Precise Identification of the Worst-Case Voltage Drop Conditions in Power Grid Verification

Nestoras Evmorfopoulos
University of Thessaly
Department of Computer and
Communication Engineering
Volos, 38221, Greece

nestevmo@uth.gr

Dimitris Karampatzakis
University of Thessaly
Department of Computer and
Communication Engineering
Volos, 38221, Greece

ofni@uth.gr

Georgios Stamoulis
University of Thessaly
Department of Computer and
Communication Engineering
Volos, 38221, Greece

georges@uth.gr

Abstract – Identifying worst-case voltage drop conditions in every module supplied by the power grid is a crucial problem in modern IC design. In this paper we develop a novel methodology for power grid verification which is based on accurately constructing the space of current variations of the supplied modules and locating its precise points that yield the worst-case voltage drop conditions. The construction of the current space is performed via plain simulation and statistical extrapolation using results from extreme value theory. The method overcomes limitations of past methods which either relied on loosely bounding the worst-case voltage drop, or abstracted the current space in a vague and incomplete set of bound-type constraints. Experimental results verify the potential of the proposed method to identify worst-case conditions and demonstrate the pessimism inherent in previous bound-type approaches.

I. INTRODUCTION

The deterioration of the voltage level supplied on the active cells or modules by the lines of the power distribution network (voltage-drop or IR-drop) constitutes one of the biggest reliability problems in modern nanometer-scale VLSI circuits as it adversely affects circuit speed and noise margins [1]-[2]. Upcoming generations of ICs are going to experience substantial voltage drops (due to increased currents and parasitics), which combined to the reduced supply levels (and increased drop-to-supply ratios) will make the situation extremely harsh. To circumvent these problems designers are interested in checking if the designed power grid is *robust*, i.e. if it constantly maintains a safe voltage level at all active modules under all possible loading conditions.

The voltage drop is a dynamic phenomenon, its value at any time instant being dependent on the transient current waveforms drawn over time by the active modules (acting as current sinks). These current waveforms can be determined by means of simulation of the *nonlinear* signal network representing the digital circuit, and the voltage drop may be subsequently computed by solution of the *linear* network representing the power grid where the sink currents are denoted as independent (time-varying) current sources (this is really an approximation since the current sources are voltage-dependent on the supplied level, but the error of ignoring this interdependence and performing the computation in two steps is usually negligible). Since the output of the simulation for the sink currents is a function of the sequence of input patterns applied on the digital circuit, to fully check grid robustness one has to repeat the above procedure for all possible input patterns which is obviously impractical. This means that dynamic analyses can only be executed for a representative set of input patterns. Thus most existing methodologies complement dynamic analyses with some kind of static (DC) analysis in which one vector (or maybe a collection of vectors) of static values for the sink currents is used to represent the transient current waveforms over all input patterns [1]-[2]. Apart from its comprehensive coverage quality, the static analysis is also simpler to implement in terms of modeling and solution of the power grid, since only the resistive model of the latter needs to be extracted and utilized for the calculation of voltage drop. Two obvious choices for the static currents are the average and maximum values of the transient waveforms over all input patterns. Average current values are relatively easy to estimate but may clearly overlook the worst case and miss out conditions for which the grid is essentially non-robust. On the other hand maximum currents do present an upper bound on voltage drop at all sinks, and hence can guarantee grid robustness under all possible input conditions. Estimating maximum sink currents over all input patterns is a challenging problem by itself, and has been addressed in the past by some papers [3]-[4], although more recent techniques for overall maximum power/current estimation of the circuit based on simulation and statistical extrapolation [5]-[8] are more appropriate and accurate for the estimation of maximum current at each sink as well. The problem of using maximum currents as static values for grid verification is that they are overly pessimistic and may erroneously indicate problems that do not exist, since the currents drawn from different sinks are highly correlated and never attain their maximum values simultaneously.

In this paper we develop a novel methodology for deriving a collection of static current vectors that provide a realistic worst-case voltage drop (not just a loose upper bound) at each sink over all input patterns, and also calculate this worst-case voltage drop. The methodology is based on constructing an accurate profile of the space (or locus) of simultaneous sink current variations and locating the specific points that yield the worst-case voltage drop at each sink. Of course in order to find the space of current variations in every detail one has to simulate the circuit for all input patterns which brings about the same obstacle. However, a scaled image of the current space can be obtained via a sample of vectors of sink currents, which can be subsequently extrapolated to the entire space by using results from *extreme value theory* (EVT) [9]. Some past approaches [10]-[11] tried to abstract the current space by prompting the user to enter constraints that express relationships between sink currents. The problem is that the user cannot determine with much accuracy relationships between different sinks and formulate them as constraints, and even if he did, he cannot possibly capture all different relationships and interdependencies in the form of a small number of constraints. A large part of them will surely be missed and the calculated worst-case voltage drops (which only reflect the constraints externally provided) will fall short of the observed reality. On the other hand, it is certainly much more accurate and easier for him to simulate the design for a set of input patterns, and construct the current space exclusively out of the simulation results. In addition, constraints in the previous approaches were in the form of vague upper bounds and thus do not generate the actual current space, but

only a superset of it which is still much pessimistic. The approach in this paper *accurately constructs* the current space (or, specifically, its portion containing the worst-case points for voltage drop) without enclosing it into bounds and accounting for *all* potential correlations between sink currents.

The rest of the paper is organized as follows. The next section provides the necessary background for conditions on sink currents that lead to the worst-case voltage drops. Section III builds on the results of the previous section to formulate a practical methodology for power grid verification. Section IV presents our experimental results from various case studies comprising of test grids and typical benchmark circuits, and finally section V gives the overall concluding remarks.

