Voltage-Drop-Constrained Optimization of Power Distribution Network Based on Reliable Maximum Current Estimates

N.E. Evmorfopoulos, D.P. Karampatzakis, and G.I. Stamoulis

University of Thessaly
Dept. of Computer and Communication Engineering
Volos, 38221, GREECE
e-mail: georges@uth.gr

Abstract **– The problem of optimum design of tree-shaped power distribution networks with respect to the voltage drop effect is addressed in this paper. An approach for the width adjustment of the power lines supplying the circuit's major functional blocks is formulated, so that the network occupies the minimum possible area under specific voltage drop constraints at all blocks. The optimization approach is based on precise maximum current estimates derived by statistical means from recent advances in the field of extreme value theory. Experimental tests include the design of power grid for a choice of different topologies and voltage drop tolerances in a typical benchmark circuit.**

I. INTRODUCTION

The massive power distribution networks of modern deep-submicron VLSI circuits are particularly susceptible to a number of reliability problems, the biggest one of which is the well-known voltage drop or IR-drop problem [1]-[2]. This effect characterizes the lowering of the effective voltage level supplied on the active devices of the circuit due to the finite resistance of the power and ground wires, and can have an adverse effect on circuit speed and noise margins (Fig.1), degrading performance (at best) or causing faulty logic signals and circuit malfunction [3]. Traditionally, in order to avoid significant drop of the voltage level the power distribution lines have been made excessively wide to reduce their resistance. However, the voltage drop problem becomes even more pronounced with each new generation of integrated circuits, as the increase in the number of devices and operating frequency combined with the decrease in feature size induce an increase in the currents and the effective resistance respectively, while the decrease in the supply voltage lowers its acceptable drop along the power distribution lines. Therefore simply increasing line width without any restriction cannot be maintained since this would result in a significant occupation of valuable silicon area (which must also accommodate the interconnection lines of the circuit). A method for designing power networks to satisfy certain constraints on IR-drop while occupying the minimum possible silicon area is required, and this became an important research direction lately. A good deal of work has been done previously for this purpose [4]-[9]. Yet, such a design method necessitates the knowledge of the maximum possible currents flowing at any time through the lines of the network, which in all of previous works were considered as given. However, their specific value is hard to obtain since the instantaneous current is a function of the specific pair of input vectors inflicting a transition at the circuit's logic state, the number of which is (as is well known) exponential to the number of primary inputs and prohibitively large to conduct an exhaustive examination. A number of independent approaches for the estimation of maximum currents have appeared [10]-[13], but they are mostly heuristic or over-simplified and cannot provide the accuracy needed for the design of deep-submicron ICs.
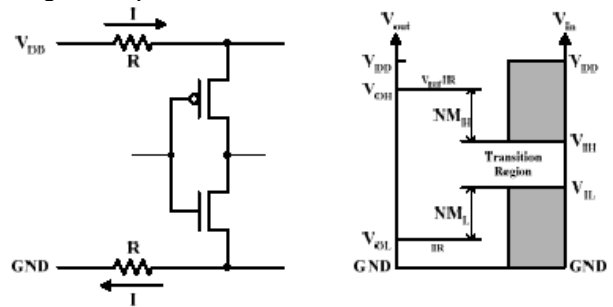

Fig. 1. Voltage drop along the power distribution wires and its impact on circuit noise margins.

In order to ensure the highest possible precision in the assembly of current data, an estimation approach has to rely on detailed circuit simulation. Since, however, it is not possible to simulate the circuit for the entire set of all possible vector pairs, statistical techniques have to be employed as the most viable solution to transfer the burden from a large population of units to a much smaller sample. The design method being proposed in the current paper takes advantage of the most recent research in statistical maximum current estimation from the discipline of *extreme value theory* (EVT) [14], which is the pertinent field of statistics for the estimation of the unknown maximum of a related population from one (or more) of its samples [15]. A few previous techniques based on statistical exploitation of simulation data did appear [16]-[18], but they were either not based on EVT or they did not make efficient use of the theory as was subsequently demonstrated in [14].

