Switching Activity Estimation Based on Conditional Independence

Radu Marculescu, Diana Marculescu
and Massoud Pedram

CENG Technical Report 95-04

Department of Electrical Engineering - Systems
University of Southern
Los Angeles, California 90089-2562
(213)740-4458

February 1995

Switching Activity Estimation Based on Conditional Independence

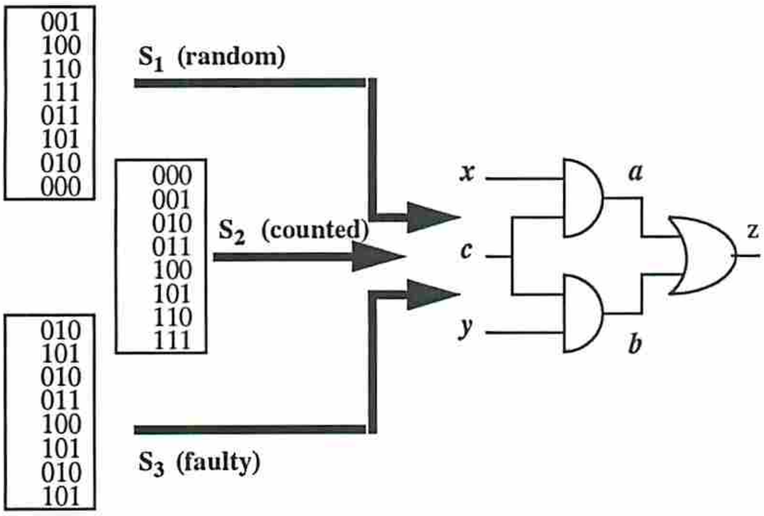
Abstract

Power estimation in combinational modules is addressed from a probabilistic point of view. The zero-delay hypothesis is considered and under highly correlated input streams, the activities at the primary outputs and all internal nodes are estimated. For the first time, the relationship between logical and probabilistic domains is investigated and two new concepts - conditional independence and isotropy of signals - are brought into attention. Based on them, a sufficient condition for analyzing complex dependencies is given. In the most general case, the conditional independence problem has been shown to be NP-complete and thus appropriate heuristics are presented to estimate switching activity. Detailed experiments demonstrate the accuracy and efficiency of the method. The results reported here are useful in low power design.

1. Introduction

Power estimation has demonstrated already its importance in today's electronics. With the growing need for low-power devices, power analysis and optimization techniques have become crucial tasks challenging the CAD community from the architectural to the device level. The key issue in power analysis was from the very beginning switching activity estimation because charging and discharging different load capacitances is by far the most important source of energy dissipation in digital CMOS circuits. Power estimation techniques must be fast and accurate in order to be applicable in practice. Not surprisingly, these two requirements interfere with one another and at some point they become contradictory. General simulation techniques can provide sufficient accuracy, but the price we have to pay for that is too high; one can extract switching activity information by exhaustive simulation on small circuits, but it is unrealistic to rely on simulation results for larger circuits. Few years ago, probabilistic techniques came into the picture and demonstrated their feasibility at least for limited purposes [1], [3]; at that time, it was a good bargain to process combinational and sequential circuits in a few seconds even if the results provided by such an analysis were inaccurate for practical purposes. The reason behind this inaccuracy is that the results were extracted using only the circuit description and assuming the input independence. Signal probability estimation techniques based on global Ordered Binary decision Diagrams (OBDDs) can capture dependencies among internal signal lines, but they are impractical to use on anything other than fairly small circuits [2]. Common digital circuits are dominated by the reconvergent fan-out (RFO) problem; over the years, people working in testing, timing and more recently in power areas have been faced with difficult problems arising from the fan-out reconvergence, mostly when they want to calculate the signal probability [3], [4], [5]. In general, accounting for structural dependencies is a difficult task but when combined with temporal dependencies on circuit inputs is even harder; unfortunately, power estimation techniques must consider all of these dependecies in order to produce accurate results. Let us consider a simple case to illustrate these concepts.

The circuit in Fig.1 is fed successively by three input sequences, S_1, S_2 and S_3; S_1 is an exhaustive random sequence, S_2 is also an exhaustive sequence generated by a 3-bit counter and S_3 is obtained by a 'faulty' 3-bit counter. All three sequences have the same signal probability on lines x, y and c ($p = 0.5$), but even intuitively they are completely different. There are two other measures which differentiate these sequences, namely transition and conditional probabilities, and switching activity calculations should rely on them. In fact, these sequences exercise the circuit such that the number of transitions N_1, N_2, N_3 on each signal line a, b, z becomes quite different once we feed S_1, S_2, or S_3 respectively. In order to accurately compute the number of transitions, calculations based on signal probabilities should undoubtedly account for the influence of reconvergent fan-out, specifically in the previous figure, a and b cannot be considered independent signal lines. This problem could be solved (but only for small circuits) by building global OBDD's in terms of primary inputs, but even so, neglecting the correlations among primary inputs can lead to incorrect results. As we see in this example, assuming input independence for sequences S_2 and S_3, is an unrealistic hypothesis because the patterns in each of them are temporally correlated (for example, each pattern in sequence S_2 is obtained from the previous one by adding a binary 1). Even more than that, transitions as $0 \rightarrow 1$ or $1 \rightarrow 0$ on apparently independent signal lines (like x and c in Fig. 1) are correlated and a detailed analysis on these input streams will reveal strong spatial relationship. Consequently, to compute accurately the switching activity one has to account for both spatial and temporal dependencies starting from the primary inputs and continuing throughout the circuit.



	a	b	z	Total
N_1	2	2	2	6
N_2	2	4	4	10
N_3	0	2	2	4

Fig. 1

Recently, a few approaches which accounts for correlations have been proposed: using an event-driven probabilistic simulation technique, Tsui et al. accounts in [6] only for first-order spatial correlations among probabilistic waveforms. Kapoor in [7] suggests an approximate technique to deal with structural dependencies, but on average the accuracy of the approach is modest. In [8] the authors rely on lag-one Markov Chains and account for temporal correlations; unfortunately, they assume independent transition probabilities among the primary inputs and use global OBDDs to evaluate switching activity (that limits severely the size of the circuits they can process). In [9], an analytical model accounting for spatiotemporal correlations and a technique which gives good results for moderate sized combinational circuits (under random or biased input sequences) are presented; however, the run time is still a problem for large circuits (e.g. more than 2000 gates).

The approach presented in this paper improves the state-of-the art in two ways: theoretically by providing a deep insight about the relationship between the logical and probabilistic domains, and practically by offering a sound mathematical framework and an efficient technique for power analysis. For the first time to our knowledge, the mathematical concept of *conditional independence* is brought into attention and based on it a complete analytical model for power analysis is developed. Defining a new working hypothesis based on the notion of *almost isotropic signals*, this paper presents theoretical and practical evidences that conditional independence is a concept powerful enough to overcome the difficulties arising from structural dependencies as well as highly correlated input streams; more precisely, based on conditional independence and signal isotropy, we give a formal proof showing that the statistics taken for pairwise correlated signals are sufficient enough to characterize larger sets of dependent signals. The practical value of these results becomes particularly evident during optimization and synthesis for low power; a detailed analysis presented here evidentiates the importance of being accurate line-by-line (not only for the total power consumption) and identifies potential drawbacks in previous approaches if the patterns feeding the inputs are highly correlated. To support the potential impact of this research, experimental results are presented for benchmark circuits.

The paper is organized as follows. Section 2 presents in detail the concepts of conditional independence, isotropy and their relationship with switching analysis problem. In section 3 we present a a Markov Chain based approach and an incremental technique for correlation manipulations. Section 4 is devoted to practical aspects: an efficient heuristic for run time improvement and a detailed analysis concerning highly correlated inputs are provided. In section 4 we give our experiences on benchmark circuits ranging from hundreds to thousands of gates. Finally, we summarize our contributions.