II. BACKGROUND FOR WORST-CASE VOLTAGE DROP CONDITIONS

The model that we will assume for the power grid is the resistive linear model with time-varying current sources in place of sinks, since we are seeking static current vectors for DC analysis. Instead, however, of using one single DC vector (such as the vector of overall average or maximum currents) to represent all transient waveforms for all possible input patterns, we will adopt a *multi-cycle* DC current scheme [1] in which only the cycle-accurate current waveform for each sink (corresponding to one specific input pattern) is substituted by a constant DC value (Fig. 1). Based on this scheme we will subsequently attempt to find those cycle-DC current vectors that provide the worst-case voltage drop at any sink for the given power grid. This particular scheme is able to account for the diverse current workloads which arise for dissimilar input patterns (due to the different number of devices being switched), and can even faithfully reproduce the transient behavior of the grid on condition that there exists sufficient decoupling capacitance to smooth out large dynamic peaks of voltage drop during the period of a cycle (especially right after a clock edge when multiple devices switch simultaneously). Of course the latter may not be entirely true in modern ICs with operating frequencies above 1GHz, and that is why sample dynamic analyses may still have to be run. However those analyses can never cover all input pattern combinations and thus do not overshadow the value of a static analysis, especially one that is based on the considerably more accurate multi-cycle DC current scheme.
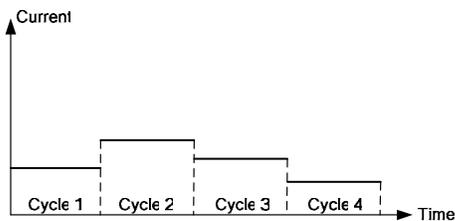

Fig. 1. Multi-cycle DC current scheme.

We assume that the power grid (or the portion of the power grid that needs verification) has been extracted as a linear network of resistive branches that meet at $q + p$ nodes, of which p nodes are connected to the external power supply via power pads (which are either located at the periphery of the grid in the case of a wire-bond package, or are scattered across the entire grid area in a C4 methodology), and the remaining q nodes are divided to n sink nodes (with current sources to an external ground node) and $q - n$ internal nodes (for which usually $q \gg n$). According to the *modified nodal analysis* (MNA) [12], the $q \times 1$ vector of instantaneous voltages $\underline{U}(t)$ (relative to ground) at all internal and sink nodes is determined by the instantaneous currents drawn by the sinks and the specific network structure, and is given by the following matrix equation:

(1) $\qquad \underline{G} \cdot \underline{U}(t) = -\underline{I}(t) + \underline{G} \cdot \underline{V}_{dd}$

where $\underline{G}$ is the $q \times q$ *conductance* matrix of the network (formed by the conductances of the network branches), $\underline{I}(t)$ is a $q \times 1$ vector of current excitations at the nodes (with positive currents to the ground at sink nodes and zero entries at all internal nodes), and $\underline{V}_{dd}$ is another $q \times 1$ vector with all entries equal to the supply voltage V_{dd}. By defining $\underline{V}(t) = \underline{V}_{dd} - \underline{U}(t)$ as the voltage drop at all nodes we can rewrite the above network equation in a form which can be solved directly for the voltage drop values:

(2) $\qquad \underline{G} \cdot \underline{V}(t) = \underline{I}(t)$

In order to obtain appropriate expressions for the voltage drops at the n sink nodes which are of interest, we must enumerate them first in the matrix $\underline{G}$ and vectors $\underline{I}(t)$ and $\underline{V}(t)$, and then we can write eq. (2) as follows:

(3) $\qquad \begin{bmatrix} \underline{G}_{11} & \underline{G}_{12} \\ \underline{G}_{21} & \underline{G}_{22} \end{bmatrix} \cdot \begin{bmatrix} \underline{V}_s(t) \\ \underline{V}_i(t) \end{bmatrix} = \begin{bmatrix} \underline{I}_s(t) \\ \underline{0} \end{bmatrix}$

where $\underline{G}_{11}$, $\underline{G}_{12}$, $\underline{G}_{21}$, and $\underline{G}_{22}$ are submatrices of $\underline{G}$ with sizes $n \times n$, $n \times (q-n)$, $(q-n) \times n$, and $(q-n) \times (q-n)$ respectively, while $\underline{V}_s(t)$ and $\underline{I}_s(t)$ are vectors of size $n \times 1$ (representing the voltage drops and current excitations at the sink nodes respectively) and $\underline{V}_i(t)$ is a vector of size $(q-n) \times 1$ (representing the voltage drops at the internal nodes). The latter equation can be solved with respect to $\underline{V}_s(t)$ and after some calculations gives:

(4) $\qquad \underline{V}_s(t) = (\underline{G}_{11} - \underline{G}_{12} \underline{G}_{22}^{-1} \underline{G}_{21})^{-1} \cdot \underline{I}_s(t) = \underline{R} \cdot \underline{I}_s(t)$

where the matrix $\underline{R} = (\underline{G}_{11} - \underline{G}_{12} \underline{G}_{22}^{-1} \underline{G}_{21})^{-1}$ is of size $n \times n$ and consists only of non-negative values since $\underline{G}$ is an M-matrix. The process of power grid verification typically involves checking that the voltage drop at all sink nodes does not exceed a safety threshold voltage V_0 (e.g. $V_0 = 0.1 V_{dd}$) at all time instants t, i.e. $\underline{V}_s(t) < V_0$, $\forall t \in \Re$. Since the latter is equivalent to $\max_{t \in R} \underline{V}_s(t) < V_0$ (where the "max" operator is interpreted *component-wise* in vector $\underline{V}_s(t)$), we need to find the maximum voltage drop $\max_{t \in R} V_k(t)$ at each sink $1 \le k \le n$. Supposing that the matrix $\underline{R}$ has rows and elements:

$$\underline{R} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix} = \begin{bmatrix} \underline{r}_1^T \\ \underline{r}_2^T \\ \vdots \\ \underline{r}_n^T \end{bmatrix}$$

then each component $V_k(t)$ in vector $\underline{V}_s(t)$ would be as follows:

(5) $\qquad V_k(t) = r_{k1} I_1(t) + r_{k2} I_2(t) + \cdots + r_{kn} I_n(t) = \underline{r}_k^T \cdot \underline{I}_s(t)$

Thus the verification problem is concerned about maximizing the function $V_k(t)$ given by the above equation. Note that the standard practice of taking maximum current values at each sink effectively executes:

$$\max_{t \in \Re} V_k(t) = \max_{t \in \Re} [r_{k1} I_1(t) + r_{k2} I_2(t) + \cdots + r_{kn} I_n(t)]$$

$$\le r_{k1} \max_{t \in \Re} I_1(t) + r_{k2} \max_{t \in \Re} I_2(t) + \cdots + r_{kn} \max_{t \in \Re} I_n(t) = \underline{r}_k^T \cdot \max_{t \in \Re} \underline{I}_s(t)$$

which obviously leads to a very conservative upper bound.