The rest of the paper is organized as follows. The next section formulates the problem of power network design as a constrained optimization problem. Section III gives the basic statistical results for the estimation of maximum currents using EVT. Section IV describes the ensuing procedure for the optimum selection of power line widths, and section V presents our experimental results from various case studies on a typical benchmark circuit. Finally, section VI gives the overall concluding remarks.

479

II. PROBLEM FORMULATION

The voltage drop at any time along a branch j of the power distribution network follows the fluctuations of the instantaneous current waveform $I_j(t)$ and is given by:

$$(1) \quad V_j(t) = I_j(t)R_j = I_j(t)\frac{\rho}{t_j}\frac{l_j}{w_j} = I_j(t)R_{sh,j}\frac{l_j}{w_j}$$

where l_j, w_j, t_j, and ρ are the length, width, thickness, and resistivity of branch j respectively, while $R_{sh,j} = \rho/t_j$ is the sheet resistance of the metal layer where the branch lies. This paper will only deal with the optimization of tree-shaped power networks which, although very popular in CAD environments for automatic IC design (e.g. standard-cell ASIC design), have not yet received a satisfactory solution for their effective implementation. A tree topology has the convenient feature that all changes in the widths (and resistances) of the wires do not affect the currents flowing inside the network, the latter being solely determined by the currents drawn from the active devices. A tree also constitutes the minimum area topology under some given constraints for voltage drop, as was explicitly proved in [19]. On the other hand, in the case of general graph networks typically encountered in contemporary custom ICs, the width adjustment procedure would alter the currents flowing inside the network and should be performed in an iterative fashion until a proper convergence criterion is satisfied.

We will apply width modification on the global power lines supplying the major functional blocks (as they are placed during floorplanning), ignoring the local lines reaching the final devices of the circuit. This substantially reduces the number of parameters and constraints involved without any serious compromises, since only those lines carrying large currents and inevitably having the biggest influence in total area are taken into consideration. We will further shrink the number of parameters and constraints by considering the composite straight lines instead of the individual branches in the network, since all collinear branches comprising every one straight line will have common widths eventually.

Taking all the above into account, the objective function describing total network area is given by:

$$(2) \quad A(w_1, w_2, \ldots, w_p) = \sum_{j=1}^{p} l_j w_j$$

where p is the number of global distribution lines in the network, and the optimization is to take place with respect to the width variables w_j (since the lengths l_j are considered as fixed and given by the specific grid topology issuing from the floorplanning scheme). The respective IR-drop constraints must be formulated in such way that the difference $V_{DD} - V(t)$ at all functional blocks stays below a safety threshold voltage V_0 (typically set at 10% of the supply voltage V_{DD} [20]) at all time instants t. This formulation will employ *separate* maximum current values for each new line in the tree hierarchy and not merely use the sum of the maximum currents of its successor lines (or even the end blocks), since the latter are obviously correlated and cannot possibly appear at the same time. Specifically, if $I_{mx,j}$ is the maximum current flowing downstream through line j, the desired IR-drop constraints can be expressed as:

$$(3) \quad \sum_{j \in T_i} I_{mx,j} R_{sh,j} \frac{l_j^*}{w_j} \leq V_0, \quad \forall i = 1,2,\ldots,q_r$$

where q_r is the number of *outermost* functional blocks located at the edges of the power lines (to which we can obviously confine since each of them exhibits the largest voltage drop with respect to all former ones along the same line), and T_i is the *unique* (for tree-shaped networks) path from the main power pad to the outermost block i (see Fig. 2). Notice that the line lengths are denoted as l_j^* since they do not belong entirely to the path T_i but only partially.
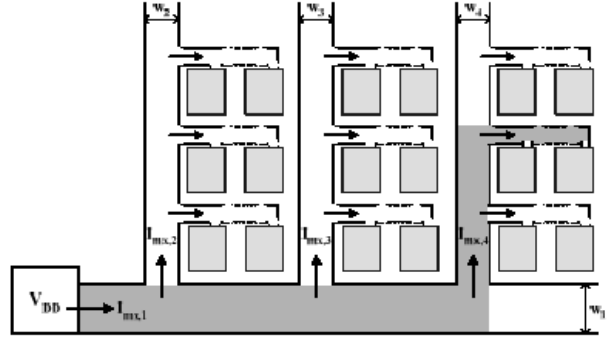

Fig. 2. Tree-shaped power distribution network and voltage drop path from the main power pad to an outermost block.

