# Efficient Power Estimation for Highly Correlated Input Streams*

Radu Marculescu, Diana Marculescu, Massoud Pedram

Department of Electrical Engineering - Systems
University of Southern California, Los Angeles, CA 90089

**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 logic 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.**

## I. Introduction

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 tag 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. A few years ago, probabilistic techniques came into the picture and demonstrated their feasibility at least for limited purposes [1], [2]; 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 for this inaccuracy was 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 spatial and temporal dependencies on circuit inputs it becomes even harder. To accurately compute the switching activity one has to account for both spatial and temporal dependencies starting from the primary inputs and continuing throughout the circuit. Recently, a few approaches which account for correlations have been proposed: using an event-driven probabilistic simulation technique, Tsui et al. account 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 (this severely limits the size of the circuits that can be handled). In [9], an analytical model accounting for spatiotemporal correlations and a technique which gives good results for moderate sized combinational circuits are presented; however, the run time is still a problem for large circuits.

The approach presented in this paper improves the state-of-the art in two ways: theoretically, by providing a deep insight into the relationship between the logic 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

*This research was supported in part by the National Science Foundation's Young Investigator Award under contract no. MIP-9457392.

evident during optimization and synthesis for low power; a detailed analysis presented here demonstrates the importance of being accurate *line-by-line* (not only for the total power consumption) and identifies potential drawbacks in previous approaches. 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 Markov Chain based approach and an incremental technique for power estimation. 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 5 we give our experiences on benchmark circuits ranging from hundreds to thousands of gates. Finally, we summarize our contributions.

## II. An Axiomatic Approach to Conditional Probability

### A. Stochastic Independence

Conventional probability models consist of triplets ($\Omega$, $\Sigma$, $P$) describing an experiment; more precisely, $\Omega$ represents the set of all possible outcomes of an experiment, $\Sigma$ is 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 $P(\cdot|\cdot)$: $\mathcal{A} \times \mathcal{B} \rightarrow \mathbf{R}^+$ is the basic probability function that is considered here; this is in some sense motivated by the intuition that, *every probability is in reality conditional* [14].

***Definition 1.*** (Conditional Probability)
If ($\Omega$, $\Sigma$, P) is a probability space, $B \in \Sigma$ with $P(B) > 0$, then the *conditional probability of A given B* is:

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

❒

***Note:*** $P(A|B)$ satisfies the axioms of probability; in particular, we have that $0 \le P(A|B) \le 1$.

***Definition 2.*** (Stochastic Independence)
Let ($\Omega$, $\Sigma$, P) be a discrete probability space and let $A$ and $B$ be two events. $A$ and $B$ are said to be *independent* iff

$$P(A\,B) = P(A)\,P(B) \qquad (2)$$

❒

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.*** If $P(A_i \mid A_{i+1} \cap \ldots \cap A_n) = P(A_i)$, $i = 1,\ldots, n - 1$ then:

$$P(A_1 \cap \ldots \cap A_m) = \prod_{i=1}^{m} P(A_i) \qquad 1 < m \le n \qquad (3)$$

(Due to space limitation, all proofs are given in [15].)

When (3) holds then we say that $A_1, A_2,\ldots, A_n$ are *mutually (*or *universally) independent* events without regard to any other conditions such as $P(A_i) > 0$. Since the sequence $A_1, A_2,\ldots, A_n$ has $(2^n - n - 1)$ subsequences of two or more events, the mutual independence of $A_1, A_2,\ldots, A_n$ is equivalent to satisfying $(2^n - n - 1)$ equations. One natural restriction may arise now: we may consider that equation (3) is true only for each pair of events in the sequence. This then requires only $\binom{n}{2}$ equations to be satisfied. This situation corresponds to a *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 $n \ge 3$ events are mutually independent then they are for sure pairwise independent, but not vice-versa. Thus, *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, $1 \le k \le n$, and if no $A_k$ is used more than once, then $B_1, B_2,\ldots, B_m$ are mutually independent.❒

