# Characterization of the Worst-Case Current Waveform Excitations in General RLC-Model Power Grid Analysis


Nestor Evmorfopoulos
University of Thessaly
Dept. of Computer and Communication Engineering
Volos, 38221, Greece
nestevmo@uth.gr

Maria-Aikaterini Rammou
University of Thessaly
Dept. of Computer and Communication Engineering
Volos, 38221, Greece
marammou@uth.gr

George Stamoulis
University of Thessaly
Dept. of Computer and Communication Engineering
Volos, 38221, Greece
georges@uth.gr

John Moondanos
University of Thessaly
Dept. of Computer and Communication Engineering
Volos, 38221, Greece
jmoondan@uth.gr



***Abstract*** – Validating the robustness of power distribution in modern IC design is a crucial but very difficult problem, due to the vast number of possible working modes and the high operating frequencies which necessitate the modeling of power grid as a general RLC network. In this paper we provide a characterization of the worst-case current waveform excitations that produce the maximum voltage drop among all possible working modes of the IC. In addition, we give a practical methodology to estimate these worst-case excitations on the basis of a sample of the excitation space acquired via plain circuit simulation. In the course of characterizing the worst-case excitations we also establish that the voltage drop function for RLC grid models has nonnegative coefficients, which has been an open problem so far.


## I. INTRODUCTION

The power distribution system is one of the most crucial and decisive factors in determining performance and noise immunity in modern nanometer-scale VLSI circuits [1]. In order for the correct and fast operation of the circuit, the power distribution network (or power grid) is expected to be robust, i.e. deliver a reliable voltage level to the circuit functional modules under all possible operating conditions. Verifying the robustness of the power grid, however, is a very difficult problem because the supplied voltage level depends on the currents that the modules themselves (acting as current sinks) draw as excitations from the grid during their operation.

There have been in the recent past some approaches for power grid verification [2]-[4], covering the whole spectrum of power grid models (resistive-only, RC, or full RLC), that operated more or less on the basis of prompting the user to enter constraints that express relationships between sink currents. There are, however, certain problems and limitations with the concept of these so-called "current constraints". First of all, 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 limited 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 probably fall short of the observed reality. Secondly, current constraints have the form of vague upper bounds and thus will only generate a pessimistic upper bound of voltage drop rather than a tight approximation. Lastly, these constraints only involve linear relationships between sink currents and thus give rise to an excitation space which is equivalent to the feasible set of a linear program (LP), despite this space being very nonlinear.

Because of the above problems, the approaches based on current constraints are only really suited for the early stages of the design cycle where the circuit has not been designed yet and the designer only wants to have a very rough idea about the resources that he needs to allocate for the power distribution system. On the other hand, for power grid signoff at the final stages of the design cycle (when the circuit design is fully or almost complete), the designer must be based on detailed simulation of the circuit in order to provide realistic current waveforms that can be drawn as excitations from the power grid. The question is what an appropriate set of current waveforms might be. The possible operating modes of the circuit are really countless (as is the number of possible input binary vectors) and the designer would be in great need of a methodology that provides a set of waveform excitations which can be considered as "worst-case", in the sense that they produce (exactly or approximately) the maximum voltage drop at any node of the grid.

A methodology was proposed in [5] that attempts to do precisely that, but for a resistive-only model of the power grid in which the excitations are not waveforms but static (DC) currents. The methodology relied on the fact that the system matrix of a resistive grid model is inverse-nonnegative (i.e. its inverse has only nonnegative elements) in order to characterize the worst-case subset of the (static) excitation space, and then proceeded to estimate this subset by sampling the excitation space and statistically projecting the sample's own worst-case subset to its expected position in the excitation space. One objective of this paper is to carry out the extension to a more general power grid model that is being stimulated by time-varying waveforms instead of static currents. The extension to an RC grid model is not extremely difficult because it is easy to show that the system matrix is also inverse-nonnegative in this case. The latter fact was the basis of the approach in [6], which nevertheless relied on the concept of current constraints rather than exploiting information from circuit simulation, and is thus targeted at the early stages of the design cycle.

RC grid models, however, are no longer appropriate for the analysis of power distribution in modern circuits with operating frequencies above 1GHz, where the LdI/dt noise from on-chip inductance is extremely important and must be taken into account (see e.g. [7]). Unfortunately, the system of equations that describe the grid behavior with RLC models has completely different structure and properties, and it is still unknown whether the system matrix maintains its inverse-nonnegative property, at least for the part related to voltage drop. A considerable portion (and the second objective) of this paper is devoted to demonstrate that this is true, a fact which can be of significant interest in its own right and might have other and more far reaching consequences (e.g. it establishes a kind of monotonicity property for the power grid, even in the case of general RLC models where this was not thought possible before).