III. STATISTICAL MAX. CURRENT ESTIMATION

As has already been mentioned above, the difficulty in obtaining the maximum currents $I_{mx,j}$ is that the current waveform $I_j(t)$ flowing through each power line j is a function of the vector pair $(\underline{v}_1, \underline{v}_2)$ generating a state transition at a particular clock edge. If the peak value $I_{P,j}$ of $I_j(t)$ is taken over one clock cycle as a function of $(\underline{v}_1, \underline{v}_2)$, then $I_{mx,j}$ is equal to the maximum value of this *cycle-accurate peak* current among all possible pairs $(\underline{v}_1, \underline{v}_2)$. From a statistical viewpoint, the entire set of vector pairs can establish an initial population of size 4^k (where k is the number of primary inputs - or the number of bits in each input vector) in which the quantity of the cycle-accurate peak current is regarded as a random variable X. Supposing that X is characterized by a cumulative distribution function (cdf) $F(x)$ (and associated density function $f(x) = dF(x)/dx$), which is assumed to be *continuous* and *differentiable* (see Fig. 3 for examples of such distributions for different power lines in a typical benchmark circuit), then the problem of determining the overall maximum current can be cast as a problem of estimating the unknown maximum of a given statistical population with cdf $F(x)$.

The latter quantity, known as the *upper endpoint* ω_F, is formally defined as the least upper bound of the range of values of the variable X (*supporting domain* or *support* of $F(x)$):

(4) $\omega_F = \sup\{x \in \Re : 0 < F(x) < 1\}$

which becomes $\omega_F = F^{-1}(1)$ if X is bounded from above or $\omega_F = +\infty$ in the opposite case.

Let now $\underline{X}_k = [X_{k1}, X_{k2}, \ldots, X_{kn}]^T$ $(k = 1,2,\ldots,m)$ be m samples of size n from the population in hand, where all units are drawn in random so that they form *independent and identically distributed* (iid) random variables with cdf equal to $F(x)$. In our effort to estimate ω_F we seek to construct a new sample $\underline{X}_{mx} = [Z_1, Z_2, \ldots, Z_m]^T$ of the *maxima* units $Z_k = \max(X_{k1}, X_{k2}, \ldots, X_{kn})$ from each initial sample $\underline{X}_k$, which for an adequate size n is known to approach asymptotically a well-defined *parametric* cdf $G(x)$ where ω_F is passed within its parameters a_n, b_n (provided that the parent $F(x)$ is continuous and differentiable). The analytic functional forms of $G(x)$ and a_n, b_n are determined by a fundamental limit theorem of EVT [15], which can be found collectively in [14]. The target estimate $\hat{\omega}_F$ of ω_F can then be extracted by conventional *maximum likelihood* (ML) estimation [21] of parameters a_n, b_n upon the sample $\underline{X}_{mx}$. For the estimation act we will focus on the *dominant* sub-case $G_0(x)$ of $G(x)$, defined as the *Gumbel* probability distribution in [14], which can be shown to characterize the vast majority of practical situations (and largely corresponds to the normal distribution of the central limit theorem being the dominant sub-case within the more general family known as *stable* [22]). This dominance of the Gumbel distribution for maxima has been validated both theoretically [23] as well as experimentally [14],[24]. We cannot use the general cdf $G(x)$ in place of $G_0(x)$ for the estimation of ω_F because, in the vast amount of situations following the latter one, the variance of $\hat{\omega}_F$ (along with the resulting error) would approach *infinity* and thereby render the estimation process inherently inefficient or inaccurate (as was clearly demonstrated in [14]). On the other hand, in the very few cases where the Gumbel assumption is not justified there will only be an overestimation of the actual maximum and a consequent conservative design, which is by all means reasonable compared to a potentially disastrous underdesign. Here, we replicate the important results for this sub-case from [14], referring the reader to that publication for details on their derivation. The upper endpoint estimate is given by:

(5) $\hat{\omega}_F = \hat{a}_n + \dfrac{\hat{b}_n}{1 + n\sqrt{\pi \log n}\left(\mathrm{erf}(\sqrt{\log n}) - 1\right)}$

where $\mathrm{erf}(x) = \dfrac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)\,dt$ is the well-known *error function* which is found in tabular form in many mathematical textbooks (e.g. [25]), and $\hat{a}_n, \hat{b}_n$ are the ML estimates of parameters a_n, b_n as determined by maximization of the subsequent log-likelihood function:

(6)

$\log L(a_n, b_n) = -\sum_{i=1}^{m}\left(\dfrac{Z_i - a_n}{b_n} + \exp\left(-\dfrac{Z_i - a_n}{b_n}\right) + \log b_n\right)$

A confidence interval (corresponding to a confidence level of $(1-\delta)\times 100\%$) can also be constructed for the final estimate $\hat{\omega}_F$, as follows [21]:

(7) $\left|\hat{\omega}_F - \omega_F\right| \le \dfrac{z_{\delta/2}}{\sqrt{m}}\dfrac{\hat{b}_n\sqrt{6}}{\pi}$

$\cdot \sqrt{(\gamma-1)^2 + \dfrac{\pi^2}{6} + \dfrac{2(1-\gamma)}{1 + n\sqrt{\pi \log n}\left(\mathrm{erf}(\sqrt{\log n})-1\right)} + \dfrac{1}{\left(1 + n\sqrt{\pi \log n}\left(\mathrm{erf}(\sqrt{\log n})-1\right)\right)^2}}$

where $z_{\delta/2}$ is the $\delta/2$ quantile point of the standard normal distribution and $\gamma \approx 0.5772\ldots$ is an explicit irrational number known as the *Euler gamma* constant [25].
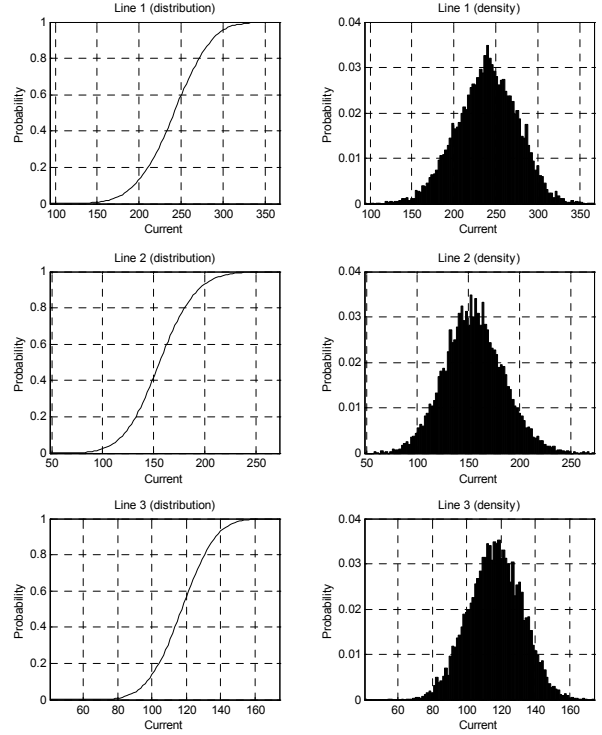
Fig. 3. Empirical distribution function and density function of the cycle-accurate peak current on three major distribution lines (for the c6288 benchmark circuit).

