43] provide introductions to this topic. Our goal in providing this characterization is primarily to help place the notion of a network and its associated energy complexity in the context of familiar terrain.

A. Switching and the RABRAM model of computing

In an earlier paper [6], this author introduced the RABRAM model for energy-aware algorithm analysis and design. Akin to the standard random access machine (or RAM) model of computing, its relationship to a switching step was the subject of Section VII-B which implies an immediate reduction between the two models—in the RABRAM the address decoder is abstracted away and is a potential source of additional computational power. A central contribution of this earlier work ([6]) is the demonstration of asymptotic energy savings in the RaBRAM model, in the context of the basic question of detecting whether a given vector of n elements which are drawn from the set $\{0, 1\}$, contains at least one element which is equal to 0. This problem is referred to as the *distinct vector problem*, for variants of which the following results are established (in [6]). Using lower bounds for the deterministic case and upper bounds for the probabilistic case which has a probability of error bound from above by $\frac{1}{n^c}$, *asymptotic energy savings* are show to grow as $\Omega(n)$ using a *randomized value amplification* technique [6]. An interesting aspect of this result is that (as far as can be determined), it is the first asymptotic demonstration of energy savings derived from a probabilistic algorithm when compared to any deterministic counterpart, wherein the complexity of the running time is *identical* to $(\Theta(n))$ in both cases. This result demonstrated that the energy savings are due to probabilistic "switching" as opposed to the (trivial) case of being a by-product of an improvement to the running time achieved by randomization since intuitively, a lower running time may imply lower energy consumption. Similar results have also been obtained by this author for variants of the well-known string-matching problem using randomized *fingerprinting*-based approaches due to Karp and Rabin [44].

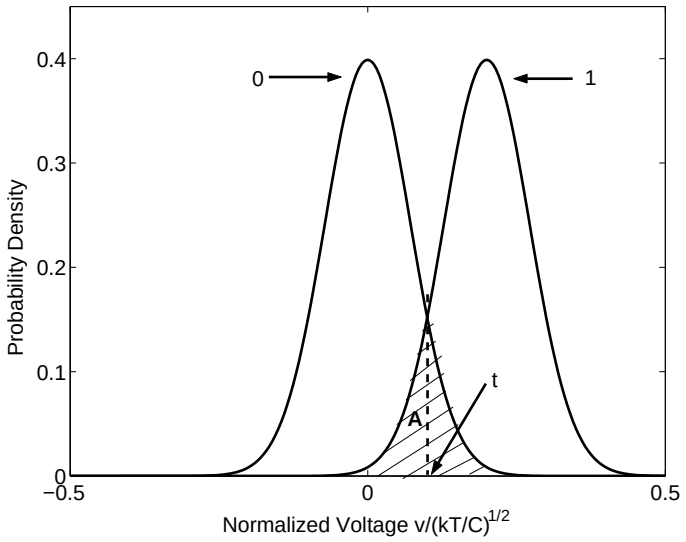


Fig. 18. Canonical realization of a PBIT using voltages

B. Networks and Circuit Complexity

We will use the term *circuit* to refer to the form of boolean circuits that have become ubiquitous in the study of the complexity of computing, and *networks* to refer to the particular model for computing introduced in this paper. Thus, a crucial point of reference for circuits, using this terminology, is embodied in the early work of Pippenger and Fischer [12]; we will use this paper as a basis for definitions. We recognize that since this early characterization, significant strides have been made in circuit complexity, notably in clarifying the power of "monotonicity" as embodied in the work of Razborov [45], [46].