In the latter equation the function of time $V_k(t)$ can be seen as a *composite* function of a vector variable (i.e. a multivariable function) $V_k(\underline{I}_s)$, where $\underline{I}_s = \underline{I}_s(t)$ is a vector-valued function (i.e. a vector of functions) $\underline{I}_s : \Re \to \Re^n$ of the independent scalar variable $t \in \Re$. If $D = \underline{I}_s(\Re) \subset \Re^n$ is the *range space* (or simply range) of $\underline{I}_s(t)$ within the n-dimensional space $\Re^n$, then it constitutes the *domain* on which the composite function $V_k(\underline{I}_s)$ is

defined, i.e. $V_k : D \rightarrow \Re$. The maximum of $V_k(t)$ over all $t \in \Re$ would then be equal to the maximum of $V_k(\underline{I}_s)$ within its domain $D \subset \Re^n$, i.e. $\max_{t \in \Re} V_k(t) = \max_{\underline{I}_s \in D} V_k(\underline{I}_s)$. The latter is effectively a problem of maximizing an objective function $V_k(\underline{I}_s)$ over a *feasible set D* that is formed by imposing constraints on the parameter vector $\underline{I}_s$. The problem is thus transferred to constructing a realistic profile of the feasible domain *D* of current variations (via analytical relationships or using a finite number of points in $\Re^n$), or a part of it where the maximum is located, and then finding the maximum of $V_k(\underline{I}_s)$ over this profile.

From eq. (5) the function $V_k(\underline{I}_s)$ is perceived to be a *linear* function of the vector variable $\underline{I}_s$. Also the domain *D* is obviously bounded in $\Re^n$ (since no current can have an infinite value) and therefore has a *boundary ∂D* (which also belongs to *D*, making it a *compact* set in $\Re^n$). Based on these observations, the following facts will help constructing the domain *D* and ease the task of finding the maximum of $V_k(\underline{I}_s)$ over it.

Definition 1. A point $\underline{I}$ is called a *maximal* (or *noninferior*) point of the *partially ordered* set $D \subset \Re^n$ [13] if for every $\underline{I}' \in D$ the relation $\underline{I}' \geq \underline{I}$ implies $\underline{I}' = \underline{I}$, or equivalently if there does not exist a $\underline{I}' \in D$ such that $\underline{I}' \geq \underline{I}$ (component-wise) with at least one component $1 \leq k \leq n$ being $I'_k > I_k$.

Theorem 1. Let $V(\underline{I}) = \underline{r}^T \cdot \underline{I}$, $\underline{r} > 0$ (component-wise), be a linear function of the vector variable $\underline{I}$ that is defined on a compact set $D \subset \Re^n$. If $\underline{I}^*$ is a maximizer of $V(\underline{I})$ (i.e. a point that maximizes $V(\underline{I})$), then $\underline{I}^*$ is a maximal point of *D*.

Proof. Since $V(\underline{I}) = \underline{r}^T \cdot \underline{I}$ is a continuous function defined on a compact set, a maximizer $\underline{I}^*$ always exists due to the famous Weierstarss theorem [13]. Suppose this maximizer $\underline{I}^*$ is not a maximal point. Then there would exist a point $\underline{I}' \in D$ such that $\underline{I}' \geq \underline{I}^*$ and $I'_k > I^*_k$ for at least one component $1 \leq k \leq n$. This means that $\underline{r}^T \cdot \underline{I}' > \underline{r}^T \cdot \underline{I}^*$ (since $\underline{r} > 0$) and therefore $\underline{I}^*$ cannot be a maximizer of $\underline{r}^T \cdot \underline{I}$, which is a contradiction.

Theorem 2. Let *P* be the set of maximal points of a compact set $D \subset \Re^n$. Then *P* is a subset of the boundary ∂D, i.e. $P \subset \partial D$.

Proof. Obvious, since any point $\underline{I} \in D$ that does not belong to the boundary ∂D is an *interior* point of *D*, for which there always exists a $\underline{I}' \in D$ such that $\underline{I}' \geq \underline{I}$ and $I'_k > I_k$ for at least one component $1 \leq k \leq n$, which means that $\underline{I}$ cannot be a maximal point.

The above theorems state that the maximum of a linear function $V_k(\underline{I}_s) = \underline{r}_k^T \cdot \underline{I}_s$, $\underline{r}_k > 0$ over a compact set $D \subset \Re^n$ will be located at a maximal point of the boundary ∂D. This is intuitively shown in Fig. 2, where the maximum of the function $\underline{r}_k^T \cdot \underline{I}_s$ is the specific value of the constant *c* for which the *hyperplane* $\underline{r}_k^T \cdot \underline{I}_s = c$ in $\Re^n$ just touches the boundary of *D* (the maximizer $\underline{I}_s^*$ is then the common point of the hyperplane and the boundary ∂D). It is also obvious from above that if $D_k = [\alpha_k, \omega_k] \subset \Re$ is the range of each individual component I_k, $1 \leq k \leq n$ (i.e. the projection of the set $D \subset \Re^n$ on the *k*-axis), then there exists (at least) one maximal point $\underline{I}_s \in D$ for which $I_k = \omega_k = \max_{\underline{I}_s \in D} I_k$.
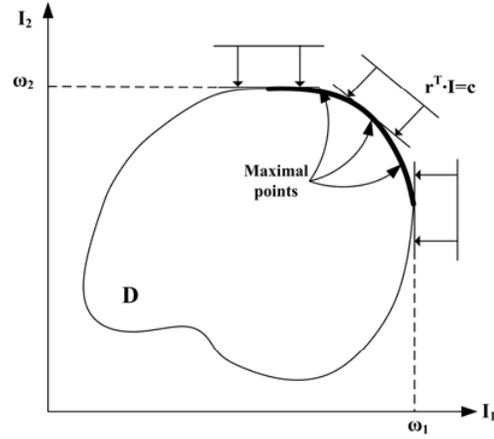

Fig. 2. Maximal points of a compact set $D \subset \Re^n$ and maximization of a linear function over it.