IV. POWER GRID OPTIMIZATION PROCEDURE

The formulated power grid optimization procedure consists of two main parts. The first part is to obtain the estimates of maximum currents along each line of interest, which can be performed using the statistical concepts presented in the previous section. A major assumption here is that the parent cdf $F(x)$ of the cycle-accurate peak current for each line is a continuous and differentiable function. Clearly, this assumption cannot be proved or guaranteed for any possible circuit and any power distribution line. However, it can be most frequently justified in practice for situations where the circuit has a moderate number of primary inputs (which leads to an almost infinite population of vector pairs) and each line has sufficient transistors beneath, so that their combined current (for different vector pairs) is diverse enough to form a continuous and clearly shaped cdf (for example, the empirical cdfs for the three major lines in Fig. 3 can be safely assumed to be continuous and differentiable).

481

In fact, most other researchers have resorted to this particular assumption when dealing with problems of similar nature [12],[16]-[18], and thus we will proceed likewise. Apart from the above assumption, another requirement to ensure the validity of the estimation process is to perform *random* sampling upon the population of vector pairs, so as to obtain current data samples which are comprised of iid units. This requirement is most typically accomplished by means of a computer-based *random number generation* technique. With the above observations in mind, the ensuing procedure for the estimation of maximum current along each power distribution line consists of the following steps:

1. Generate a total of $n \cdot m$ random vector pairs for the circuit under consideration.

2. Simulate circuit (e.g. with Powermill) under all generated pairs and, for each vector pair, record peak current in all relevant power lines within a clock cycle.

For each power line separately then:

3. Arrange the cycle-accurate peak current values in m samples $\underline{X}_k$ of size n each.

4. Construct the sample $\underline{X}_{mx}$ of the maxima from each initial sample $\underline{X}_k$.

5. Perform ML estimation of parameters a_n, b_n on sample $\underline{X}_{mx}$ via maximization of (6).

6. Determine the estimate of overall maximum current from (5), and (optionally) its confidence interval for a desired confidence level $1 - \delta$ from (7).

From the above steps the most subtle one is the ML estimation procedure which again (like our broader problem) involves optimization (maximization here) of an objective function with respect to a set of independent variables. This step can obviously be handled in the same way as the global optimization problem described below in brief. Initial estimates of the parameters a_n, b_n necessary for the algorithmic implementation can be found by equating the first two statistical moments (mean and standard deviation) of the sample $\underline{X}_{mx}$ with those of the population with cdf $G_0(x)$, in the fashion described in [14]. The sizes n and m of the samples participating in the above procedure are selected as $n = 100$ and $m = 100$ respectively (giving a total of $n \cdot m = 10000$ units), which as demonstrated in [14] can lead to a relative estimation error $\varepsilon = |\hat{\omega}_F - \omega_F| / \hat{\omega}_F$ of approximately 5% at a confidence level of 95%.

After determining the maximum current estimates in all lines of interest, we come to the second part of the optimization process which involves application of a proper optimization algorithm to the objective function (2) with the constraints given by (3). Constrained optimization is a very broad field and many algorithms are available which differ from each other in terms of speed and overall efficiency. However, the purpose of this paper is not to elaborate on numerical techniques but to demonstrate the potential of EVT-based maximum current estimation as well as make a smart problem formulation so that it can be approached by any relevant algorithm. That said, a Lagrange-Newton (or SQP) algorithm attempting a direct solution of the *Kuhn-Tucker* optimality conditions was finally adopted, the theory and implementation details of which can be found in [26]-[28].