Following Pippenger and Fischer [12], two Turing machines M, M' *simulate* each other, if when they are started with the same string of symbols on their input tapes, they produce the same string of symbols *on-line* on their output tapes; two machines that simulate each other do so on-line if the shifts of the input and output heads occur in the same order (but not necessarily at the same steps) for both machines. This notion of simulation can be naturally extended to involve *intersimulation* between Turing machines, networks, and circuits, viewed as language recognizers within stated resource bounds. As in the case of circuits, a potentially infinite set of networks $\mathcal{N}_1, \mathcal{N}_2, \ldots$ correspond to a single Turing machine, in one-one correspondence with each distinct input size $1, 2, \ldots, M$. In the context of on-line simulation, the energy complexity of networks are related to the number of steps taken by any Turing machine, whereas the size of a circuit is related to the number of steps of an *oblivious* Turing machine. Thus, the energy complexity of a network and the size of a circuit are separated by a gap determined by the separation between oblivious and non-oblivious Turing machines using time-complexity as a measure.

XIV. PHYSICAL REPRESENTATION OF PBITS AND ENERGY SAVINGS

All of the energy bounds developed here are based on a novel approach to representing PBITs as detailed by this author in [47] and [6]. Recall that in all of this work, the value of a PBIT is detected through an instantaneous "measurement" of the existence of a *microstate* from classical thermodynamics. An example of an instantaneous measurement is the detection of the position of a single molecule of gas in a cylinder containing it, as explained by Feynman [29] (and in Section XII). By contrast with this representation, conventional representations of a PBIT would use a value such as voltage, whose *mean* is the value of the PBIT as shown in Figure 18. Following Stein [9]—the mean, a normally distributed noise signal denotes the PBIT value. Thus, the probability of detecting a particular voltage—assuming an arbitrarily precise instantaneous measurement device—is determined by the normal distribution, and the actual value will be a function of the current mean. The overlap area **A** shown in the figure represents the region where a measurement can lead to an erroneous measurement. Quantitatively, the area of overlap between the two density functions, **A**, represents the two-sided error of a PBIT value of 0 being erroneously detected to be 1 and vice-versa. Thus, using the notation from this paper, the one-sided error $(1-p) = \frac{\mathbf{A}}{2}$.

Comparisons between the energy savings using randomization, between the canonical realization of a PBIT and its novel counterpart were determined in collaboration with Cheemalavagu and Korkmaz [11]; for a detailed discussion, the reader is referred there. As shown in Figure 19, while energy savings are possible in both realizations, for a fixed probability p, the novel realization requires lower energy than the canonical representation.

XV. QUESTIONS, CAVEATS AND DIGRESSIONS

In this section, we will address various topics, somewhat germane to the central thesis of this paper; an interested reader might find them useful.

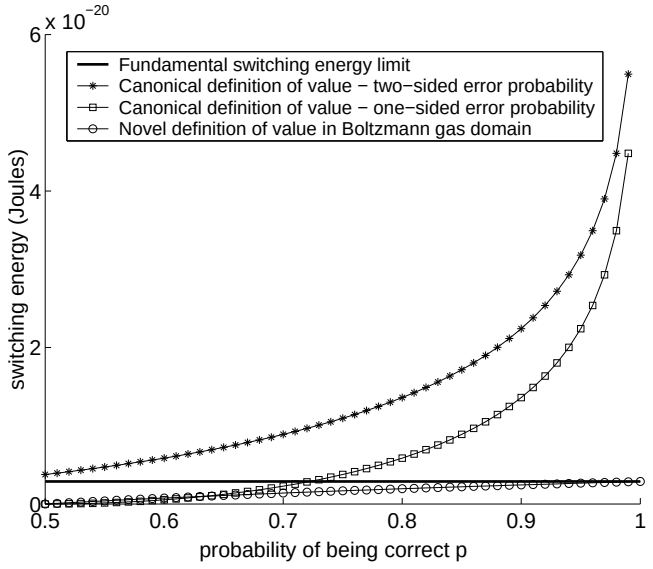


Fig. 19. Comparing energy savings due to probabilistic switching based on conventional and novel definitions of value