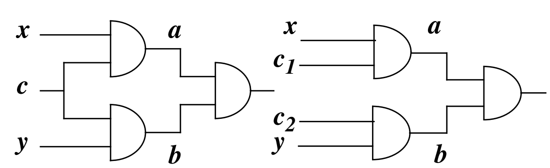
### B. Logic Independence

Based on Proposition 2 and notion of support of a boolean function (i.e. the set of variables on which it depends), 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:




(a) (b)

Fig.1

In (a) $x$, $y$, $c$ are mutually independent; signals $a$ and $b$ are not stochastically independent because $p(a\,b) \ne p(a)\,p(b)$. Moreover, lines $a$ and $b$ are not logically independent. For (b), $a$ and $b$ are stochastically independent and because $Sup(a) \cap Sup(b) = \varnothing$, they are also logically independent.

On the other hand, in (a), if we have an input sequence such that $p(c) = 1$ (i.e. $c$ is constant one), then $a$ 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 $a$ and $b$ are stochastically dependent even if they are logically independent. This is not a contradiction; it rather shows that logic and stochastic independence are different concepts if the assumption of input independence is dropped. Intuitively, neither stochastic, nor logic independence are sufficient concepts to be used in real circuits where structural dependencies are dominant.

*C. 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 *conditionally independent* (notation **c.in.**) with respect to $C$ iff $P(A\ B \mid C) = P(A \mid C)\ P(B \mid C)$.❐

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 $\{x_1, x_2,..., x_n\}$ and an index $i$ ($1 \le i \le n$), we say that the subset $\{x_1, x_2,..., x_{i-1}, x_{i+1},..., x_n\}$ is *conditionally independent* with respect to $x_i$ if $p(\prod_{1 \le j \le n,\, j \ne i} x_j \mid x_i) = \prod_{1 \le j \le n,\, j \ne i} p(x_j \mid x_i)$.❐

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

For boolean functions, we may state the following:
***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.in.** with respect to $c$ that is, $p(f\,g \mid c) = p(f \mid c)\ p(g \mid c)$.❐

In Fig.1, signals $a$, $b$ are **c.in.** with respect to $c$. It is worthwhile to note that, in order to compute $p(a\ b\ c)$, if $a$ and $b$ are **c.in.** with respect to $c$, we may use only pairwise signal probabilities as follows: $p(a\ b\ c) = p(a\ b \mid c)\ p(c) = p(a \mid c)\ p(b \mid c)\ p(c) = p(a\ c)\ p(b\ c)\ /\ p(c)$ which reduces the problem of evaluating the probability of three correlated signals to that of considering only pairwise correlated signals.

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 $\{x_1, x_2,..., x_n\}$ such that the remaining set of ($n$ - 1) signals is **c.in.** with respect to $x_i$ is a difficult problem (actually it is **NP**-complete [15]).

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 in the case of **c.in.** with respect to a single signal. Thus, from a computational point of view, this does not seem to be useful.

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.in.** which holds for correlated inputs as well as for uncorrelated ones.
***Definition 5.*** (Isotropy)
Given the set of $n$ signals $\{x_1,x_2,...,x_n\}$, we say that the **c.in.** relation is *isotropic*, if it is true with respect to every signal $x_i$, $i$ = 1, 2,..., $n$; more precisely, taking out all $x_i$'s one at a time, the subset of the remaining ($n$ - 1) signals is **c.in.** with respect to the taken $x_i$.❐

Returning to our example in Fig.1 (a), given the set of signals $\{a, b, c\}$ we have that $\{a, b\}$ is **c.in.** with respect to $c$, but the sets $\{a, c\}$ or $\{b, c\}$ are not **c.in.** with respect to $b$, or $a$, respectively; it follows that **c.in.** is not isotropic in this particular case. Intuitively, the concept of isotropy as defined above, is very restrictive by its 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 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 $x_1, x_2,..., x_n$ ($n \ge 3$), we say that the subset $\{x_j\}_{1 \le j \le n, j \ne i}$ is *almost conditionally independent* (notation **a.c.in.**) with respect to $x_i$ ($i$ = 1, 2,..., $n$) if there exists an ε ($0 \le \varepsilon < 1$) such that:

$$\left| \frac{\prod_{1 \le j \le n,\, j \ne i} p(x_j \mid x_i)}{p(\prod_{1 \le j \le n,\, j \ne i} x_j \mid x_i)} - 1 \right| \le \varepsilon \qquad (4)$$