V. EXPERIMENTAL RESULTS

The circuit selected for the experimental validation of the method was the c6288 (a 16×16 multiplier with 32 inputs, 32 outputs, and 2406 gates), which is the largest and most complex circuit of the ISCAS85 benchmark suite. The circuit was implemented in a contemporary $0.13\ \mu m$ process with 6 copper metal layers and supply voltage $V_{DD} = 1.2\ V$, and was broken up into 5 major functional blocks. These were subsequently arranged in two different floorplanning schemes, which are henceforth referred to as scheme I and scheme II. The estimation of maximum currents at each power line was performed first, according to steps 1 to 6 described in the previous section, and the derived results are reported in Table I. The main information includes the floorplanning scheme in column 1, the number of vector pairs used for estimation in column 2, the line current identifier in column 3, the maximum current in the sample used for estimation (useful to assess the potential of the statistical method) in column 4, the actual maximum current estimate in column 5, the confidence intervals for confidence levels 95%, 99% and 99.99% in columns 6 to 8, and the relative estimation errors for these confidence levels in columns 9 to 11 (derived from columns 6 to 8 by division with column 5), which in the case of the 95% level are close to 5% as expected for the 10000 pairs used for estimation. Note that in the case of scheme I we have additionally conducted estimation for 30000 vector pairs which are found to give around 5% relative error at a higher confidence level of 99.99%.

Table I. Results for the estimated values of maximum current along the three major distribution lines.

Floor-planning scheme	Number of vector pairs	Line current	Sample maximum (mA)	Estimated maximum (mA)	Confidence interval			Relative estimation error		
					95%	99%	99.99%	95%	99%	99.99%
I	10000	$I_{mx,1}$	367.249	471.365	±23.999	±31.540	±47.639	5.1%	6.7%	10.1%
		$I_{mx,2}$	273.032	386.230	±25.428	±33.418	±50.476	6.5%	8.6%	13.1%
		$I_{mx,3}$	174.183	220.125	±10.844	±14.252	±21.526	4.9%	6.5%	9.8%
I	30000	$I_{mx,1}$	376.935	470.862	±13.844	±18.195	±27.482	2.9%	3.9%	5.8%
		$I_{mx,2}$	287.910	379.042	±14.097	±18.527	±27.983	3.7%	4.9%	7.4%
		$I_{mx,3}$	177.075	218.243	±6.033	±7.929	±11.976	2.7%	3.6%	5.5%
II	10000	$I_{mx,1}$	367.249	471.365	±23.999	±31.540	±47.639	5.1%	6.7%	10.1%
		$I_{mx,2}$	275.065	366.538	±24.280	±31.909	±48.197	6.6%	8.7%	13.1%
		$I_{mx,3}$	234.556	302.157	±18.189	±23.905	±36.107	6.0%	7.9%	11.9%
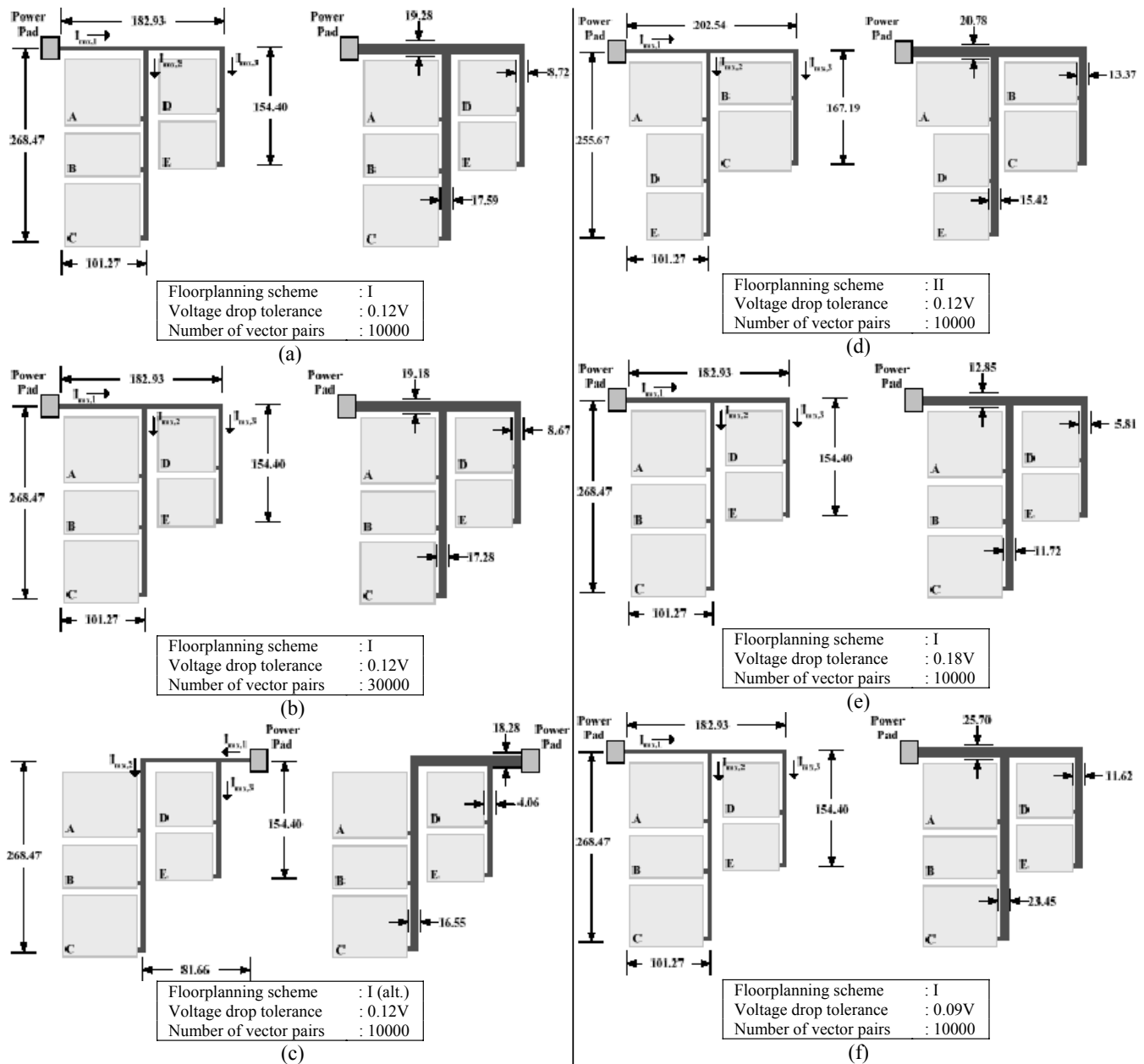
Fig. 4. Various experimental case studies and resulting optimum line widths. All line sizes are expressed in microns and all widths are initially set to $1.2\ \mu m$. The sheet resistance is $0.0143\ \Omega/sq$ and the supply voltage is $1.2\ V$.