2. An Axiomatic Approach to Conditional Probability

2.1. Stochastic Independence

We take the conditional probability as the basic building block in defining a probability structure useful for our purposes. The conventional probability models consist of triplets (Ω, Σ, P) describing an experiment; more precisely, Ω represents the set of all possible outcomes of an experiment, Σ the class of events that are of interest and P is the probability on the basic class of events. If A is an event of the basic class, then the probability of A can be determined by an experiment or may be described on the basis of an earlier known event B; thus the value $P(A \mid B)$ (read as 'probability of A given B') depending on both A and B, becomes the target probability. Consequently, if $\mathcal{B}$ is the set of known events prior to the experiment (but related to it), and $\mathcal{A}$ is the class of events of interest, then $P(\cdot|\cdot)$: $\mathcal{A} \times \mathcal{B} \rightarrow \mathrm{R}^+$ is the basic probability function taken into consideration; this is in some sense motivated by the intuition that, *every probability is in reality conditional* [14].

Definition 1. (Conditional Probability)

If (Ω, Σ, P) is a probability space, $B \in \Sigma$ with $P(B) > 0$, then the *conditional probability of A given B* is:

$$P(A \mid B) = P(A \cap B) \,/\, P(B) \qquad\qquad A \in \Sigma, B \in \Sigma \tag{1}$$

❒

Note: since $P(A \mid B)$ becomes indeterminate for $P(B) = 0$, such events are excluded from the above definition. $P(A \mid B)$ satisfies the axioms of probability; in particular, we have that $0 \leq P(A \mid B) \leq 1$.

Definition 2. (Stochastic Independence)

Let (Ω, Σ, P) be a discrete probability space and let A and B be two events. A and B are said to be *independent* iff

$$P(A\,B) = P(A)\,P(B) \tag{2}$$

❒

Note: if $P(A) > 0$ and $P(B) > 0$ then either all three of the following are true or all three are false: $P(B \mid A) = P(B)$, $P(A\,B) = P(A)\,P(B)$, $P(A \mid B) = P(A)$. Independence of events is primarily a numerical fact about probabilities rather than a fact about their relationship. To emphasize this feature, we will use the term "stochastically independent" instead of saying simply "independent".

Proposition 1. Let $\{A_k\}_{1 \leq k \leq n}$ be events from (Ω, Σ, P). If $P(A_2 \cap A_3 \cap \ldots \cap A_n) > 0$, then

$$P(A_1 \cap \ldots \cap A_{n-1} | A_n) = \prod_{k=1}^{n-1} P(A_k | A_{k+1} \cap \ldots \cap A_n) \tag{3}$$

❒

Proof: $$P(A_1 \cap \ldots \cap A_{n-1} | A_n) = \frac{P(A_1 \cap \ldots \cap A_n)}{P(A_n)} = \prod_{i=1}^{n-1} \frac{P(A_i \cap \ldots \cap A_n)}{P(A_{i+1} \cap \ldots \cap A_n)}$$

$$= \prod_{i=1}^{n-1} P(A_i | A_{i+1} \cap \ldots \cap A_n)$$

■

Corollary 1. If $P(A_i \mid A_{i+1} \cap \ldots \cap A_n) = P(A_i)$, $i = 2, \ldots, n-1$ then (3) becomes

$$P(A_1 \cap \ldots \cap A_m) = \prod_{i=1}^{m} P(A_i) \qquad 1 < m \leq n \tag{4}$$

❐

When (4) holds then we say that $A_1, A_2, \ldots, A_n$ are *mutually* (or *universally) independent* events without regard to any side conditions such as $P(A_i) > 0$; the sequence $A_1, A_2, \ldots, A_n$ has $(2^n - n - 1)$ subsequences of two or more events, then the mutual independence of $A_1, A_2, \ldots, A_n$ is equivalent with the validity of $(2^n - n - 1)$ equations. One natural restriction may arise now: we may consider that equation (4) is true only for each pair of events in the sequence and this involves only $\binom{n}{2}$ equations to be satisfied. This situation corresponds to the *weaker* concept of independence and will be termed as *pairwise independence*. Except for the case $n = 2$, the concepts of mutual independence and pairwise independence are distinct. If $n \geq 3$ events are mutually independent then they are for sure pairwise independent, but not vice-versa; one may construct examples where $n \geq 3$ events pairwise independent are not mutually independent. For example, let us consider the following set of events corresponding to two independent tosses of a fair coin:

- A_1: head on the first toss;
- A_2: head on the second toss;
- A_3: outcomes of the two tosses are different.

As we can see, $P(A_1 A_2) = P(A_1)\, P(A_2)$, $P(A_2 A_3) = P(A_2)\, P(A_3)$ and $P(A_1 A_3) = P(A_1)\, P(A_3)$ so that A_1, A_2 and A_3 are pairwise independent. However, $0 = P(A_3 \mid A_1 A_2) \neq P(A_3) = 0.5$, and hence they are *not* mutually independent. Thus, we may conclude that mutual independence is stronger than pairwise independence.

Proposition 2. [10] If $A_1, A_2, \ldots, A_n$ are mutually independent events and each of $B_1, B_2, \ldots, B_m$ equals the intersection of *some* A_k's, $1 \leq k \leq n$, and if no A_k is used more than once, then $B_1, B_2, \ldots, B_m$ are independent.

❐

Proof: Follows immediately by induction on j $(1 \leq j \leq m)$; first, we have to consider two events B_1, B_2 with $B_1 = A_{n1} A_{n2} \ldots A_{nk}$ and $B_2 = A_{r1} A_{r2} \ldots A_{rs}$ where $n_1, \ldots, n_k, r_1, \ldots, r_s$ are distinct positive integers less or equal to n, then three events $B_1 = A_{n1} A_{n2} \ldots A_{nk}$, $B_2 = A_{r1} A_{r2} \ldots A_{rs}$, $B_3 = A_{t1} A_{t2} \ldots A_{tq}$, and so on, and to keep in mind that reordering independent events does not destroy their independence.

■

2.2. Logic Independence

Based on Proposition 2 and notion of support of a boolean function, we give the following definition:
Definition 3. (Logic Independence)
Two boolean functions f and g are said to be *logically independent* (notation $f \perp g$) iff $Sup(f) \cap Sup(g) = \varnothing$; if they are not logically independent then f and g must share at least one common input variable.
❒

Note: It can be seen from the above definition of f and g that logic independence is a functional notion and does not use any information about the statistics of the inputs. If the hypothesis of independent inputs is satisfied, the two concepts (stochastic and logic independence) coincide due to Proposition 2.
Let us consider the following simple circuits where the primary inputs x, y, c and x, y, c_1, c_2 respectively, are assumed to be stochastically independent:

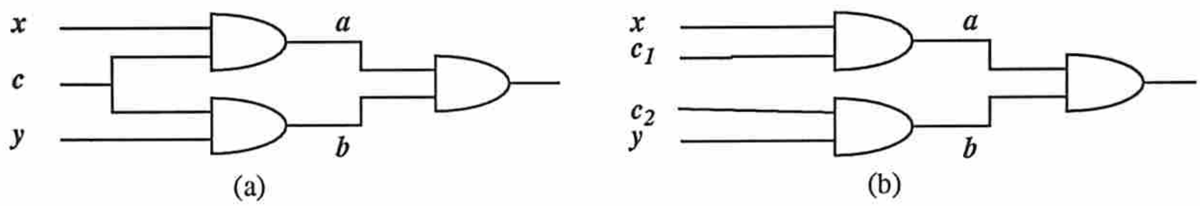


Fig. 2

In (a), we have $p(a) = p(x\,c) = p(x)\,p(c)$, and $p(b) = p(y\,c) = p(y)\,p(c)$ because the pairs (x, c) and (y, c) are stochastically independent; in turn, x, y are stochastically independent and because $p(x\,y\,c) = p(x)\,p(y)\,p(c)$, we may conclude that x, y, c are mutually independent. Signals a and b are not stochastically independent because $p(a\,b) \neq p(a)\,p(b)$. Moreover, lines a and b are not logically independent because $Sup(a) = \{x, c\}$, $Sup(b) = \{y, c\}$ and therefore $Sup(a) \cap Sup(b) = \{c\}$.