❐

In Fig.1, signals $a$ and $b$ are **a.c.in.** with respect to signal $c$ with $\varepsilon = 0$ (in fact, they are **c.in.**); also, there exist $0 \le \varepsilon_1$, $\varepsilon_2 < 1$ such that $a$, $c$ are **a.c.in.** with respect to $b$, and $b$, $c$ are **a.c.in.** with respect to $a$. In other words, almost conditional independence is an approximation of conditional independence within given bounds of relative error.
***Definition 7.*** (Almost Isotropy)
The property of conditional independence for a set of $n$ signals $\{x_j\}_{1 \le j \le n}$ is called *almost isotropic* (notation **a.is.**) if there exists some ε ($0 \le \varepsilon < 1$) so that it is satisfied within ε relative error for any signal $x_i$.❐

Differently stated, **a.is.** is an approximation of isotropy within given bounds of relative error. In practice, is appropriate to consider **a.is.** as an approximation of pure isotropy. Based on the previous definition, we get:

***Proposition 4.*** Given an **a.is.** 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:

$$p\left(\prod_{j=1}^{n} x_j\right) \approx \frac{\left(\prod_{1 \le i < j \le n} p(x_i x_j)\right)^{\frac{2}{n}}}{\left(\prod_{i=1}^{n} p(x_i)\right)^{\frac{n-2}{n}}} \qquad (5)$$

❒

This proposition provides us a very strong result: given that $n$ signals are **a.is.** 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.

## III. A Probabilistic Model for Switching Activity Analysis

### *A. Spatiotemporal Correlations*

In order to characterize the signals in the probabilistic domain, we use the model presented in [9]. Two useful concepts defined in that paper are the *signal probability* $p(x = i)$ and *transition probability* $p(x_{i \to j})$ for a given signal $x$ and $i, j = 0, 1$. Pairwise correlated signals are characterized by *signal (SC)* and *transition (TC) correlation coefficients*:

$$SC^{xy}_{ij} = \frac{p(x = i \wedge y = j)}{p(x = i)p(y = j)} \quad TC^{xy}_{ij,\,kl} = \frac{p(x_{i \to k} \wedge y_{j \to l})}{p(x_{i \to k})p(y_{j \to l})} \qquad (6)$$

where $i, j, k, l = 0, 1$ [9].

Starting with this model for capturing the spatiotemporal correlations, we are able to develop a new, more efficient technique, based on the **almost conditional independence** hypothesis. Two approaches are 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 *TCs* are propagated through the circuit.

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

### *B. An Incremental Propagation Mechanism Using Almost Conditional Independence*

If the almost conditional independence property is satisfied, Proposition 4 may be easily extended to boolean functions represented by OBDDs. Let $f$ be a boolean function of $n$ variables $x_1, x_2,..., x_n$ which may be defined through the following two sets of OBDD paths:

- $\Pi_1$ - the set of all paths in the ON-set of $f$
- $\Pi_0$ - the set of all paths in the OFF-set of $f$

Based on this representation, we give the following result:

***Proposition 5.*** Given $f$ a boolean function of variables $x_1, x_2,..., x_n$, then:

a) If the set $\{x_j^*\}_{1 \le j \le n}$ ($x_j^*$ denotes either $x_i$ or $\overline{x}_i$) is **a.is.** 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:

$$p(f = i) \approx \sum_{\pi \in \Pi_i} \frac{\left(\prod_{1 \le k < l \le n} p(x_k = i_k \wedge x_l = i_l)\right)^{\frac{2}{n}}}{\left(\prod_{k=1}^{n} p(x_k = i_k)\right)^{\frac{n-2}{n}}} \qquad (7)$$