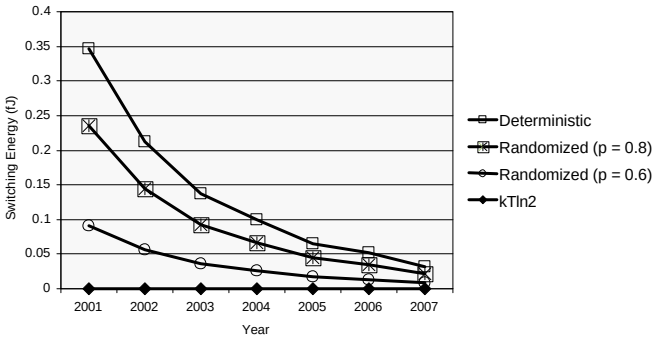


Fig. 20. The energy dissipated by a single switching and a comparison between the deterministic and probabilistic cases

A. Dissipation

A significant issue in modern computing devices, largely based on CMOS semiconductor material, is *dissipation*. Feynman [29], in Chapter 7, discusses this topic and its impact on the idealized devices that he describes in the earlier chapters. The switching energy today tends to be several orders of magnitude above the ideal dissipationless limiting value of $\kappa t \ln 2$ and the idealized models used in this paper. To factor in some of these pragmatic concerns and relate our work to the domain of dissipative switches, we will now clarify the relationship to the context of computing with dissipation, using projections from the *national semiconductor roadmap*.

In Figure 20, we show the trends in energy consumed by switches *with dissipation* by extrapolating the ITRS 2002 roadmap. The non-idealization, or the model for dissipation that we use follows Meindl's "lumped parameter" γ [10]. The projected switching energy in femto Joules for the deterministic as well as the probabilistic switching cases with probability parameters $p = 0.6$ and $p = 0.8$ respectively, are shown. For example, the switching energy associated with deterministic switching in 2006 is achieved in 2002 by a probabilistic switch with $p = 0.6$. Thus, using accepted engineering approximations for switching, probabilistic switches, as shown in Figure 20, accelerate Moore's law by about two generations. All of the above predictions are empirical; for those interested in a theoretical approach to modeling and analyzing dissipation, the reader is referred to Gupta's work [48].

In Figure 21, we have shown the rate at which deterministic switching energy drops with progress in CMOS technology. As shown there, *static energy consumed* in the year 2013 is comparable to that consumed by switching. Thus, switching energy tends to be the dominant issue till that point in time, and the improvements projected by randomization shown in Figure 20 will be a significant factor till then. Based on current projection, after 2016, additional improvements to overcome static energy consumption—efforts underway by technologists in the semiconductor arena now—become necessary to glean further improvements from probabilistic switches.

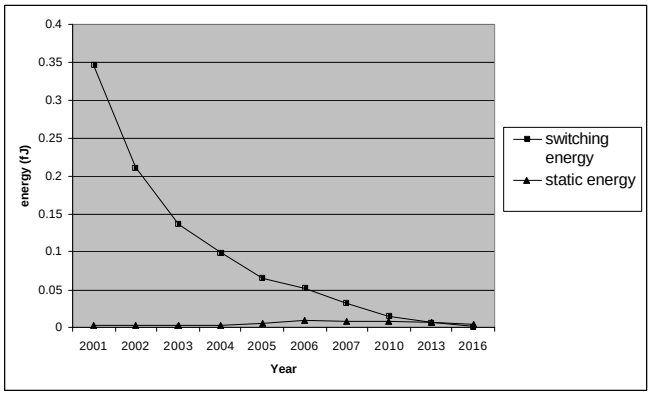


Fig. 21. The relationship between switching energy and static power dissipation

B. Assumptions underlying the idealization

An important mathematical conundrum that emerged from the postulates guiding the development of the kinetic theory of gases is the celebrated *recurrence paradox*. Following Newton's success in providing a (deterministic) mechanistic view of the universe, and Laplace's influence, there was a significant interest in providing a similar framework to explaining the essential *irreversibility*, experimentally noticed in thermodynamics: flow of heat from hot to cold bodies, and not vice-versa. While Boltzmann's statistical approach helped provide an accepted explanation of irreversibility, statistical methods and the concomitant framework of randomness were not universally appealing as part of scientific discourse at the time, with Boltzmann vigorously defending his views against the approaches and criticisms of the determinists.