For (b), $p(a) = p(x\,c_1) = p(x)\,p(c_1)$, $p(b) = p(y\,c_2) = p(y)\,p(c_2)$ and $p(a\,b) = p(x\,c_1\,y\,c_2) = p(x)\,p(c_1)\,p(y)\,p(c_2)$, and hence $p(a\,b) = p(a)\,p(b)$ so a and b are stochastically independent. Because $Sup(a) \cap Sup(b) = \varnothing$, they are also logically independent.

On the other hand, if we have a sequence on the input such that $p(c) = 1$ (i.e. c is constant one), then $p(a\,b) = p(x\,c\,y\,c) = p(x\,y) = p(x)\,p(y)$ and $p(a) = p(x\,c) = p(x)$, $p(b) = p(y\,c) = p(y)$ so that $p(a\,b) = p(a)\,p(b)$, thus a and b are stochastically independent. For (b), if c_1 has the same behavior as c_2, i.e. $c_1 = c_2$, and c_1 is not constant 1 or 0, we get that $p(a\,b) = p(x\,c_1\,y\,c_2) = p(x)\,p(c_1)\,p(y) \neq p(x)\,p^2(c_1)\,p(y) = p(a)\,p(b)$ thus a and b are stochastically depedent even if they are logically independent. This is not a contradiction; it is rather an example showing that logical independence and stochastical independence are different concepts if the assumption of input independence is dropped.
Coming back to our example (a), both cofactors of a and b with respect to c are logically independent; indeed, $Sup(a^c) = \{x\}$, $Sup(b^c) = \{y\}$ and then $Sup(a^c) \cap Sup(b^c) = \varnothing$. Intuitively, neither stochastic, nor logic independence are sufficient concepts to be used in real circuits where structural dependencies are dominant.

2.3. Conditional Independence

Definition 4. (Conditional Independence)
Let (Ω, Σ, P) be a discrete probability space and let A, B and C be three events; the events A and B are

conditionalyl independent (notation c.i.) with respect to C iff

$$P(A\,B \mid C) = P(A \mid C)\,P(B \mid C) \tag{5}$$

❐

The above definition may be extended to digital signals and to any number of signals as follows:

Definition 5.

Given the set of n signals $\{x_1, x_2, \dots, x_n\}$ and an index i $(1 \le i \le n)$, we say that the subset $\{x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$ is conditionally independent with respect to x_i if the following holds:

$$p(x_1 x_2 \dots x_{i-1} x_{i+1} \dots x_n | x_i) = \prod_{1 \le j \le n,\, j \ne i} p(x_j | x_i)$$

❐

Note: It should be pointed out that if the set $\{x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$is c.i. with respect to x_i, it might not be c.i. with respect to $\bar{x}_i$. However, the corresponding set in which *any* variable (or subset of variables) is complemented, is still c.i. with respect to x_i if the conditions from Definition 5 are met.

For boolean functions, we may give the following property:

Proposition 3. Let f and g be two boolean functions and f^c, g^c the cofactors of f and g with respect to a common variable c; if $f^c \perp g^c$ then f and g are c.i. with respect to c that is,

$$p(f\,g \mid c) = p(f \mid c)\,p(g \mid c) \tag{6}$$

❐

Proof: Shannon's decomposition for f and g gives $f = c\,f^c + c\,f^c$, $g = c\,g^c + c\,g^c$ respectively; consequently, $f\,g = c\,f^c\,g^c + c\,f^c\,g^c$. Applying Definition 1 to calculate $p(f\,g \mid c)$ we get:

$$p(f\,g \mid c) = p(f\,g\,c)\,/\,p(c) = p(c\,f^c\,g^c)\,/\,p(c) = p(f^c\,g^c) = p(f^c)p(g^c)$$

because $Sup(f^c) \cap Sup(g^c) = \varnothing$. We have also:

$$p(f \mid c) = p(f\,c)\,/\,p(c) = p(c\,f^c)\,/\,p(c) = p(f^c)$$

$$p(g \mid c) = p(g\,c)\,/\,p(c) = p(c\,g^c)\,/\,p(c) = p(g^c)$$

therefore $p(f \mid c)\,p(g \mid c) = p(f^c)p(g^c)$ and this concludes our demonstration.

■

In Fig.2, signals a, b are c.i. with respect to c; indeed, we have:

$$p(a\,b \mid c) = p(a\,b\,c)\,/\,p(c) = p(x\,c\,y\,c\,c)\,/\,p(c) = p(x\,y\,c)\,/\,p(c) = p(x)\,p(y)$$

$$p(a \mid c)\,p(b \mid c) = p(a\,c)\,p(b\,c)\,/\,p^2(c) = p(x\,c)\,p(y\,c)\,/\,p^2(c) = p(x)\,p(y)$$

It's worthwhile to note that, in order to compute $p(a\,b\,c)$, if a and b are c.i. with respect to c, we may use only pairwise signal probabilities; as we may deduce by simple manipulations:

$$p(a\,b\,c) = p(a\,b \mid c)\,p(c) = p(a \mid c)\,p(b \mid c)\,p(c) = p(a\,c)\,p(b\,c)\,/\,p(c)$$

which reduces the problem of evaluating the probability of three correlated signals to the one of considering only pairwise correlated signals (if the hypothesis of conditional independence is satisfied).

Consequently, the conditional independence concept can lead to efficient computations even in very complex situations. In fact, proposition 3 gives us a *sufficient* condition for conditional independence and this is very useful from a practical point of view, because all events appearing in digital logic are somehow logically correlated. However, the general problem, to determine a variable x_i from a set of n signals $\{x_1, x_2, \dots, x_n\}$ such that the remaining set of $(n - 1)$ signals is c.i. with respect to x_i is a complex problem; we will prove in the following that it is actually an **NP**-complete problem.

Proposition 5. (Conditional Independence Problem - CIP)

Given a set of n boolean functions $\{x_1, x_2, ..., x_n\}$, an index i and $k \le n - 1$, deciding whether there are at least k signals from the remaining subset c.i. with respect to x_i, is an **NP**-complete problem.

❐

Proof: First, we have to prove that CIP is in **NP**. Indeed, given a particular instance of the problem, it may be verified in polynomial time whether the requirements are met.

We will prove that CIP is **NP**-complete using a reduction from the Set Packing Problem [11]:

*Given a collection C of finite sets $\{ S_1, S_2, ..., S_n \}$, a positive integer $k \le |C|$, deciding whether C contains at least k mutually disjoint sets is **NP**-complete.*

Let C and k be as above, $n = |C| + 1$ and x be a boolean function such that $Sup(x) \cap S_j = \emptyset$ for every $j = 1, 2, ..., n - 1$ where $S_j \in C$. We build the following boolean functions x_j, $j = 1, 2, ..., n - 1$:

$$x_j = xf_j + \bar{x}g_j$$

where f_j and g_j are boolean functions such that $Sup(f_j) = S_j$ and g_j is an arbitrary boolean function.. We can see that there is a subset of at least k signals from $x_1, x_2, ..., x_{n-1}$ which are c.i. with respect to x iff there exists a permutation $i_1, i_2, ..., i_{n-1}$ of $1, 2, ..., n - 1$ such that the following is true:

$$p\left(\prod_{j=1}^{K} x_{i_j} \middle| x\right) = \prod_{j=1}^{K} p(x_{i_j}|x) \qquad \text{where } K \ge k$$

Using the definition of conditional probability and the expression of x_j, we get $\dfrac{p(x\prod_{j=1}^{K} f_{i_j})}{p(x)} = \prod_{j=1}^{K} \dfrac{p(xf_{i_j})}{p(x)}$.