III. DEVELOPMENT OF A PRACTICAL POWER GRID VERIFICATION METHODOLOGY

As outlined in the beginning of the previous section, we will assume a multi-cycle DC scheme for the vector of sink currents $\underline{I}_s(t)$ where each waveform $I_k(t)$, $1 \leq k \leq n$, is substituted by a constant DC value during the period of a clock cycle. This means that the time-dependent current vector $\underline{I}_s(t)$ is actually a function of the pair of binary vectors $\{v_p, v_n\}$ being applied on the digital circuit before and after a clock edge, i.e. $\underline{I}_s(t) = \underline{I}_s(\{v_p, v_n\})$, or alternatively stated, the vector-valued function $\underline{I}_s : \Omega \rightarrow \Re^n$ is defined on the set Ω of all possible pairs $\{v_p, v_n\}$ for the specific circuit (with the cardinality of Ω being $|\Omega| = 4^r$, where *r* is the number of primary inputs).

One can use any type of DC values for the cycle period in a multi-cycle DC current scheme. In this paper we have employed cycle-accurate maximum values, i.e. $I_k(\{v_p, v_n\}) = \max_{t \in [0,\tau]} I_k(t)$, $1 \leq k \leq n$ (where τ is the clock period), which provide an upper bound on voltage drop. However, these values may indeed be slightly pessimistic since they effectively make the assumption that all sinks attain their maximum current simultaneously during the same clock cycle. This means that other per-cycle DC values such as the cycle average $I_k(\{v_p, v_n\}) = \frac{1}{\tau} \int_0^\tau I_k(t) dt$ or the cycle RMS $I_k(\{v_p, v_n\}) = \sqrt{\frac{1}{\tau} \int_0^\tau I_k^2(t) dt}$ might be more accurate eventually, even though they are not entirely foolproof (these can also be computed more efficiently by gate-level simulators instead of expensive transistor-level ones). Nonetheless, the differences inside the same clock cycle are not expected to be very large (compared to the corresponding differences over the entire set Ω of all input pairs).

As maintained in the previous section, in order to be able to find the maximum of $V_k(t)$ we have to construct an accurate profile of the space of current variations *D* (or particularly the maximal points of the boundary ∂D) which now becomes the range space of the function $\underline{I}_s(\{v_p, v_n\})$. We can obtain a sampled image of *D* by *randomly* selecting a number of its points $\underline{I}_{s,i} \in D$, $1 \leq i \leq m$ [i.e for *m* random pairs $\{v_p, v_n\}$] which make up the random sample $\bar{\underline{I}}_s = [\underline{I}_{s,1}, \underline{I}_{s,2}, ..., \underline{I}_{s,m}]^T$ (a visual example in two dimensions taken from an actual circuit is shown in Fig.3). Of course we cannot perform maximization of $V_k(\underline{I}_s)$ solely within

this sample since this would be equivalent to maximizing $V_k(t)$ over the pairs that we have selected for the sample. We must somehow make an inference about the entire space D (or its portion containing the maximal boundary points) on the basis of the available sample.
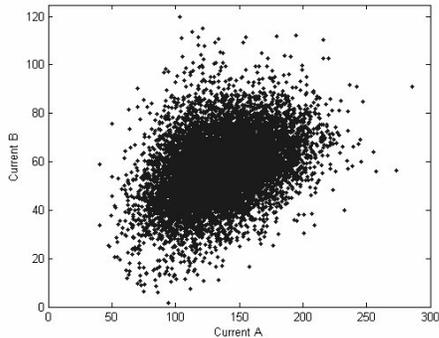

Fig. 3. Sample space of current variations for two sinks in a typical benchmark circuit.

For this we adopt a statistical viewpoint and consider the set Ω of input pairs as a statistical population and the vector function $\underline{I}_s : \Omega \to \Re^n$ as a vector of random variables or a *random vector*. The required range space $D = \underline{I}_s(\Omega) \subset \Re^n$ will then be equal to the range of the random vector $\underline{I}_s$. If the latter is characterized by a multivariate distribution with pdf $f(\underline{I}_s)$ (in which univariate distributions for the constituent random variables I_k can be extracted as *marginals*), then the range coincides with the *supporting domain* or *support* of $f(\underline{I}_s)$ i.e. the set $D = \left\{ \underline{I}_s \in \Re^n : f(\underline{I}_s) > 0 \right\}$ (an example of a bi-variate pdf for a vector of two sink currents in an actual circuit is shown in Fig. 4).
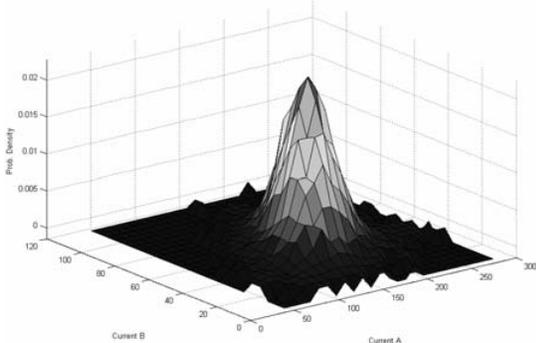

Fig. 4. Bi-variate pdf for a random vector of two sink currents in a typical benchmark circuit.

In the univariate case, the range space of each individual random variable I_k in the vector $\underline{I}_s$ (i.e. the support of its pdf $f_k(I_k)$) is the interval $D_k = [\alpha_k, \omega_k]$ of the real line, where the maximal point of the boundary set $\partial D_k = \{\alpha_k, \omega_k\}$ is the single point $\omega_k = \max_{\underline{I}_s \in D} I_k$ (the maximal point of a compact subset of the real line - where complete order exists - is only one and coincides with its supremum) which is formally known as the *upper endpoint* [9]. The estimation of the upper endpoint of a (bounded) random variable from an existing sample has been performed successfully in [8] (for the problem of maximum power estimation) using elements from univariate extreme value theory (EVT). Specifically, if I_k is a random variable and $\bar{I}_k = [I_{k,1}, I_{k,2}, \dots, I_{k,m}]^T$ is a random sample of size m (with the acquired units $I_{k,i}$, $1 \le i \le m$, forming

themselves "iid" random variables with the distribution of I_k), and if $\bar{I}_k$ is partitioned into m/l sub-samples of size l from which the *maxima* units $Z_{k,j} = \max(I_{k,(j-1)l+1}, \dots, I_{k,jl})$, $1 \le j \le m/l$, are taken out to create a new sample $\bar{Z}_k = [Z_{k,1}, Z_{k,2}, \dots, Z_{k,m/l}]^T$ of size m/l, then an estimate for the upper endpoint ω_k of I_k can be computed as follows:

$$(6) \qquad \hat{\omega}_k = \hat{a}_k + \frac{\hat{b}_k}{1 + l\sqrt{\pi \log l}\left(\text{erf}(\sqrt{\log l}) - 1\right)}$$

where $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$ is the well-known *error function* and $\hat{a}_k, \hat{b}_k$ are *maximum likelihood* (ML) estimates of parameters a_k, b_k that characterize the *asymptotic* distribution of the sample $\bar{Z}_k$ (provided that the size l of the sub-samples is sufficiently large). The latter is assumed to be the so-called *Gumbel* distribution which constitutes the *dominant* asymptotic distribution for maxima, as it was proved in [14] and validated experimentally in [8] (just as the normal distribution is actually the dominant - and not the only - asymptotic distribution of the central limit theorem inside a more general family known as *stable* [15]). This means that $\hat{a}_k$ and $\hat{b}_k$ are to be obtained by maximization of the following log-likelihood function of Gumbel distribution evaluated on the sample $\bar{Z}_k$:

$$(7) \qquad \log L(a_k, b_k) = -\sum_{j=1}^{m/l} \left(\frac{Z_{k,j} - a_k}{b_k} + \exp\left(-\frac{Z_{k,j} - a_k}{b_k} \right) + \log b_k \right)$$

A confidence interval (corresponding to a confidence level of $(1 - \delta) \times 100\%$) has also been constructed in [8] for the estimate (6), and is given by:

$$(8) \qquad |\hat{\omega}_k - \omega_k| \le \frac{z_{\delta/2}}{\sqrt{m/l}} \frac{\hat{b}_n \sqrt{6}}{\pi}$$

$$\cdot \sqrt{(\gamma - 1)^2 + \frac{\pi^2}{6} + \frac{2(1-\gamma)}{1 + l\sqrt{\pi \log l}\left(\text{erf}(\sqrt{\log l}) - 1\right)} + \frac{1}{\left(1 + l\sqrt{\pi \log l}\left(\text{erf}(\sqrt{\log l}) - 1\right)\right)^2}}$$

where $z_{\delta/2}$ is the $\delta/2$ quantile point of the standard normal distribution and $\gamma \approx 0.5772\dots$ is the *Euler gamma* constant.

It must be noted here that in order for the above results to be valid, the distribution of each random variable I_k must be approximately *continuous* even though the population Ω of input pairs is discrete and finite. This is generally safe to assume in practice for circuits with a fair number of primary inputs (in which the population of pairs increases exponentially) and for sink currents I_k with largely diverse values for each different input pair $\{v_p, v_n\}$. The latter observation limits somewhat the applicability of the method to a higher level of hierarchy in the grid where sinks are actually functional modules or rows of standard cells instead of end devices (a total of 50 to 100 devices per sink is quite sufficient to produce a diverse current for each different input pair).

Unfortunately, extreme value theory is not yet developed up to the point where deductions about the boundary and maximal points of the support of a multivariate pdf can be extracted (see [16]-[18] for some recent facts on multivariate EVT). Nevertheless, given a multivariate sample $\bar{\underline{I}}_s$ we can always estimate the upper endpoint ω_k in each of the coordinate axes $1 \le k \le n$ by applying the above results on the univariate samples $\bar{I}_k$ of $\bar{\underline{I}}_s$. Now the sample $\bar{\underline{I}}_s$ (henceforth referred to as the "sample space"), being an image of the current space D, forms a boundary of outermost points in $\Re^n$ and has a set of maximal points of its own. The formed boundary of

this small image, however, will be contracted with respect to the boundary of the vast space D (since there will always be points $\underline{I}_s \in D$ lying outside the outermost boundary of $\overline{\underline{I}}_s$), and in particular the set of its maximal points (or maximal front) will be scaled *down* in each individual coordinate $1 \le k \le n$ (Fig. 5). A sound approximation for this down-scaling of the maximal front as a *whole* in each $1 \le k \le n$ would be $\omega_k - \max_{1 \le i \le m} I_{k,i}$, where $\max_{1 \le i \le m} I_{k,i}$ is the maximum value of each univariate sample $\overline{I}_k$ (i.e. the maximum of the space $\overline{\underline{I}}_s$ in each coordinate axis). Writing this succinctly in vector form for all $1 \le k \le n$ as follows:

(9) $\qquad \underline{\omega} - \max_{1 \le i \le m} \underline{I}_{s,i}$

we obtain a difference vector by which we can shift the sample maximal front in order to transfer it to the location of the maximal front of D in $\Re^n$. The shifted sample maximal points will then be the ones to be selected as static vectors for DC analysis in (4) in order to compute the worst-case voltage drop (of course the maximal front of D will have much different structure and include many more points than the maximal front of $\overline{\underline{I}}_s$, but the maximum value of a linear function is fairly insensitive to the structure of the maximal front and instead depends primarily on its global position in $\Re^n$). To find the maximal points in the space $\overline{\underline{I}}_s$ consisting of m points we must compare each point to all others (to determine whether a specific point is *not* dominated by *any* others in *all* components, according to Definition 1), which leads to a total of m^2 comparisons. However, it can be shown [19] that the necessary comparisons can be reduced to at most $O\left(m(\log_2 m)^{n-2}\right)$, where n is the dimension of the space and its constituent points.
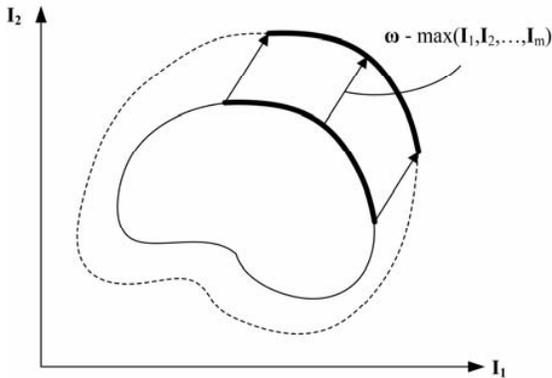

Fig. 5. Sample current space and shift of its maximal points to meet the maximal points of the entire space.