The recurrence paradox was an important approach aimed at settling this issue mathematically, and the discussion was led significantly by Poincare [49] and his student Zermelo [50]. Ultimately, these debates led to the use of statistical methods and randomness in providing an explanation to physical phenomenon—eventually inspiring Planck to characterize energy as being *quantized* [20] and laying the foundations to modern quantum mechanics. The role of randomness, and more significantly ergodicity, remained an issue that was at the heart of the debate towards the end of the nineteenth century. It finally led to a mathematical characterization of *ergodicity* itself! This concept is especially relevant in the context of our unimolecular assumption in earlier sections, and we will briefly outline some of the relevant issues below.

Based largely on an encyclopedia article of the Ehrenfests [51], it came to be believed that Maxwell and Boltzmann both assumed the *ergodic hypothesis* as part of their development of the theory of gases. It suffices to say that in the switch constructions used here with the walls of a cylinder having irrational angles suffices following the arguments put forward by Boltzmann himself to justify the challenges to ergodicity. Leaping forward to the next century, based respectively on the work of Lebesgue and Brauer, Plancherel [52] and Rosenthal [53] proved the impossibility of systems that satisfy the ergodic hypothesis. In fact, a careful analysis of the Ehrenfests article shows that a weaker form of the ergodic hypothesis—the *quasi-ergodic hypothesis* will suffice from the standpoint of the the Maxwell-Boltzmann developments and we rely on this hypothesis as well.

Abstractly, following Boltzmann (and Maxwell), we will assume all initial states with the same energy to be equally likely; subsequently, Gibbs used this notion formally in the definition of a thermodynamic system using averages over ensembles and defined the *microcanonical ensemble* [19]. In our modeling of a switch, for example in Section XII, we will adopt this strategy of assuming that the initial conditions or Gibbs' "complexions" of the molecule(s) follow these postulates from classical thermodynamics. Specifically the statistical viability of a unimolecular switch follows from a consideration of $N \to \infty$ switches, which, following Gibbs, can be also achieved by N switching steps of a small number (say 2 in the AND gate construction) of switches. However following Boltzmann [18] all possible initial states are equally probable for each of the switching steps. The fact that all volume elements are traversed then follows from the weaker quasi-ergodic hypothesis following Boltzmann's arguments of introducing a cylinder into the volume of gas [54], for example.

ACKNOWLEDGMENT

I wish to thank Lakshmi Chakrapani, Pinar Korkmaz and Rodric Rabbah, in that order, for helping with typesetting and other related editorial issues. Rodric Rabbah helped verify the results pertaining to the relationships between networks, circuits and turing machines. Zeke Zalcstein's comments on the previous paper and his critisism of the abstract notion of a logical mechanism introduced by this author in [6], were crucial in motivating this paper. Jack Schwartz clinched this issue by insisting on a concrete notion of a "switch" to help clarify the subtleties of this work. He also pointed out the importance of clarifying the deeper issues underlying the ergodicity of the gaseous system, as well as the concerns caused by dissipation in the domain of actual switching, when it is not idealized. For all of this, I am very grateful. Suresh Cheemalavagu has been of immense

help, as a “sounding board” in ensuring the engineering relevance of this work, and in verifying the correctness of several of the subtler relationships between the physical domain and the mathematical formalisms; he also helped with numerous editorial comments and provided the data used in Section XV.