The construction of x_j's was done such that x and f_j's have disjoint supports so, according to Definition 3, they are logically independent, and thus equivalently we get $p(\prod_{j=1}^{K} f_{i_j}) = \prod_{j=1}^{K} p(f_{i_j})$ which is true iff f_{ij} are logically independent i.e. their supports are mutually disjoint: $Sup(f_{i_j}) \cap Sup(f_{i_l}) = \emptyset$ or $S_{i_j} \cap S_{i_l} = \emptyset$ for any $j, l \le K$

Thus, the set of signals built above has at least k signals c.i. with respect to x iff C has at least k mutually disjoint sets. To conclude, CIP is **NP**-complete.

■

One may extend the notion of conditional independence with respect to a single signal to that with respect to a subset of signals. The disadvantage is that, even if we find such a set, we may not express the probability of complex events in terms of probabilities of pairs of events as it is the case with c.i. with respect to a single signal. Thus, from a computational point of view, this does not seem to be useful.

It may seem that the problem we are trying to solve (i.e. finding from a given set of signals, the one which is needed for satisfying the c.i.) is too hard for our purposes since one might try to identify subsets of signals, each being c.i. with respect to different signals. However, this approach is not useful in practice unless the subsets are mutually independent. Since we deal with inputs which are not independent, information about the logic (structural) independence of any subsets of signals is not particularly useful as any *logically uncorrelated* signals may become *stochastically correlated* due to input dependencies. In the following, we will use an approximation of c.i.which holds for correlated inputs.

Definition 5. (Isotropy)

Given the set of n signals $\{x_1, x_2, \ldots, x_n\}$, we say that the c.i. relation is *isotropic*, if it is true for all signals $x_1, x_2, \ldots, x_n$; more precisely, taking out all x_i's one at a time, the subset of the remaining $(n - 1)$ signals is c.i. with respect to the taken x_i.

❐

Returning to our example in Fig.2 (a), given the set of signals $\{a, b, c\}$ we have that $\{a, b\}$ is c.i. with respect to c, but the sets $\{a, c\}$ or $\{b, c\}$ are not c.i. with respect to b, or a, respectively; it follows that c.i. is not isotropic in this particular case. Intuitively, the concept of isotropy as defined above, is very restrictive by its very nature and it is hardly conceivable that a set of signals taken randomly from a target circuit will satisfy Definition 5. Our goal, however, is not to use this concept as it is, but to make it more practical for our purposes. As we shall see later, the main advantage of isotropy is that it offers a canonical approach to the estimation of different kinds of probabilities in digital circuits.

Definition 6. (Almost Conditional Independence)

Given n signals $x_1, x_2, \ldots, x_n$ $(n \geq 3)$, we say that the subset $\{x_j\}_{1 \leq j \leq n, j \neq i}$ is *almost conditionally independent* (notation a.c.i.) with respect to x_i $(i = 1, 2, \ldots, n)$ if there exists an ε $(0 \leq \varepsilon < 1)$ such that:

$$\left| \frac{\prod_{1 \leq j \leq n, j \neq i} p(x_j | x_i)}{p\left(\prod_{1 \leq j \leq n, j \neq i} x_j \middle| x_i\right)} - 1 \right| \leq \varepsilon \qquad (7)$$

❐

In Fig.1, signals a and b are a.c.i. with respect to signal c with $\varepsilon = 0$ (in fact, they are c.i.); also, there exist $0 \leq \varepsilon_1, \varepsilon_2 < 1$ such that a, c are a.c.i. with respect to b, and b, c are a.c.i. with respect to a:

$$\left| \frac{p(a|b)p(c|b)}{p(ac|b)} - 1 \right| \leq \varepsilon_1 \qquad \text{and} \qquad \left| \frac{p(b|a)p(c|a)}{p(bc|a)} - 1 \right| \leq \varepsilon_2$$

In other terms, almost conditional independence is an approximation of conditional independence within given bounds of relative error. We showed that the problem of finding the signal x_i such that the remaining signals are c.i. with respect to x_i, is NP-complete; hence, this less restricted notion seems to be more useful for practical purposes.

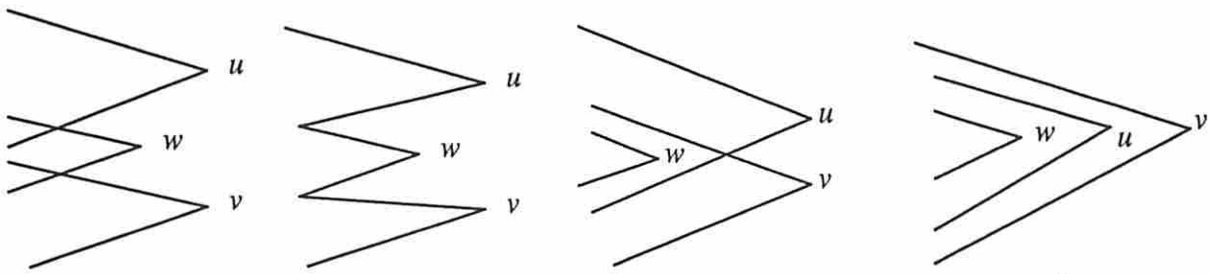
Definition 7. (Almost Isotropy)

The property of conditional independence for a set of n signals $\{x_j\}_{1 \leq j \leq n}$ is called *almost isotropic* (notation a.i.) if there exists some ε $(0 \leq \varepsilon < 1)$ so that it is satisfied within ε relative error for any possible permutation of signals x_i:

$$\left| \frac{\prod_{1 \leq j \leq n, j \neq i} p(x_j | x_i)}{p\left(\prod_{1 \leq j \leq n, j \neq i} x_j \middle| x_i\right)} - 1 \right| \leq \varepsilon \quad \text{for any } i = 1, 2, \ldots, n \qquad (8)$$

❐

Differently stated, a.i. is an approximation of isotropy within given bounds of relative error. A natural question may arise now: in practice, how often it is appropriate to consider a.i. as an approximation of pure isotropy (Definition 5). To answer this question, we consider in Fig.3 several common situations involving the set of signals $\{u, v, w\}$ and the relative position of their logic cones (each cone illustrates the dependence of signals u, v, w on the primary inputs). Whilst the isotropy is completely satisfied only in (*b*), the a.i. concept is applicable in all other cases; more precisely, the c.i. relation is partially satisfied in (*a*) with respect to w, in (*c*) with respect to u and v and in (*d*) with respect to u and v.



a) *b)* *c)* *d)*

Fig. 3

Based on the previous definition, we get the following:

Proposition 6. Given an a.i. set of signals $\{x_j\}_{1 \le j \le n}$ for some ε, the probability of the composed signal $p(\prod_{j=1}^{n} x_j)$ may be estimated within ε relative error as:

$$p\left(\prod_{j=1}^{n} x_j\right) = \frac{\left(\prod_{1 \le i < j \le n} p(x_i x_j)\right)^{\frac{2}{n}}}{\left(\prod_{i=1}^{n} p(x_i)\right)^{\frac{n-2}{n}}} \tag{9}$$

❐

Proof: From the definition of a.i., we get the following, for every $i = 1, 2, \ldots, n$:

$$\left| \frac{\prod_{1 \le j \le n, j \ne i} p(x_j | x_i)}{p\left(\prod_{1 \le j \le n, j \ne i} x_j \middle| x_i\right)} - 1 \right| \le \varepsilon$$

which may be re-written using the definition of conditional probability as:

$$1-\varepsilon \le \frac{\dfrac{\prod_{1\le j\le n,\, j\ne i} p(x_i x_j)}{p^{n-2}(x_i)}}{p\left(\prod_{j=1}^{n} x_j\right)} \le 1+\varepsilon \qquad \text{for each } i = 1, 2 .., n$$

Multiplying all inequalities we get exactly the above claim.

■

This proposition provides us a very strong result: given that n signals are a.i. for some ε, the probability of their conjunction may be estimated within ε relative error using only the probabilities of pairs of signals, thus reducing the problem complexity from exponential to quadratic. This is similar to the concept of pairwise correlation coefficient introduced in [5] and generalized in [9]. However, their approach is mainly intuitive and, unfortunately, does not provide sufficient accuracy for highly correlated signals as we shall see later.