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.is.** 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:

$$p(f_{i \to j}) \approx \sum_{\pi \in \Pi_i} \sum_{\pi' \in \Pi_j} \frac{\left(\prod_{1 \le k < l \le n} p(x_{k_{i_k \to j_k}} \wedge x_{l_{i_l \to j_l}})\right)^{\frac{2}{n}}}{\left(\prod_{k=1}^{n} p(x_{k_{i_k \to j_k}})\right)^{\frac{n-2}{n}}} \qquad (8)$$

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

***Corollary 1.*** Given sets of signals as in Proposition 5 and a boolean function $f$ of variables $\{x_j\}_{1 \le j \le n}$, signal and transition probabilities may be estimated within ε relative error as:

$$p(f = i) \approx \sum_{\pi \in \Pi_i} \left(\prod_{1 \le k < l \le n} SC^{x_k x_l}_{i_k i_l}\right)^{\frac{2}{n}} \prod_{k=1}^{n} p(x_k = i_k) \quad \text{and}$$

$$p(f_{i \to j}) \approx \sum_{\pi \in \Pi_i} \sum_{\pi' \in \Pi_j} \left(\prod_{1 \le k < l \le n} TC^{x_k x_l}_{i_k i_l,\, j_k j_l}\right)^{\frac{2}{n}} \prod_{k=1}^{n} p(x_{k_{i_k \to j_k}}). ❒$$

This result has also been extended to the calculation of correlation coefficients (*SCs* or *TCs*) between two signals in the circuit (see [15]). From a practical point of view, this becomes an important piece in the propagation mechanism of probabilities and coefficients through the boolean network.

## IV. Issues in Performance Management

### *A. 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 ‘‘degenerate’’ in 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 simple gates (two inputs ANDs, ORs, XORs, etc.). At each level $j$ ($1 < j \leq \log_2(k)$) we need to compute for each gate $(4^j - 1) / 3$ correlation coefficients, which add 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 than 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 the following result to be useful:

***Proposition 6.*** 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 $C_j \propto (C_{j-l})^{\left(\frac{2}{n+1}\right)^l}$ where $n$ represents the average fan-in value in the circuit.❒

***Corollary 2.*** If $l \to \infty$ then the signals behave as uncorrelated.❒

In other words, we do not have 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 the value of l*. It is worthwhile to note that the **c.in.**, more specifically, the **a.is.**, is essential for this conclusion. The approach based on spatiotemporal correlations *only,* does not provide sufficient conditions for this conclusion.

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**.

### *B. Highly Correlated Signals*

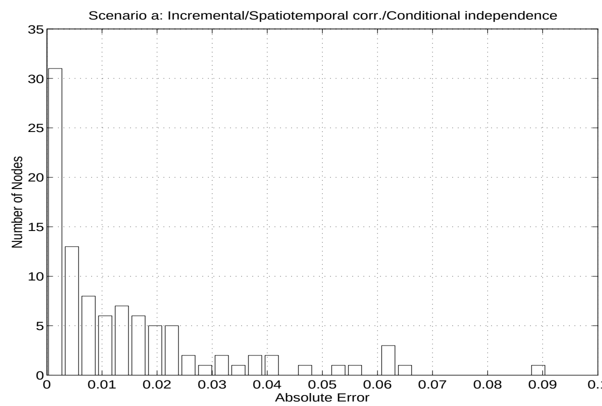
Accurate estimation of the switching activity is particularly important in low-power design scenarios when we are interested primarily in *point-by-point comparisons* among different nodes in the boolean network rather than the total power consumption in the circuit; this need precludes the classical 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. 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 uncorrelated, slightly correlated and highly correlated, respectively.