The rest of the paper is organized as follows. The next section describes the RLC model of the power grid and reviews its analysis in the time domain. Section III determines the worst-case waveform excitations in the case of RLC grid models, and proves along that even for such models the inversion of the system matrix will result in nonnegative coefficients in the voltage drop function. Section IV builds on the characterization of the worst-case excitations in order to formulate a practical methodology for power grid verification. Section V presents our experimental results from various case studies comprising of test grids and typical benchmark circuits, and finally section VI gives the overall concluding remarks.

## II. MODEL AND TRANSIENT ANALYSIS OF THE POWER GRID

As already stated, we will be concerned with the full RLC model of the power grid. We assume that the grid is composed of $b$ branches and $n+p$ nodes, of which $p$ nodes are connected to the external power supply via power pads, and the remaining $n$ nodes are divided to $m$ sink nodes (with current sources to an external ground node) and $n-m$ internal nodes. Especially for networks representing power grids we typically model each wire segment (between two contacts) as a resistance in series with an inductance, with capacitances to ground at the two contact nodes (Fig. 1). Thus, in the analysis that follows we will consider the $b$ branches of the grid as *composite* resistive-inductive (R-L) branches.

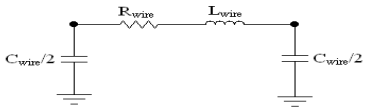


**Fig. 1.** Typical model of a wire segment in power grids.

We next review briefly the Modified Nodal Analysis (MNA) framework for the analysis of RLC power grids [8], in which inductor currents constitute additional variables alongside node voltages. The Kirchhoff's current and voltage laws for the linear network representing the power grid are:

(1) $\mathbf{A}_{rl}\mathbf{i}_b(t)+\mathbf{i}_c(t)=\mathbf{e}_n(t)$

(2) $\mathbf{A}_{rl}^T\mathbf{v}_n(t)=\mathbf{v}_b(t)$

where $\mathbf{A}_{rl}$ is the $n\times b$ incidence matrix of the directed composite R-L branches (with elements $a_{ij}=\pm1$ or $a_{ij}=0$ depending on whether branch $j$ leaves/enters or is not incident with node $i$), $\mathbf{v}_n(t)$, $\mathbf{v}_b(t)$, and $\mathbf{i}_b(t)$ are the $n\times1$, $b\times1$, and $b\times1$ vectors of node voltages, branch voltages, and branch currents respectively, $\mathbf{i}_c(t)$ is a $n\times1$ vector of currents of the additional capacitive branches which appear at the $n$ nodes, and $\mathbf{e}_n(t)$ is a $n\times1$ vector of excitations from independent sources (either current or voltage ones) at the nodes.

The current-voltage relationships of the $n$ capacitive branches and the $b$ composite R-L branches are as follows:

(3) $\mathbf{i}_c(t)=\mathbf{C}_n\dot{\mathbf{v}}_n(t)$

(4) $\mathbf{v}_b(t)=\mathbf{R}_b\mathbf{i}_b(t)+\mathbf{L}_b\dot{\mathbf{i}}_b(t)$

where $\dot{\mathbf{v}}_n(t)$ and $\dot{\mathbf{i}}_b(t)$ are the time derivatives of vectors $\mathbf{v}_n(t)$ and $\mathbf{i}_b(t)$ respectively, $\mathbf{C}_n$ is a $n\times n$ diagonal matrix of the node capacitances, and $\mathbf{R}_b$, $\mathbf{L}_b$ are $b\times b$ matrices of the resistances and inductances of the composite R-L branches. The matrix $\mathbf{R}_b$ is a diagonal matrix, while $\mathbf{L}_b$ is either diagonal if there are only self-inductances at the branches, or a full matrix if there are also mutual inductances between branches. We assume that each R-L branch has nonzero self-inductance and each node has nonzero capacitance, so that the matrices $\mathbf{L}_b$ and $\mathbf{C}_n$ are nonsingular (invertible).