3. A Probabilistic Model for Switching Activity Analysis

3.1. Spatiotemporal Correlations

In order to characterize the signals in the probabilistic domain, we use the model presented in [9] which is briefly summarized in this section..

The behavior of line x is described as a lag-one Markov Chain $\{x_n\}_{n>1}$ over the state set $S = \{0,1\}$ through the transition matrix Q [12]:

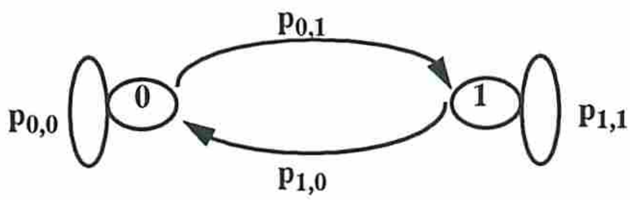


Fig. 4

$$x_n = \begin{cases} 0 \text{ if } x = 0 \\ 1 \text{ if } x = 1; \end{cases} \qquad\qquad Q = \begin{bmatrix} p^x_{0,0} & p^x_{1,0} \\ p^x_{0,1} & p^x_{1,1} \end{bmatrix} \tag{10}$$

Every entry $p_{i,j}$ in the Q matrix represents a conditional probability and may be viewed as the one-step transition probability to state i at step n from state j at step n-1:

$$p^x_{i,j} = p((x(n) = j) \mid (x(n-1) = i)) \qquad \text{for } i, j = 0, 1 \tag{11}$$

A lag-one Markov Chain has the property that one-step transition probabilities do not depend on the 'history', i.e they are the same irrespective of the number of previous steps. Assuming the stationarity of the process $\{x_n\}_{n>1}$, the signal probability may be expressed as follows:

$$p(x = i) = \frac{p^x_{k,i}}{p^x_{1,0} + p^x_{0,1}} \qquad \text{for } i = 0, 1, k = \bar{i} \tag{12}$$

Another useful concept defined and characterized in the original paper is the transition probability for a signal line x:

$$p(x_{i \to j}) = p\left((x(n) = j) \wedge (x(n-1) = i) \right) \qquad \text{for } i, j = 0, 1 \tag{13}$$

It has been showed that transition probabilities are related to conditional probabilities by:

$$p(x_{i \to j}) = \frac{p^x_{k,i} p^x_{i,j}}{p^x_{1,0} + p^x_{0,1}} \qquad \text{for } i, j = 0, 1, k = \bar{i} \tag{14}$$

Definition 8. For any given line x, the switching activity is:

$$sw(x) = p(x_{0 \to 1}) + p(x_{1 \to 0}) = 2\frac{p^x_{1,0} p^x_{0,1}}{p^x_{1,0} + p^x_{0,1}} \tag{15}$$

❒

Note: If the inputs are assumed temporally independent, the switching activity is estimated as:

$$sw(x) = 2 \cdot p(x = 0) \cdot p(x = 1) = 2\frac{p^x_{1,0} p^x_{0,1}}{(p^x_{1,0} + p^x_{0,1})^2}$$

which is quite incorrect; the error made depends on the values of $p^x_{0,1}$ and $p^x_{1,0}$ and thus the results can be well off the mark. This is the reason why previous approaches based on this hypothesis fail to accurately evaluate switching activity.

Temporal correlations are not sufficient for the probabilistic characterization of the circuit: two different signals may be correlated due to RFO or input pattern dependency. In order to capture simultaneous transitions of two signals, we used a lag-one Markov Chain with 4 states (states 0, 1, 2, 3 which stand for encodings 00, 01, 10, 11 of (x,y)):

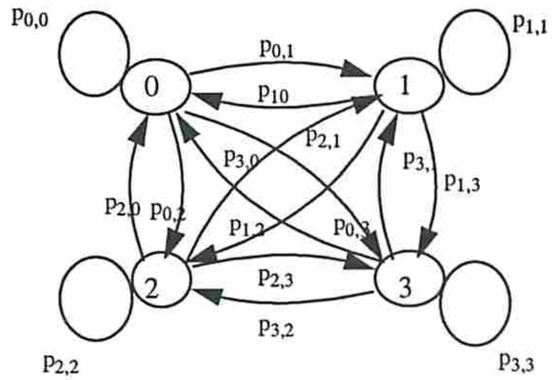


Fig.5

Pairwise correlated signals are characterized by *signal (SC)* and *transition (TC) correlation coefficients*([9]):

$$SC^{xy}_{ij} = \frac{p(x = i \wedge y = j)}{p(x = i)p(y = j)} \qquad TC^{xy}_{ij,kl} = \frac{p(x_{i \to k} \wedge y_{j \to l})}{p(x_{i \to k})p(y_{j \to l})} \tag{16}$$

where $i, j, k, l = 0, 1$.

Starting with this model for capturing the spatiotemporal correlations, we were able to develop a new, more efficient technique, based on the **almost conditional independence** hypothesis. Two approaches were used:

- *The global approach* - for each node, the OBDD is built as a function of the primary inputs;
- *The incremental approach* - for each node, the OBDD is built in terms of its immediate fanin and the transition probabilities and the *TC*'s are propagated through the circuit.

Whilst the first method is more accurate and time/memory consuming, the second provides a sufficient level of accuracy within reasonable bounds of time and space complexity.

3.2. An Incremental Propagation Mechanism Using Almost Conditional Independence

If the almost conditional independence property is satisfied, Proposition 6 may be easily extended to boolean functions represented by OBDDs. Let f be a boolean function of n variables $x_1, x_2, \ldots, x_n$ which may be defined through the following two sets of OBDD paths:

- Π_1 - the set of all paths in the ON-set of f
- Π_0 - the set of all paths in the OFF-set of f

Based on this representation, we give the following result:

Proposition 7. Given f a boolean function of variables $x_1, x_2, \ldots, x_n$, the following hold:

a) If the set $\{x_j\}_{1 \le j \le n}$ in which each variable is either direct or complemented is a.i. for some ε ($0 \le \varepsilon < 1$), then the *signal probability* $p\,(f = i)$ with $i = 0, 1$ may be expressed within ε relative error as:

$$p(x = i) = \sum_{\pi \in \Pi_i} \frac{\left(\prod_{1 \le k < l \le n} p(x_k = i_k \wedge x_l = i_l) \right)^{\frac{2}{n}}}{\left(\prod_{k=1}^{n} p(x_k = i_k) \right)^{\frac{n-2}{n}}} \tag{17}$$

where i_k is the value taken by variable x_k in the cube $\pi \in \Pi_i$.

b) If the set $\{x_{j_{k \to l}}\}_{1 \le j \le n,\ k, l = 0, 1}$ is a.i. for some ε ($0 \le \varepsilon < 1$), then the *transition probability* $p\,(f_{i \to j})$ with $i, j = 0, 1$ may be expressed within ε relative error as:

$$p(f_{i \to j}) = \sum_{\pi \in \Pi_i} \sum_{\pi' \in \Pi_j} \frac{\left(\prod_{1 \le k < l \le n} p(x_{k_{i_k \to j_k}} \wedge x_{l_{i_l \to j_l}}) \right)^{\frac{2}{n}}}{\left(\prod_{k=1}^{n} p(x_{k_{i_k \to j_k}}) \right)^{\frac{n-2}{n}}} \tag{18}$$

where i_k, j_k are the values taken by the variable x_k in the cubes $\pi \in \Pi_i$. and $\pi' \in \Pi_j$.

❐

Proof: a) We may define the event 'f takes the value i' as:

$$(f = i) = \sum_{\pi \in \Pi_i} \prod_{k=1}^{n} (x_k = i_k)$$

In the probabilistic domain, this becomes:

$$p(f = i) = \sum_{\pi \in \Pi_i} p\left(\prod_{k=1}^{n} (x_k = i_k)\right)$$

because the paths $\pi \in \Pi_i$ are disjoint. Since the input variables are a.i. for some ε $(0 \le \varepsilon < 1)$, we may use Proposition 6 to express the probability of each cube, getting exactly (17).

b) The proof is similar to a), considering instead the event 'f switches from i to j' expressed as:

$$f_{i \to j} = \sum_{\pi \in \Pi_i} \sum_{\pi' \in \Pi_j} \prod_{k=1}^{n} x_{k_{i_k \to j_k}}$$

■

The above result may be reformulated using signal and transition correlation coefficients; it may be used in signal probability and switching activity estimation, given that the a.i. conditions are met.