Based on the previous maximum current estimates, we have applied the constrained optimization algorithm on six different case studies for the design of power grid of the c6288 circuit. These are collectively shown in Fig. 4. In the left part of each figure the initial network (placed in the uppermost metal layer) is shown with all line lengths explicitly denoted. At the right part the final network following the optimization procedure is demonstrated. The sheet resistance for the copper metal layer is $0.0143\ \Omega/sq$, and initial line widths are $1.2\ \mu m$ in all occasions.

Fig. 4a shows the floorplanning scheme I and the resulting optimum line widths at the typical voltage drop tolerance (10% of V_{DD}, i.e. $V_0 = 0.12\ V$ for $V_{DD} = 1.2\ V$). Fig. 4b

corresponds to the same situation with maximum currents being estimated using 30000 vector pairs, thus giving 5% relative error at a 99.99% confidence level. The resulting line widths are nearly identical, confirming that estimation at a 95% confidence level is fair enough. Fig. 4c shows the same floorplanning scheme supplied by a different (tree-shaped) power network topology. Here, the instantaneous (and maximum) currents are equal but the line lengths are different and this affects the optimum line widths. Fig. 4d illustrates the resulting situation for the floorplanning scheme II, where both line lengths and maximum currents are different. The remaining two cases in Figs. 4e and 4f illustrate the situation for scheme I but with the voltage drop

tolerance being relaxed to 15% of V_{DD} ($V_0 = 0.18\,V$) and tightened to 7.5% of V_{DD} ($V_0 = 0.09\,V$) respectively. We can observe the decrease in optimum line widths in the first case and increase in the second case as was naturally expected.