REFERENCES

[1] J. T. Schwartz, “Fast probabilistic algorithms for verification of polynomial identities,” *J ACM*, vol. 27, pp. 701–717, 1980.
[2] M. O. Rabin, “Probabilistic algorithms,” *Algorithms and Complexity, New Directions and Recent Trends (ed. J. F. Traub)*, pp. 29–39, 1976.
[3] G. J. Chaitin and J. T. Schwartz, “A note on monte carlo primality tests and algorithmic information theory,” *Communications on Pure and Applied Mathematics*, vol. 31, pp. 521–527, 1978.
[4] R. Balian, *From Microphysics to Macrophysics Vol 1,2*. Springer-Verlag, 1991.
[5] U. V. Vazirani and V. V. Vazirani, “Efficient and secure pseudo-random number generation (extended abstract),” *Annual Symposium on Foundations of Computer Science*, pp. 458–463, 1984.
[6] K. V. Palem, “Energy aware computation: From algorithms and thermodynamics to randomized (semiconductor) devices,” Georgia Institute of Technology, Tech. Rep. GIT-CC-03-10, Feb. 2003.
[7] ——, “Energy aware computing through randomized switching,” Georgia Institute of Technology, Tech. Rep. GIT-CC-03-16, May 2003.
[8] L. Szilard, “On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings,” *in Maxwell’s Demon: Why warmth disperses and time passes by Leff. H, Rex, E.*, 1998.
[9] K.-U. Stein, “Noise-induced error rate as limiting factor for energy per operation in digital ics,” *IEEE Journal of Solid-State Circuits*, vol. SC-31, no. 5, 1977.
[10] J. D. Meindl, “Low power microelectronics: Retrospect and prospect,” *Proceedings of IEEE*, pp. 619–635, Apr. 1995.
[11] K. Palem, S. Cheemalavagu, and P. Korkmaz, “The physical representation of probabilistic bits (pbits) and the energy consumption of randomized switching,” *CREST Technical report*, june 2003.
[12] N. Pippenger and M. J. Fischer, “Relations among complexity measures,” *J. ACM*, vol. 26, pp. 361–381, Apr. 1979.
[13] N. Pippenger, “On simultaneous resource bounds (preliminary version),” in *20th Annual Symposium on Foundations of Computer Science*, 1979, pp. 307–311.
[14] S. Carnot, *Reflections On The Motive Power Of Fire (Reflexions sur la puissance motrice du feu et sur les machines propres a developper cette puissance)*, 1824.
[15] R. Clausius, “Uber die bewegende kraft der warme,” *Annalen der Physik*, 1850.
[16] J. C. Maxwell, “Illustrations of the dynamical theory of gases,” *Phil. Mag.*, vol. 19, pp. 19–32, 1860.
[17] ——, “On the dynamical theory of gases,” *Philosophical Transactions of the Royal Society of London*, vol. 157, pp. 49–88, 1867.
[18] L. Boltzmann and S. G. Brush, *Lectures on Gas Theory (English translation by Stephen G. Brush)*. Dover Publications, 1995.
[19] J. W. Gibbs, *Elementary Principles in Statistical Mechanics*. (New York : Scribner), 1902.
[20] M. Planck and A. T. Ogg, *Treatise on Thermodynamics*. Dover Publications Inc., 1922.
[21] J. von Neumann, *in Fourth University of Illinois lecture in Theory of Self-Reproducing Automata (A. W. Burks editor)*. University of Illinois Press, 1966.
[22] R. Landauer, “Irreversibility and heat generation in the computing process,” *IBM J. Research and Development*, vol. 3, pp. 183–191, July 1961.
[23] C. H. Bennett, “Logical reversibility of computation,” *IBM J. Research and Development*, vol. 17, pp. 525–532, Nov. 1973.
[24] E. Fredkin and T. Toffoli, “Conservative logic,” in *Proceedings of the Physics of Computation Conference*, 1982, pp. 219–253.
[25] M. O. Rabin and D. S. Scott, “Finite automata and their decision problems,” *IBM J. Res. Develop.*, vol. 3, no. 2, pp. 115–125, 1959.
[26] M. O. Rabin, “Probabilistic automata,” *Information and Control*, vol. 6, pp. 230–245, 1963.