__*Corollary 2.*__ Given a set of signals $\{x_j\}_{1 \le j \le n}$ as in Proposition 7 and a boolean function f of variables $\{x_j\}_{1 \le j \le n}$, the following inequalities hold:

a) $$\left| \frac{p(f = i) - \sum_{\pi \in \Pi_i} \left(\prod_{1 \le k < l \le n} SC_{i_k i_l}^{x_k x_l} \right)^{\frac{2}{n}} \prod_{k=1}^{n} p(x_k = i_k)}{p(f = i)} \right| \le \varepsilon \qquad (19)$$

b) $$\left| \frac{p(f_{i \to j}) - \sum_{\pi \in \Pi_i} \sum_{\pi' \in \Pi_j} \left(\prod_{1 \le k < l \le n} TC_{i_k i_l, j_k j_l}^{x_k x_l} \right)^{\frac{2}{n}} \prod_{k=1}^{n} p(x_{k_{i_k \to j_k}})}{p(f_{i \to j})} \right| \le \varepsilon \qquad (20)$$

Proof: a) and b) Using the definition of transition probability and Proposition 7, we easily get the above inequalities.

■

This result can also be extended to the calculation of correlation coefficients (*SCs* or *TCs*) between two signals in the circuit. From a practical point of view, this becomes an important piece in the propagation mechanism of probabilities and coefficients through the boolean network. We get the following:

__*Proposition 8.*__ Given a set of signals $\{x_j\}_{1 \le j \le n}$ as in Proposition 7, a boolean function f of variables $\{x_j\}_{1 \le j \le n}$, and x a signal from the circuit, the correlation coefficients (*SCs* and *TCs*) satisfy:

a) $$\left|\frac{SC_{ij}^{fx} - \dfrac{\sum_{\pi \in \Pi_i}\left(\prod_{1 \le k < l \le n} SC_{i_k i_l}^{x_k x_l}\right)^{\frac{2}{n+1}} \prod_{k=1}^{n}\left(p(x_k = i_k)\left(SC_{i_k j}^{x_k x}\right)^{\frac{2}{n+1}}\right)}{p(f = i)}}{SC_{ij}^{fx}}\right| \le \varepsilon$$

(21)

b) $$\left|\frac{TC_{ip,jq}^{fx} - \dfrac{\sum_{\pi \in \Pi_i}\sum_{\pi' \in \Pi_j}\left(\prod_{1 \le k < l \le n} TC_{i_k i_l, j_k j_l}^{x_k x_l}\right)^{\frac{2}{n+1}} \prod_{k=1}^{n}\left(p(x_{k_{i_k \to j_k}})\left(TC_{i_k p, j_k q}^{x_k x}\right)^{\frac{2}{n+1}}\right)}{p(f_{i \to j})}}{TC_{ip,jq}^{fx}}\right| \le \varepsilon$$

where $i, j, p, q = 0, 1$.

❐

Proof: a) and b) follow directly from the definition of *SCs* and *TCs* and using the events '$f = i$ and $x = j$ simultaneous' and 'f switches from i to j and x from p to q simultaneously', respectively.

■

These results provide a new heuristic for signal and transition probabilities estimation in the hypothesis of spatiotemporal correlations. As we shall see later, the combination of spatiotemporal correlations and conditional independence provides the best results.

4. Issues in Performance Management

4.1. Inherently Complex Circuits

In real examples, we may have to estimate power consumption in huge circuits like ISCAS benchmarks C6288, C7552, 32-bits multipliers, etc. where global approaches are totally impractical; in such cases, incremental approaches based on correlation coefficients are still applicable, despite the significant amount of CPU time they need for switching activity analysis [9]. Surprisingly enough, there are some other circuits, much simpler (in terms of gate count and structure), which raise a lot of problems in terms of running time; in such cases, the incremental approaches need a large number of backtracks in order to compute the correlations among different signals and in some sense they "degenerate" to global approaches, that is, they tend to behave almost alike at least as far as the running time is concerned.

To begin with, let us consider first ordinary tree circuits with k primary inputs consisting of common type gates (two inputs ANDs, ORs, XORs, etc.). At each level j ($1 < j \le \log_2(k)$) we need to compute for each gate $(4^j - 1) / 3$ correlation coefficients, which adds up to a total of $\theta(k^2)$ calculations for the entire circuit. The running time for tree circuits is thus about 4-5 times that of non-tree circuits with the same number of gates and circuit inputs. This worst-case computation requirement is not present in non-tree circuits. In order to reduce the running time, we found useful the following result:

Proposition 9. If C_j is a correlation coefficient (*SC* or *TC*) at level j (given by a topological order from inputs to outputs of the circuit), then it is related to C_{j-l} ($0 < l < j$) by a proportionality relationship expressible as $C_j \sim (C_{j-l})^{\left(\frac{2}{n+1}\right)^l}$ (24)

where n represents the average fan-in value in the circuit.

❐

Proof: Follows immediately from Proposition 8 if the conditions required in Proposition 9 are satisfied.

■

Corollary 3. If $l \rightarrow \infty$ then the signals behave as uncorrelated.

❐

Proof: Follows immediately from Proposition 9; more precisely, $(C_{j-l})^{\left(\frac{2}{n+1}\right)^l} \rightarrow 1$ when $l \rightarrow \infty$ and that represents the condition of uncorrelation in our approach.

■

In other words, we do not need to compute the coefficients which are beyond some level l in the circuit; instead, we may assume them equal to 1 without decreasing the level of accuracy. Also, *the larger the average fanin n of the circuit, the smaller value for l* may be used. It is worthwhile to note that the c.i. independence, more specifically, the a.i., is essential for this conclusion. The approach based on spatiotemporal correlations *only,* does not provide a sufficient rationale for such a limitation.

This is actually a very important heuristic to use in practice and its impact on running time is huge; limiting the number of calculations for each node in the boolean network to a fixed amount (which depends on the value set as threshold for l) reduces the problem of coefficients estimation **from quadratic to linear complexity.**

4.2 Highly correlated signals

The degree in which the signals are correlated is reflected in the actual values of correlation coefficients; for instance, given $TC^{xy}_{ij,kl} = 1$, $TC^{zt}_{ij,kl} = 4$ and $TC^{uv}_{ij,kl} = 256$, then we may say that the pairs (x, y), (z, t) and (u, v) are uncorelated, slightly correlated and highly correlated, respectively. In general, large values for the coefficients cause a lot of problems in the propagation mechanisms of the coefficients, the main rationale behind this being the approximation formulas/hypotheses used throughout the calculations. Accurate estimation of the switching activity is particularly important in low-power design scenarios when we are interested mostly in point-by-point comparisons among different nodes in the boolean network rather than the total power consumption in the circuit; this need precludes the classical approaches (which do not account for correlations) to have any success in real applications and made us aware of the importance of high signal correlations.

Highly correlated signals may arise everywhere in the circuits, even starting at the primary inputs; for example, considering once again the tree circuit referred in the previous section, if $k = 64$ and we assume its 64 primary inputs fed by a 64-bit binary counter, we should expect values like 2^{63} for some TC coefficients on the inputs. Consequently, we need a really good mechanism to control the error level throughout the circuit; to confirm that our approach indeed keeps the error small, let us consider the benchmark **f51m** and the following two scenarios:

a) Low Correlations: the input patterns are generated by a Linear Feedback Shift Register (LFSR) [13] which implements the primitive polynomial: $p(x) = 1 \oplus x \oplus x^2 \oplus x^7 \oplus x^8$;

b) High Correlations: the input patterns are generated using the state lines of an 8-bit counter.