In MNA we replace (3) into (1) and (4) into (2), and combine the two resulting systems of differential equations into the following system:

(5) $\widetilde{\mathbf{G}}\mathbf{x}(t)+\widetilde{\mathbf{C}}\dot{\mathbf{x}}(t)=\mathbf{e}(t)$

where $\widetilde{\mathbf{G}}=\begin{bmatrix}\mathbf{0} & \mathbf{A}_{rl}\\ -\mathbf{A}_{rl}^T & \mathbf{R}_b\end{bmatrix}$, $\widetilde{\mathbf{C}}=\begin{bmatrix}\mathbf{C}_n & \mathbf{0}\\ \mathbf{0} & \mathbf{L}_b\end{bmatrix}$, $\mathbf{x}(t)=\begin{bmatrix}\mathbf{v}_n(t)\\ \mathbf{i}_b(t)\end{bmatrix}$, $\mathbf{e}(t)=\begin{bmatrix}\mathbf{e}_n(t)\\ \mathbf{0}\end{bmatrix}$

In the above system $\mathbf{v}_n(t)$ is the vector of node voltages, but the system is easily re-expressed with respect to the *voltage drops* at the nodes by omitting the independent voltage sources in the excitation vector $\mathbf{e}_n(t)$ and reversing the sign of the current sources (from – to +). From now on we will denote $\mathbf{v}_n(t)$ as the vector of voltage drops at the nodes.

By the Backward Euler approximation for a fixed time step $h$, we can replace $\dot{\mathbf{x}}(t)$ by its finite difference formula $\dot{\mathbf{x}}(t)\approx[\mathbf{x}(t)-\mathbf{x}(t-h)]/h$ in (5) and obtain $(\widetilde{\mathbf{G}}+\widetilde{\mathbf{C}}/h)\mathbf{x}(t)=(\widetilde{\mathbf{C}}/h)\mathbf{x}(t-h)+\mathbf{e}(t)$ for $t=kh$, $k=1,2,...$, or:

(6) $\mathbf{x}(kh)=(\widetilde{\mathbf{G}}+\widetilde{\mathbf{C}}/h)^{-1}\mathbf{e}(kh)+(\widetilde{\mathbf{G}}+\widetilde{\mathbf{C}}/h)^{-1}(\widetilde{\mathbf{C}}/h)\mathbf{x}((k-1)h)$
$\equiv\mathbf{B_1}\mathbf{e}(kh)+\mathbf{B}\mathbf{x}((k-1)h)$

where $\mathbf{B_1}=(\widetilde{\mathbf{G}}+\widetilde{\mathbf{C}}/h)^{-1}$ and $\mathbf{B}=(\widetilde{\mathbf{G}}+\widetilde{\mathbf{C}}/h)^{-1}(\widetilde{\mathbf{C}}/h)=\mathbf{B_1}(\widetilde{\mathbf{C}}/h)$.

The latter recursive relation is used to calculate all node voltage drops and all branch currents at a particular time instant $t=kh$, $k=1,2,...$, based on the voltage drops and branch currents at the previous instant $t=(k-1)h$.

## III. DETERMINATION OF THE WORST-CASE CURRENT EXCITATIONS

### *A. Instantaneous and cycle-mean voltage drop as function of the excitation waveforms*

By successive substitutions of the recursive expressions for $\mathbf{x}((k-1)h)$, $\mathbf{x}((k-2)h)$, … , $\mathbf{x}(h)$ into (6) we obtain:

$$\mathbf{x}(kh)=\mathbf{B_1}\mathbf{e}(kh)+\mathbf{B}\mathbf{B_1}\mathbf{e}((k-1)h)+\cdots+\mathbf{B}^{k-1}\mathbf{B_1}\mathbf{e}(h)+\mathbf{B}^k\mathbf{x}(0)$$
$$=\sum_{j=0}^{k-1}\mathbf{B}^j\mathbf{B_1}\mathbf{e}((k-j)h)+\mathbf{B}^k\mathbf{x}(0),\ k=1,2,...$$

However, both for timing and noise purposes, the peak instantaneous voltage drop is not as important as the integral of voltage drop (or the mean voltage drop) within a specified time interval $T=Nh$, which may be equal or smaller than the clock period – e.g. an interval of high activity within the clock cycle (a large instantaneous voltage drop will not severely affect timing but a large cumulative voltage drop over a time interval will). The mean vector of voltage drops and branch currents within such an interval is:

$$\bar{\mathbf{x}}=\frac{1}{N}[\mathbf{x}(h)+\mathbf{x}(2h)+\cdots+\mathbf{x}(Nh)]$$
$$=\frac{1}{N}[\mathbf{B_1}\mathbf{e}(Nh)+(\mathbf{I}+\mathbf{B})\mathbf{B_1}\mathbf{e}((N-1)h)+\cdots+(\mathbf{I}+\mathbf{B}+\cdots+\mathbf{B}^{N-1})\mathbf{B_1}\mathbf{e}(h)+\mathbf{c}]$$

where $\mathbf{c}\equiv(\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N)\mathbf{x}(0)$.