[27] R. M. Karp, “Probabilistic analysis of partitioning algorithms for the traveling-salesman problem in the plane,” *Mathematics of Operations Research,(USA)*, vol. 2, no. 3, pp. 209–224, aug 1977.
[28] J. Gill, “Computational complexity of probabilistic turing machines,” *SIAM J. Computing*, vol. 6, no. 4, pp. 675–695, 1977.
[29] R. Feynman, *Feynman Lectures on Computation*. Addison-Wesley Publishing Company Inc, 1996.
[30] L. Boltzmann, *Lecture on Gas Theory*. Dover Publications.1995, Reprinted from Vorlesungen uber Gastheorie, volumes 1-2, 1896-8.
[31] G. Joos and I. Freeman, *Theoretical Physics, Third Edition*. Dover Publications Inc., 1986.
[32] C. Kittel and H. Kroemer, *Thermal physics*. W.H. Freeman and Company, 1980.
[33] T. L. Hill, *An Introduction to Statistical Thermodynamics*. Dover Publications, 1986.
[34] D. Mihalas and B. Weibel-Mihalas, *Foundations of Radiation Hydrodynamics*. Dover Publications, 1999.
[35] S. Hawking, “The direction of time,” *New Scientist*, pp. 46–49, jul 1987.
[36] ——, “The edge of spacetime,” *American Scientist*, pp. 355–359, jul 1984.
[37] H. D. Zeh, *The Physical Basis of The Direction of Time*. Springer-Verlag, 1989.
[38] Z. Kohavi, *Switching and Finite Automata Theory*. New York, McGraw-Hill, 1970.
[39] L. Brillouin, “Maxwell’s demon cannot operate: Information and entropy. i,” *in Maxwell’s Demon: Why warmth disperses and time passes by Leff. H, Rex, E.*, pp. 134–137, 1998.
[40] C. H. Bennett and R. Landauer, “The fundamental physical limits of computation,” *Scientific American*, July 1985.
[41] C. H. Bennett, “Demons, engines and the second law,” *Scientific American*, vol. 257, no. 5, pp. 108–116, Nov. 1987.
[42] C. H. Papadimitriou, *Computational Complexity*. Addison-Wesley Pub. Co., 1994.
[43] M. Sipser, *Introduction to the Theory of Computation*. PWS Publishing Company, 1997.
[44] R. Karp and M. Rabin, “Efficient randomized pattern matching algorithms,” *IBM Journal of Research and Development*, vol. 31, no. 2, pp. 249–260, 1987.
[45] A. A. Razborov, “A lower bound on the monote complexity of the logical permanent,” *Mat. Zametki, 37(6)*, vol. English translation in Math. Notes of the Academy of Sciences of the USSR, p. 1985, 1985.
[46] ——, “Lower bounds on the monote complexity of some boolean functions,” *Dokl. Akad. Nauk SSSR*, vol. English translation in: Soviet Math. Dokl. 31, pp. 798–801, 1985.
[47] K. V. Palem, “Thermodynamics of randomized computing, a discipline for energy aware algorithm design and analysis,” Georgia Institute of Technology, Tech. Rep. GIT-CC-02-56, Nov. 2002.
[48] M. S. Gupta, “Fluctuations and dissipation in elementary one-bit information storage system,” *International Journal of Theoretical Physics*, vol. 21, no. 3/4, pp. 275–282, Apr. 1982.
[49] H. Poincare, “Sur le probleme des trois corps et les equations de dynamique,” *Acta Math.*, vol. 13,1, 1890.
[50] E. Zermelo, “Kinetic theory of gases: An anthology of classic papers with historical commentary,” *Ann Phys. (English trans. by S. G. Brush)*, vol. 3, no. 57, 1896.
[51] P. Ehrenfest and T. Ehrenfest, “The conceptual foundations of the statistical approach in mechanics, cornell university press (1959),” *Enc. Math. Wiss (English trans. by M. J. Moravcsik)*, vol. 4, no. 32, 1911.
[52] M. Plancherel, “Transport theory and statistical physics (1971),” *Arch. Sci. Phys. (English trans. by S. G. Brush)*, vol. 4, no. 33, 1912.

[53] A. Rosenthal, “Transport theory and statistical physics (1971),” *Ann. Phys.(English trans. by S. G. Brush)*, vol. 4, no. 42, 1913.
[54] L. Boltzmann, “Ueber die mechanischne analogien des zweiten hauptsatsez der thermodynamik,” *Journal fur die Reine und Angewandte Mathematik.*, vol. 100, 1887.