On leaving this section we must point out that we could have actually used the procedure described above for the univariate case to estimate directly the maximum voltage drop over all input pairs at each sink. However, this would be of no real value since the grid is likely to undergo many iterations of redesign and verification until deemed robust, during which the estimation of the maximum voltage drops [involving sample construction, maximization of (7), and evaluation of (6)] would have to be repeated, in contrast to the estimation of maximum sink currents of the digital circuit which needs to be performed only once irrespective of any changes in the power grid. Therefore the objective is actually to pinpoint the worst-case current conditions raised by the underlying circuit rather than compute the maximum voltage drops directly.

IV. EXPERIMENTAL TESTS AND RESULTS

For the experimental validation of the method we have generated a number of test power grids (since there are no universally accepted benchmarks) that will be denoted as Gq-p, where q stands for the number of internal/sink nodes and p for the number of voltage nodes on which supply pads are to be connected (e.g. the label G72-3 denotes a grid with 72 internal/sink nodes and 3 supply pads). All test grids were uniform in structure with random values for the intermediate conductances (within some process-related limits), and also had random placements of the voltage nodes across the grid area (assuming the availability of C4 pads). For the digital circuits supplied by the grids we have implemented the traditional ISCAS85 benchmarks in $0.18\mu m$ and $0.13\mu m$ technologies, and partitioned each one of them to a number of functional modules (representing the n current sinks) which were again assigned to random nodes in every test power grid.

For each digital circuit the process of creation of the sample space (by circuit simulation), univariate EVT estimation in each of the coordinate axes $1 \le k \le n$, and shifting of the maximal points of the sample space (to meet the maximal points of the entire space) is *independent* of the supplying power grid and needs to be carried out only once. The main steps in this process are summarized hereafter along with some brief remarks on their implementation and computational complexity:

- *Generate a total of $m = 5000$ random pairs of binary vectors $\{v_p, v_n\}$ for the circuit under consideration*. This step can be performed by any standard random number generator producing uniform numbers. The selection $m = 5000$ for the number of input pairs is discussed below.

- *Simulate the circuit under all generated pairs and record peak current $\max_{t \in [0,\tau]} I_k(t)$ in each sink $1 \le k \le n$ within a clock cycle.* The recorded data $\overline{I}_k = [I_{k,1}, I_{k,2}, \ldots, I_{k,m}]^T$ taken jointly for all $1 \le k \le n$ as n-dimensional vectors $[\underline{I}_{s,1}, \underline{I}_{s,2}, \ldots, \underline{I}_{s,m}]^T$ will constitute the sample current space $\overline{\underline{I}}_s$. The computational time required to complete this step is entirely up to the simulator program employed, since there are many different simulators with speeds that range considerably depending on the detail of the analysis and their algorithmic efficiency. Although larger circuits will definitely take longer to simulate for every clock cycle, we must emphasize that a total of 5000 input pairs (and consequent cycles) is sufficient to produce a reasonable EVT statistical estimate *independently* of the circuit size or sink size, as is further explained below.

- *Arrange each univariate sample $\overline{I}_k$, $1 \le k \le n$, into $m/l = 100$ sub-samples of size $l = 50$.* Here a size $l = 50$ is chosen initially to ensure the validity of the asymptotic result (6) (in [8] it was selected $l = 100$ but we have found that reducing it by two also gives satisfactory results). The number $m/l = 100$ of sub-samples (leading to a total of $m = 5000$ units) yields estimates with relative estimation error (i.e. quotient of confidence interval to estimate) of about 5% - at a confidence level 95% - for *any* sink irrespective of its size or the size of the broader circuit, as was observed in [8]. This happens because with an increase in the sink size, both the mean and the standard deviation of the distribution of sink currents are increased, but their ratio which determines the relative estimation error remains roughly constant. Only in the case where a smaller estimation error and/or a higher confidence level are desired, the number m/l of sub-samples will have to be increased (together with the total number m of input pairs).

- *For each $1 \le k \le n$ construct the sample $\overline{Z}_k$ of the maxima units from the m/l sub-samples of $\overline{I}_k$.*

- *For each $1 \le k \le n$ maximize the log-likelihood function (7) with respect to the parameters a_k, b_k, in order to obtain the ML estimates $\hat{a}_k, \hat{b}_k$.* This step is carried out via a standard unconstrained optimization algorithm (such as those described in

[20]-[21]). Although this is essentially an iterative numerical procedure, the initial guesses of $\hat{a}_k, \hat{b}_k$ found by the method of moments [8] are extremely close to the final ML values and thus the optimization algorithm usually terminates in a matter of milliseconds (and always converges to the global optimum).

- *For each $1 \le k \le n$ compute the univariate upper endpoint estimate $\hat{\omega}_k$ from (6) and - optionally - its confidence interval (for a given confidence level) from (8).*

- *Determine the maxima $\max_{1 \le i \le m} I_{k,i}$ of all univariate samples $\bar{I}_k$, $1 \le k \le n$, and in conjunction with the estimates $\hat{\omega}_k$, $1 \le k \le n$, compute the n-dimensional difference vector (9).* Here one may also wish to augment each estimate $\hat{\omega}_k$ by the upper end of its confidence interval in the computation of the difference vector, in order to incorporate the inherent statistical estimation error.

The results of executing the above steps in the case of circuits c1355 and c6288, each partitioned up to $n = 5$ modules, are reported in the following Table I. Note that the confidence intervals are for 95% confidence level, and that since the circuit c1355 is about one fifth the size of c6288, the currents were unsurprisingly much smaller in the former than in the latter case.

Table I. Sample and estimated maximum currents at each sink for two benchmark circuits.

Circuit	Current sink	Sample max. (mA)	Estim. max. (mA)	95% Conf. interval	Diff. vector
c1355	sink-A	27.24	35.42	±2.58	8.18
	sink-B	24.86	34.51	±2.45	9.65
	sink-C	28.58	38.60	±3.03	10.02
	sink-D	12.32	16.31	±1.27	3.99
	sink-E	20.37	30.01	±2.61	9.64
c6288	sink-A	273.54	349.91	±23.80	76.37
	sink-B	119.81	157.73	±11.31	37.92
	sink-C	148.28	190.96	±11.46	42.68
	sink-D	100.17	133.45	±7.49	33.28
	sink-E	102.48	119.90	±6.13	17.42

- *Locate the maximal points of the sample space $\bar{I}_s$.* As already mentioned, this step has complexity of $O\left(m(\log_2 m)^{n-2}\right)$

comparisons. The resulting number m_c of maximal points is typically much smaller than the $m = 5000$ points of the sample space (for example, the samples acquired for the circuits c1355 and c6288 had $m_c = 46$ and $m_c = 184$ maximal points respectively).