Highly correlated signals may arise everywhere in the circuits, even starting at the primary 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 up-down 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 such 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 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.2)



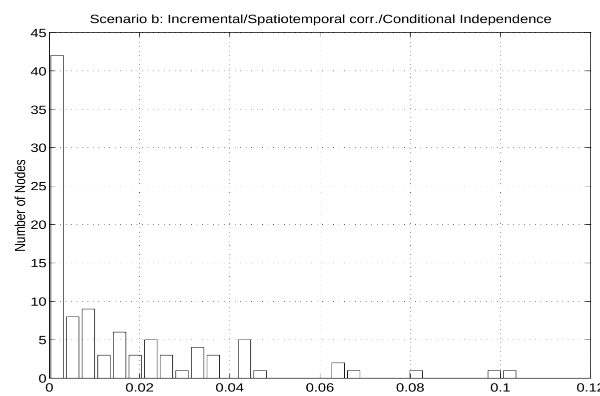




Fig.2

In Scenario a, all approaches are quite accurate. However, we point out that considering spatiotemporal correlations and conditional independence gives the highest accuracy (**100%** of the nodes estimated with error less than **0.1**). However, in Scenario b, the level of correlation strongly impacts 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 for); as expected, 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%** of the 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.

## V. 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:

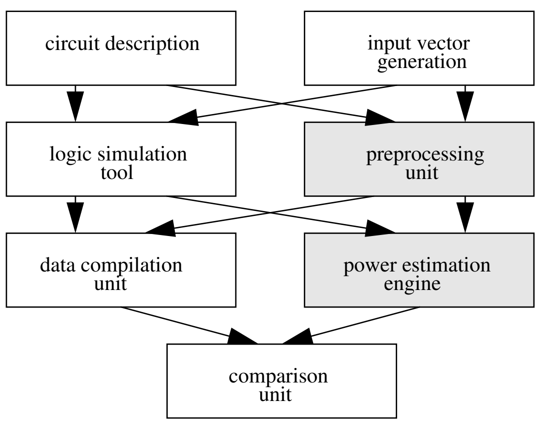



Fig.3

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, built-in random functions in the C language. 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 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

The heuristic proposed in Corollary 2 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:

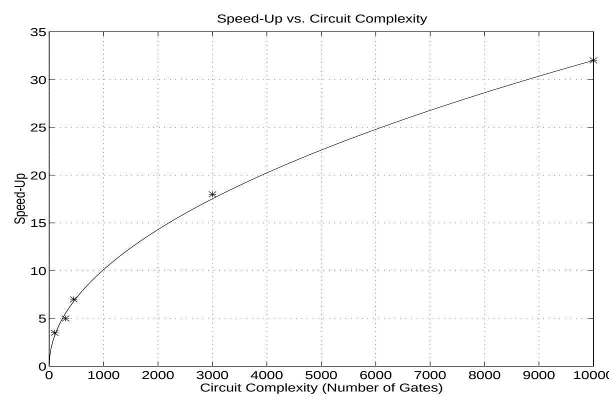



Fig.4

We can see that, whilst the speed-up is about 3 ÷ 5 times for moderate size 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 sec.** and **1052.85 sec.,** respectively. Consequently, we claim an average time of **150 sec.** necessary to process about **1K gates** if the threshold limit is set to 4; the time value is **90 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 the 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. $C_{load}$ has been estimated as a function of the fanout of the gate. Total power consumption is reported in uW @ 20 MHz and running time in seconds.

TABLE 2: LOW CORRELATIONS ON INPUTS

| Circuit | MAX | MEAN | RMS | STD | Power | Time |
|---|---|---|---|---|---|---|
| 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 | Power | Time |
|---|---|---|---|---|---|---|
| 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 **C3540**, the total power estimated under low correlated inputs was **16356.82 uW**, while this value for strongly correlated inputs was **166.25 uW** (there is a factor of **98** between the two). The same behavior has been observed for other circuits. To conclude, input pattern dependence (in particular highly correlated inputs) is an extremely important issue in power estimation, despite other claims which advocate 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.

## VI. 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 P*roc. 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 Estimation of 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 Conferenc*e

[9] R. Marculescu, D. Marculescu, and M. Pedram, ‘Switching Activity Analysis Considering Spatiotemporal Correlations’, 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 Techniques’, J. Wiley & Sons Inc. 1987

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

[15] R.Marculescu, D.Marculescu, and M.Pedram, ‘Switching Activity Estimation Based on Conditional Independence’, Technical Report CENG 95-04, Univ. of Southern California, February 1995