In order to do a fair comparison between the existing estimation techniques (including the ones which use global OBDDs) and our technique, we had to choose a small sized circuit as **f51m**.We were interested to asses the impact of the correlation level on switching activity estimation in different working hypotheses. In these experiments, two cases were considered: the pseudorandom one in Scenario a and the limit case of non-randomness in Scenario b (when, actually, the input stream is totally deterministic). The estimated values in both cases were compared against the exact values of switching activity obtained by exhaustive simulation; all internal nodes and primary outputs have been taken into consideration. (Fig. 6)

In Scenario a, all approaches are quite accurate. However, we point that only considering spatiotemporal correlations and conditional independence ensures the highest accuracy (**100%** of the nodes estimated with error less than **0.1**). As the results of Scenario b show, the level of correlation impacts strongly the quality of estimation. Specifically, it makes completely inaccurate the global approach based on input independence (despite the fact that internal dependencies due to reconvergent fan-out are accounted by building the global OBDD); as expected, this is visible mostly in Scenario b, where less than **20%** of the nodes are estimated with precision higher than **0.1**. On the other hand, even if temporal correlations are taken into account, but the inputs are assumed to be spatially uncorrelated, only **80%** of the nodes are estimated with error less than **0.1.** Accounting for spatiotemporal correlations provides excellent results for highly correlated inputs (**100%** nodes estimated with precision **0.1**), but the mean error in the hypothesis of conditional independence is anyway smaller (**90%** of the nodes are estimated with error less than **0.05**). This results clearly demonstrate that power estimation is a strongly pattern dependent problem, therefore accounting for dependencies (at the primary inputs and internally, among the different signal lines) is mandatory if accuracy is important; from this perspective, accounting for spatiotemporal correlations in the conditional independence hypothesis seems to be the best candidate to date.

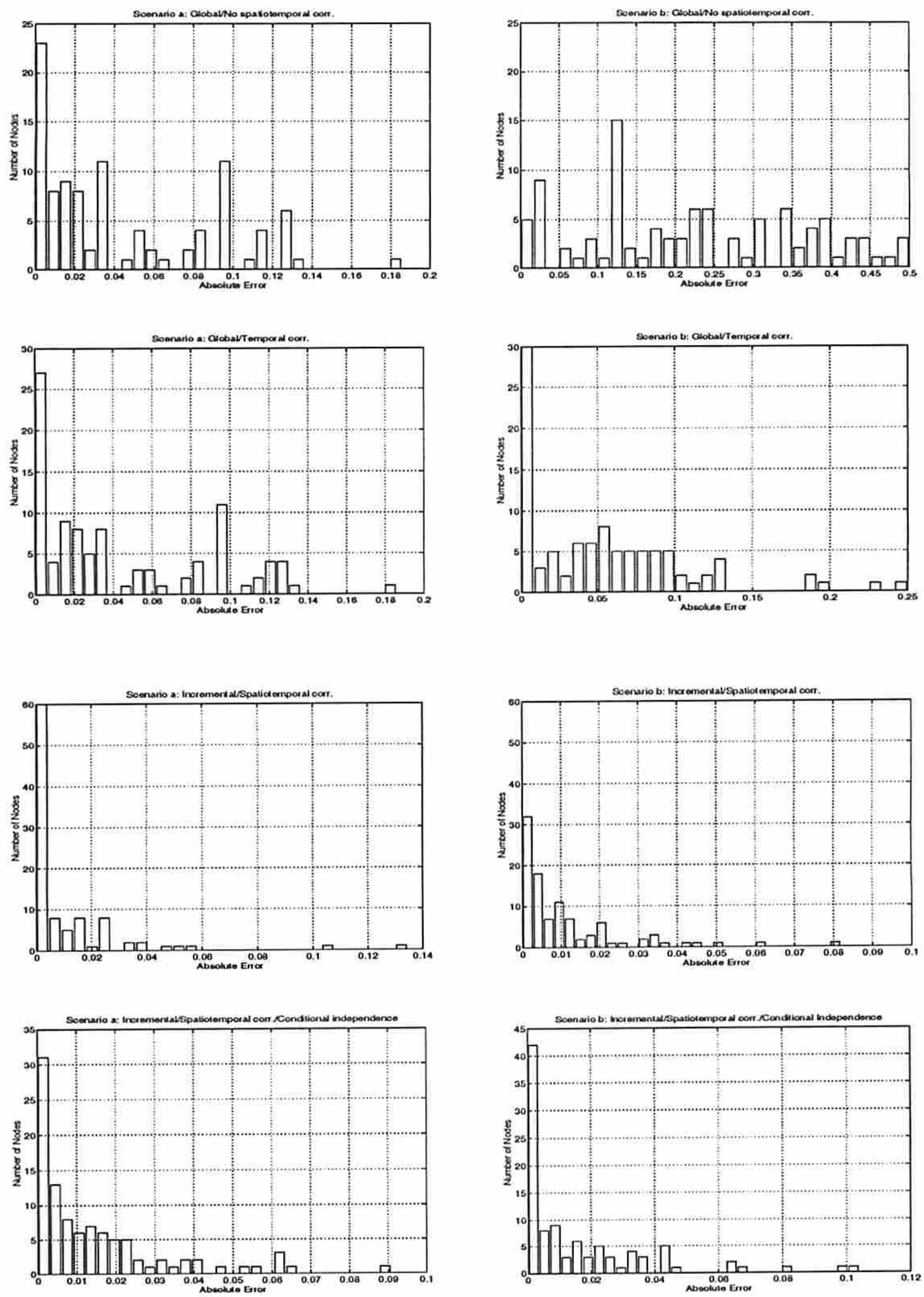


Fig. 6

5. Experimental Results

All experiments were performed in the SIS environment on a Sun Sparc II workstation with 64 Mbytes of memory; the working procedure is shown below:

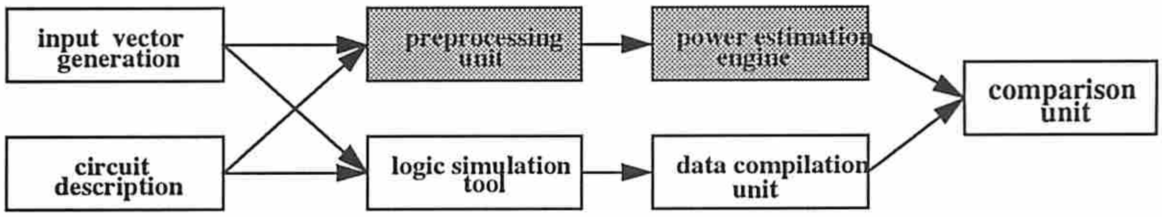


Fig. 7

To generate highly correlated inputs, we used different strategies: modified LFSR generators, generating PseudoRandom (PR) vectors at the inputs of some circuit A and then cascading A with the target circuit B, using the state bit lines of different types of counters, C built-in random functions. In short, we were mainly interested to obtain as many correlations as possible among primary inputs. For large circuits, we tried to keep time/space requirements of the simulation at a reasonable level and we used up to 2^{20} input vectors during the actual logic simulation.

We performed two types of experiments: one to assess the impact of proposed heuristic for speeding up the computation and another one to validate our model based on conditional independence. Switching activity values and power consumption were estimated at *each* internal node and primary output and compared with the ones obtained by actual logic simulations. We found that power estimation for the entire circuit is not a real measure to use in low-power design and power optimization where the switching activity at *each* node has to be accurately estimated with high degree of confidence.

a) Experiments concerning run time improvement

Heuristic proposed in section 4.1 is important in practice not only for substantially reducing the running time but also for keeping the same level of accuracy as the case when the threshold limit is set to infinity. In the following, we present a detailed analysis for the benchmark **duke2** which exhibits a typical behavior; in the first case the limit was set to infinity, in the second one the limit was 4. To report error, we used standard measures for accuracy: maximum error (MAX), mean error (MEAN), root-mean square (RMS) and standard deviation (STD); we excluded deliberately the relative error from this picture, due to its misleading prognostic for small values.