- *Shift the maximal points of the sample space $\bar{I}_s$ by the computed difference vector.* This step is performed by plain component-wise addition of the difference vector to the m_c maximal points of $\bar{I}_s$, and is a trivial one.

The output of all the above steps is a set of shifted sample maximal points for a particular circuit which approximate the position of the maximal front of its entire current space. Then the DC verification of any grid supplying the circuit is performed by the following steps:

- *Apply the m_c shifted maximal points as static current vectors in (4) to perform an equal number of DC analyses for the given power grid.* This step relies exclusively on a linear network solver, and its execution time is determined by the capability of the solver to carry out m_c DC analyses for the given grid.

- *For each sink $1 \le k \le n$ determine the maximum value among the computed DC voltage drops.* The resulting value for each sink finally constitutes the worst-case voltage drop over all input pairs.

The results for the maximum voltage drops in various test grids supplying the above benchmark circuits are shown in Table II. All computed worst-case voltage drops are compared to accurate statistical estimates obtained by directly applying the univariate EVT procedure [relations (6) to (8)] on samples of voltage drops for the same $m = 5000$ random input pairs (the indicated confidence intervals correspond to 95% confidence level). From the table it can be readily verified that the two estimates come remarkably close to each other. A slight pessimism observed for the proposed method (especially in the case of c6288) can be attributed partly to the selection of cycle-accurate maxima for multi-cycle DC values (as mentioned in the beginning of section III), and in another part to the inevitable deviation of the shifted maximal points of the sample space compared to the maximal points of the entire space which eventually seems to lie on the pessimistic side [i.e. the vector (9) slightly overestimates their relative positions]. Of course we have already stressed that a direct statistical estimation of voltage drop is not a viable solution for a power grid that is subject to many design changes with the same underlying circuit.

Table II. Maximum voltage drop (accurate and pessimistic results) at current sinks of the supplied circuits for various test grids.

Case study	Grid	Circuit	Maximum drop to sink-A (mV)				Maximum drop to sink-B (mV)			
			Our method	Statistical estimation	Pessimistic analysis	% difference	Our method	Statistical estimation	Pessimistic analysis	% difference
#1	G24-1	c1355	41.2	39.3 (±2.6)	47.6	15.5	33.7	32.9 (±2.3)	39.3	16.7
#2	G24-1	c6288	227.4	208.0 (±9.7)	273.3	20.1	242.6	217.8 (±9.8)	293.4	20.9
#3	G23-2	c1355	22.2	21.6 (±1.5)	25.9	16.8	9.2	9.0 (±0.6)	10.5	14.0
#4	G23-2	c6288	174.8	158.0 (±7.2)	207.0	18.4	146.8	127.5 (±5.5)	176.0	19.8
#5	G48-2	c1355	17.5	17.0 (±1.2)	20.4	16.5	14.3	13.7 (±0.9)	16.0	12.1
#6	G48-2	c6288	107.7	95.8 (±4.4)	126.2	17.2	99.8	87.2 (±3.7)	120.4	20.6
#7	G47-3	c1355	20.4	19.6 (±1.3)	23.6	15.3	11.3	10.8 (±0.7)	12.4	10.1
#8	G47-3	c6288	91.6	82.5 (±3.9)	108.0	17.9	132.7	116.1 (±4.6)	161.2	21.5
#9	G72-3	c1355	12.2	11.8 (±0.8)	14.1	15.2	11.8	11.3 (±0.8)	13.4	13.5
#10	G72-3	c6288	131.9	125.0 (±6.3)	158.8	20.4	142.4	125.8 (±5.8)	168.8	18.6
#11	G69-6	c6288	49.7	45.7 (±2.3)	57.2	15.2	49.1	43.1 (±1.9)	58.5	19.3
#12	G94-6	c6288	61.4	57.3 (±2.9)	72.2	17.6	53.5	48.7 (±2.4)	63.9	19.4
#13	G90-10	c6288	47.4	44.6 (±2.3)	55.8	17.7	46.7	41.4 (±2.0)	55.0	17.6
#14	G140-10	c6288	31.5	29.9 (±1.6)	35.6	13.2	24.3	22.9 (±1.2)	27.9	14.9
#15	G135-15	c6288	50.6	49.0 (±2.9)	55.7	10.0	37.1	33.2 (±1.7)	43.0	15.8

Table II. (cont'd)

Case study	Maximum drop to sink-C (mV)				Maximum drop to sink-D (mV)				Maximum drop to sink-E (mV)			
	Our method	Statistical estimation	Pess. anal.	% diff.	Our method	Statistical estimation	Pess. anal.	% diff.	Our method	Statistical estimation	Pess. anal.	% diff.
#1	33.3	31.8 (±2.2)	38.0	14.0	39.2	38.1 (±2.6)	46.3	18.1	38.8	37.6 (±2.6)	45.8	18.1
#2	231.9	210.7 (±9.5)	281.8	21.5	154.4	136.1 (±5.8)	185.2	19.9	206.8	183.1 (±7.2)	253.7	22.7
#3	19.6	19.0 (±1.3)	22.5	14.6	11.2	10.9 (±0.7)	13.2	18.1	22.8	22.8 (±1.6)	27.0	18.7
#4	182.6	161.5 (±6.5)	220.2	20.6	148.1	130.7 (±5.6)	179.3	21.0	174.1	153.9 (±6.0)	211.4	21.4
#5	17.9	17.2 (±1.2)	20.5	14.4	16.8	16.3 (±1.1)	19.9	18.4	19.6	19.3 (±1.4)	23.4	19.4
#6	131.9	117.6 (±5.1)	156.4	18.5	111.8	96.8 (±4.0)	132.4	18.4	106.5	93.9 (±3.5)	130.7	22.7
#7	21.6	20.8 (±1.5)	24.6	13.8	14.7	14.4 (±1.0)	17.3	17.9	19.2	19.2 (±1.4)	23.0	19.5
#8	167.8	145.1 (±5.6)	198.5	18.3	151.5	131.2 (±5.0)	181.5	19.8	152.5	133.5 (±5.1)	184.9	21.3
#9	13.2	12.6 (±0.9)	14.9	13.0	5.0	4.9 (±0.3)	5.9	17.6	14.3	14.0 (±1.0)	16.8	17.6
#10	93.2	84.8 (±3.8)	113.1	21.3	38.2	33.3 (±1.4)	45.3	18.7	131.5	116.1 (±4.7)	160.0	21.7
#11	53.2	46.5 (±1.8)	63.7	19.7	55.9	49.1 (±2.0)	67.2	20.3	56.9	49.9 (±1.8)	69.2	21.7
#12	58.9	53.2 (±2.4)	70.0	18.9	39.0	33.6 (±1.4)	45.4	16.3	29.0	25.5 (±0.9)	35.6	22.8
#13	50.2	44.8 (±2.0)	59.6	18.8	17.7	14.9 (±0.6)	19.6	10.6	36.8	32.9 (±1.4)	44.8	22.0
#14	23.7	21.0 (±0.9)	27.8	17.4	22.8	20.5 (±1.0)	26.2	14.8	24.6	22.8 (±1.0)	29.7	20.5
#15	11.2	10.7 (±0.6)	12.1	8.2	25.5	22.6 (±1.1)	29.7	16.2	10.5	9.6 (±0.4)	12.7	20.8