VI. CONCLUSION

A method for the optimum design of tree-shaped power networks by adjusting global line widths in the presence of IR-drop constraints has been proposed. The method relies on a combination of circuit simulation and statistics in order to obtain the necessary estimates of maximum currents, and makes use of some recent advances in the field of extreme value theory by which the latter can be estimated from an initial sample with arbitrary accuracy and confidence levels. Its potential has been demonstrated by a series of experimental case studies formulated around the physical layout of a standard benchmark circuit. The approach may be easily integrated within the design flow of digital CMOS ICs in order to help prevent power grid overdesign and point towards efficient use of routing resources, which will constitute an essential design need for the next generation of VLSI systems.

REFERENCES

[1] D. Blaauw, R. Panda, and R. Chaudhry, "Design and analysis of power distribution networks", in A. Chandrakasan, W. Bowhill, and F. Fox (eds.), *Design of High-Performance Microprocessor Circuits*, IEEE Press, 2001.

[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] H. Bakoglu, *Circuits, Interconnections, and Packaging for VLSI*, Addison-Wesley, 1990.

[4] S. Chowdhury and M. Breuer, "Minimal area design of power/ground nets having graph topologies", *IEEE Trans. Circuits and Systems*, vol. 34, pp. 1441-1450, 1987.

[5] S. Chowdhury and M. Breuer, "Optimum design of IC power/ground nets subject to reliability constraints", *IEEE Trans. Computer-Aided Design*, vol. 7, pp. 787-796, 1988.

[6] R. Dutta and M. Marek-Sadowska, "Algorithm for wire sizing of power and ground networks in VLSI designs", *J. Circuits, Systems and Computers*, vol. 2, pp. 141-157, 1992.

[7] X. Tan, C. Shi, D. Lungeanu, J. Lee, and L. Yuan, "Reliability-constrained optimization of VLSI power/ground networks via sequence of linear programmings", *ACM/IEEE Design Automation Conf.*, 1999.

[8] S. Boyd, L. Vandenberghe, and A. Gamal, "Design of robust power and ground networks", *ACM/IEEE Int. Symp. Physical Design*, 2001.

[9] X. Tan and C. Shi, "Fast power/ground network optimization based on equivalent circuit modeling", *ACM/IEEE Design Automation Conf.*, 2001.

[10] 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.

[11] 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.

[12] C. Wang and K. Roy, "Maximum power dissipation for CMOS circuits using deterministic and statistical approaches", *IEEE Trans. VLSI Systems*, vol. 6, pp. 134-140, 1998.

[13] Y. Jiang, A. Krstic, and K. Cheng, "Estimation for maximum instantaneous current through supply lines for CMOS circuits", *IEEE Trans. VLSI Systems*, vol. 8, pp. 61-73, 2000.

[14] 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.

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

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

[17] 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.

[18] 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.

[19] K.-H. Erhard and F. Johannes, "Power/ground networks in VLSI: are general graphs better than trees?", *Integration, the VLSI Journal*, vol. 14, pp. 91-109, 1992.

[20] P. Gronowski, W. Bowhill, R. Preston, M. Gowan, and R. Allmon, "High-performance microprocessor design", *IEEE J. Solid State Circuits*, vol. 33, pp. 676-685, 1998.

[21] C. Rao, *Linear Statistical Inference and its Applications, 2nd ed.*, Wiley, 1973.

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

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

[24] E. Castillo, *Extreme Value Theory in Engineering*, Academic Press, 1988.

[25] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover, 1964.

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

[27] P. Gill, W. Murray, and M. Wright, *Practical Optimization*, Academic Press, 1981.

[28] M. Powell, "Variable metric methods for constrained optimization", in A. Bachem, M. Grotschel, and B. Korte (eds.), *Mathematical Programming: The State of the Art*, Springer, 1983.