Table 1: duke2 - Speed-Up vs. Accuracy

	LOW CORRELATIONS		HIGH CORRELATIONS	
Error	**NO LIMIT**	**LIMIT=4**	**NO LIMIT**	**LIMIT=4**
MAX	0.0744	0.0710	0.0299	0.0299
MEAN	0.0133	0.0161	0.0056	0.0055
RMS	0.0223	0.0269	0.0085	0.0083
STD	0.0182	0.0219	0.0065	0.0063
TIME	760.2 s	162.7 s	777.7 s	168.8 s

As we can see, the quality of estimation is practically the same in both cases whilst the running time was significantly reduced in the second approach. It should be pointed out, that this limitation works fine also for partitioned circuits which is an essential feature in hierarchical analysis. Running extensively our estimation tool on circuits of various sizes and types (ISCAS benchmarks, adders, multipliers), we observed the following general tendency for speed-up:

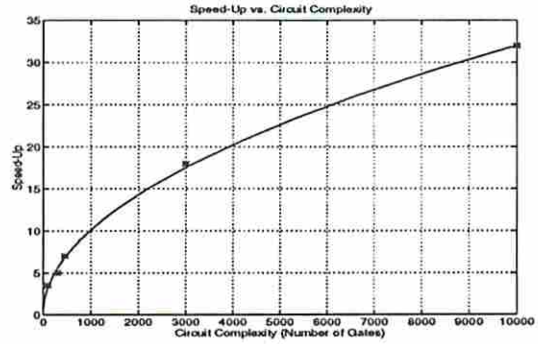


Fig. 8

We can see that, whilst the speed-up is about 3 ÷ 5 times for less complex circuits, it may become 20 ÷ 30 times for large examples; we estimated the power consumption for multipliers on 16 bits (2048 gates) and 32 bits (9124 gates) and the running times were 320.11 and 1052.85 sec. Consequently, we may claim an average time of 150 sec. necessary to process about 1K gates if the threshold limit is set to 4; the time value may be below 100 sec. if the limit is 3.

b) Experiments to validate the conditional independence hypothesis

The experiments were performed on large ISCAS examples using PR and highly correlated inputs (obtained from counted sequences of length 2^{20}); all results reported here, have been derived using the value 4 as the limit for coefficients calculations. To report error, all estimations were verified against exhaustive simulation performed with SIS logic simulator. To calculate dynamic power consumption at any node x, we have used the well-known formula: $P = 0.5\ (V_{dd}^2/T_{cycle})\ C_{load}\ sw(x)$ where V_{dd} is the supply voltage, T_{cycle} is the clock cycle period, C_{load} is the load capacitance and x is the output of the target gate. C_{load} was been estimated as a function of the fanout of the gate.

Table 2: Low correlations on inputs

Circuit	MAX	MEAN	RMS	STD	Total Power uW@20MHz	Time sec.
C432	0.1916	0.0281	0.0465	0.0374	3372.57	104.22
C499	0.0624	0.0134	0.0184	0.0126	7645.56	100.56
C880	0.0691	0.0135	0.0211	0.0164	6391.83	95.02
C1355	0.0225	0.0041	0.0051	0.0030	6797.92	48.44
C1908	0.1315	0.0091	0.0206	0.0185	7435.34	65.11
C3540	0.2010	0.0307	0.0509	0.0407	16356.82	435.14
C6288	0.0890	0.0142	0.0241	0.0196	46846.48	211.42
duke2	0.0710	0.0161	0.0269	0.0219	3611.67	162.7

Table 3: High correlations on inputs

Circuit	MAX	MEAN	RMS	STD	Total Power uW@20MHz	Time sec.
C432	0.2538	0.0225	0.0585	0.0545	306.88	94.99
C499	0.1566	0.0421	0.0760	0.0634	2283.03	107.95
C880	0.0175	0.0013	0.0040	0.0038	263.13	100.09
C1355	0.1930	0.0227	0.0520	0.0469	1865.81	45.11
C1908	0.3907	0.0294	0.0868	0.0820	3156.83	78.54
C3540	0.0279	0.0279	0.0030	0.0030	166.25	444.97
C6288	0.1773	0.0231	0.0521	0.0471	8843.72	201.8
duke2	0.0299	0.0055	0.0083	0.0063	820.87	168.8

It should be stressed that, not only the switching activities at each internal node were completely different as the level of inputs correlation changes, but also the values of total power consumption. For example, for duke2, the total power estimated under weakly correlated inputs was 3611.67 uW, while this value for strongly correlated inputs was 820.87 uW (there is a factor of 4 between the two). The same behavior has been observed for other circuits.To conclude, input pattern dependencies (in particular highly correlated inputs) is an extremely important issue in power estimation, despite other claims which assume independency and randomness on the primary inputs (or worse, throughout the circuit). From this perspective, power analysis needs analytical models to overcome this difficulty. The model we proposed here, based on conditional independence hypothesis while accounting for spatiotemporal correlations, is an efficient and robust analytical solution to this problem.

6. Conclusions

We have proposed an efficient approach for power estimation in large combinational blocks fed by input streams which exhibit high levels of correlation. The work reported here addresses the relationship between logic and probabilistic domains and gives a sufficient condition for analyzing complex dependencies. From this perspective, the new concepts of *conditional independence* and *isotropy* of signals are used in a uniform manner to fulfill practical requirements for fast and accurate estimation. Under general assumptions, the conditional independence problem has been shown to be NP-complete; consequently, efficient heuristics have been provided for probabilities and coefficients calculation. A comparative analysis with the existing techniques and evaluations on benchmarks emphasize the effectiveness and universality of our approach.

References

[1] F. N. Najm, R. Burch, P. Yang, and I. Hajj, 'Probabilistic Simulation for Reliability Analysis of CMOS VLSI Circuits', in *IEEE Trans. on CAD*, vol. CAD-9, April 1990

[2] A. Ghosh, S. Devadas, K. Keutzer, and J. White, 'Estimation of Average Switching Activity in Combinational and Sequential Circuits', in *Proc. 1992 Design Automation Conference*

[3] K. Parker, and E. J. McCluskey, 'Probabilistic Treatment of General Combinational Networks', in *IEEE Trans.on Computers*, vol. C-24, June 1975

[4] J. Savir, G. S. Ditlow, and P. H. Bardell, 'Random Pattern Testability', in *IEEE Trans.on Computers*, vol. C-33, Jan.1984

[5] S. Ercolani, M. Favalli, M. Damiani, P. Olivo, and B. Ricco, 'Testability Measures in Pseudorandom Testing', in *IEEE Trans. on CAD*, vol. 11, June 1992

[6] C.-Y. Tsui, M. Pedram, and A. M. Despain, 'Efficient Estimationof Dynamic Power Dissipation with an Application', in *Proc. 1993 Intl. Conference on Computer Aided Design*

[7] B. Kapoor, 'Improving the Accuracy of Circuit Activity Measurement', in *Proc. 1994 Design Automation Conference*

[8] P. Schneider, U. Schlichtmann, and K. Antreich, 'A New Power Estimation Technique with Application to Decomposition of Boolean Functions for Low Power', in *Proc. 1994 European Design Automation Conference*

[9] R. Marculescu, D. Marculescu, and M. Pedram, 'Switching Activity Analysis Considering Spatiotemporal Correlations', to appear in *Proc. 1994 Intl. Conference on Computer Aided Design*

[10] A. Thomasian, 'The Structure of Probability Theory with Applications', McGraw-Hill Book Co., 1969

[11] M. Garey, and D. Johnson, 'Computers and Intractability', New York: Freeman, 1979

[12] A. Papoulis, 'Probability, Random Variables, and Stochastic Processes', McGraw-Hill Co., 1984

[13] P. H. Bardell, W. H. McAnney, and J. Savir, 'Built-in Test for VLSI: Pseudorandom Techiques', J. Wiley & Sons Inc. 1987

[14] M. M. Rao, 'Conditional Measures and Applications', New York: M. Dekker, 1993