For each case a pessimistic analysis has also been performed by assuming simultaneous maximum currents at all sinks (i.e. by assigning the estimated univariate maxima $\hat{\omega}_k$, $1 \le k \le n$, to all sinks at once). We can clearly see the overestimation incurred by the standard pessimistic analysis which is about 15%-20% in most cases, irrespective of the grid density and the number of supply pads. We can also observe that the percentage difference is generally smaller in the case of c1355 than c6288 for the same supplying grid, which can be justified by the fact that there exist more complex interdependencies between sinks for larger circuits and they produce greater pessimism when assigning simultaneous maxima to all sinks. Since the ISCAS85 benchmarks are actually small circuits compared to today's standards, the differences between accurate and pessimistic analyses are expected to be even more pronounced in the case of larger designs with several current sinks and more complex interdependencies between them. All the above verify that the standard pessimistic analysis can become a source of many false alarms indicating violations in nodes that are actually legal, a fact that will usually push for an overdesign of the grid (and a waste of valuable routing resources) in an effort to prevent problems that are actually nonexistent.

V. CONCLUSION

A method for power grid verification has been developed, which relies on accurate construction of the worst-case portion of the current sink space by simulation and statistical extrapolation using results from extreme value theory. The method has been shown to accurately predict worst-case voltage drop at all sinks, which was very pessimistically bounded by previous bound-type approaches. It may be used in conjunction with a power grid routing tool in order to prevent grid overdesign and point towards efficient use of routing resources, which will constitute an essential design need for the nanometer-scale generation of VLSI circuits.

REFERENCES

[1]   A. Dharchoudhury, R. Panda, D. Blaauw, R. Vaidyanathan, B. Tutuianu, and D. Bearden, "Design and analysis of power distribution networks in PowerPC microprocessors", *ACM/IEEE Design Automation Conf.*, 1998.

[2]   G. Steele, D. Overhauser, S. Rochel, and S. Hussain, "Full-chip verification methods for DSM power distribution systems", *ACM/IEEE Design Automation Conf.*, 1998.

[3]   S. Chowdhury and J. Barkatullah, "Estimation of maximum currents in MOS IC logic circuits", *IEEE Trans. Computer-Aided Design*, vol. 9, pp. 642-654, 1990.

[4]   H. Kriplani, F. Najm, and I. Hajj, "Pattern independent maximum current estimation in power and ground buses of CMOS VLSI circuits: algorithms, signal correlations and their resolution", *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 998-1012, 1995.

[5]   A. Hill, C. Teng, and S. Kang, "Simulation-based maximum power estimation", *IEEE Int. Symp. Circuits and Systems*, 1996.

[6]   C. Ding, Q. Wu, C. Hsieh, and M. Pedram, "Statistical estimation of the cumulative distribution function for power dissipation in VLSI circuits", *ACM/IEEE Design Automation Conf.*, 1997.

[7]   Q. Wu, Q. Qiu, and M. Pedram, "Estimation of peak power dissipation in VLSI circuits using the limiting distributions of extreme order statistics", *IEEE Trans. Computer-Aided Design*, vol. 20, pp. 942-956, 2001.

[8]   N. Evmorfopoulos, G. Stamoulis, and J. Avaritsiotis, "A Monte Carlo approach for maximum power estimation based on extreme value theory", *IEEE Trans. Computer-Aided Design*, vol. 21, pp. 415-432, 2002.

[9]   J. Galambos, *The Asymptotic Theory of Extreme Order Statistics, 2nd ed.*, Krieger, 1987

[10]   D. Kouroussis and F. Najm, "A static pattern-independent technique for power grid voltage integrity verification", *ACM/IEEE Design Automation Conf.*, 2003.

[11]   H. Qian, S. Nassif, and S. Sapatnekar, "Early-stage power grid analysis for uncertain working modes", *ACM/IEEE Int. Symp. Physical Design*, 2004.

[12]   L. Pillage, R. Rohrer, and C. Visweswariah, *Electronic Circuit and System Simulation Methods*, McGraw-Hill, 1995.

[13]   H. Royden, *Real Analysis, 3rd ed.*, Prentice-Hall, 1988.

[14]   S. Resnick, *Extreme Values, Regular Variation and Point Processes*, Springer, 1987.

[15]   I. Ibragimov and Y. Linnik, *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, 1971.

[16]   H. Joe, *Multivariate Models and Dependence Concepts*, Chapman and Hall, 1996.

[17]   S. Kotz, N. Balakrishnan, and N. Johnson, *Continuous Multivariate Distributions, vol. 1*, Wiley, 2000.

[18]   S. Kotz and S. Nadarajah, *Extreme Value Distributions*, Imperial College Press, 2002.

[19]   H. Kung, F. Luccio, and F. Preparata, "On finding the maxima of a set of vectors", *J. ACM*, vol. 22, pp. 469-476, 1975.

[20]   D. Luenberger, *Linear and Nonlinear Programming, 2nd ed.*, Addison-Wesley, 1984.

[21]   R. Fletcher, *Practical Methods of Optimization, 2nd ed.*, Wiley, 1987.