Since $\mathbf{B_1}=h\mathbf{B}\widetilde{\mathbf{C}}^{-1}$ (where $\widetilde{\mathbf{C}}^{-1}=\begin{bmatrix}\mathbf{C}_n^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{L}_b^{-1}\end{bmatrix}$) we have $\mathbf{B_1}\mathbf{e}(kh)=h\mathbf{B}\mathbf{e}_C(kh)$ (where $\mathbf{e}_C(kh)\equiv\begin{bmatrix}\mathbf{C}_n^{-1}\mathbf{e}_n(kh)\\ \mathbf{0}\end{bmatrix}$) for each $k=1,2,...,N$, so

$$\bar{\mathbf{x}}=\frac{h}{N}\left[\mathbf{B}\mathbf{e}_C(Nh)+(\mathbf{B}+\mathbf{B}^2)\mathbf{e}_C((N-1)h)+\cdots+(\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N)\mathbf{e}_C(h)\right]+\frac{\mathbf{c}}{N}$$

By denoting the upper-left $n\times n$ blocks of the matrices $\mathbf{B}$, $\mathbf{B}+\mathbf{B}^2$,…, $\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N$ as $(\mathbf{B})_{n\times n}$, $(\mathbf{B}+\mathbf{B}^2)_{n\times n}$,…, $(\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N)_{n\times n}$, we have for the block $\bar{\mathbf{v}}_n$ of voltage drops within the vector $\bar{\mathbf{x}}$:

(7) $$\bar{\mathbf{v}}_n=\frac{h}{N}\left[(\mathbf{B})_{n\times n}\mathbf{C}_n^{-1}\mathbf{e}_n(Nh)+(\mathbf{B}+\mathbf{B}^2)_{n\times n}\mathbf{C}_n^{-1}\mathbf{e}_n((N-1)h)+\right.$$
$$\left.+\cdots+(\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N)_{n\times n}\mathbf{C}_n^{-1}\mathbf{e}_n(h)\right]+\mathbf{c}_{n\times1}/N$$
$$\equiv\frac{h}{N}\left[(\mathbf{B})_{n\times n}\mathbf{C}_n^{-1}\mathbf{e}_N+(\mathbf{B}+\mathbf{B}^2)_{n\times n}\mathbf{C}_n^{-1}\mathbf{e}_{N-1}+\cdots+(\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N)_{n\times n}\mathbf{C}_n^{-1}\mathbf{e}_1\right]+\frac{\mathbf{c}_{n\times1}}{N}$$

where we have written the vectors of excitations $\mathbf{e}_n(Nh)$, $\mathbf{e}_n((N-1)h)$,…, $\mathbf{e}_n(h)$ at time instants $t=kh$, $k=1,2,...,N$ as $\mathbf{e}_N$, $\mathbf{e}_{N-1}$,…, $\mathbf{e}_1$ since we will treat them as variables hereafter. Thus we have arrived at the result that the mean voltage drop at each node $i=1,2,...,n$ (i.e. each component $\bar{v}_i$, $i=1,2,...,n$ within $\bar{\mathbf{v}}_n$) is a linear function (or a linear affine function if the constant vector $\mathbf{c}$ is not $\mathbf{0}$) of the super-vector of $M\equiv mN$ dimensions:

$$\mathbf{y}\equiv(e_{N,1},e_{N,2},\ldots,e_{N,m},e_{N-1,1},e_{N-1,2},\ldots,e_{N-1,m},\ldots,e_{1,1},e_{1,2},\ldots,e_{1,m})$$

which consists of the (discretized) current waveform excitations at the $m$ sink nodes (we remind that in each $n\times1$ vector $\mathbf{e}_N$, $\mathbf{e}_{N-1}$,…, $\mathbf{e}_1$ only $m$ of the $n$ components that correspond to sink nodes are nonzero).

### *B. Maximizers of a linear (or linear affine) function with nonnegative coefficients*

The variable vector $\mathbf{y}$ in every function $\bar{v}_i\equiv\bar{v}_i(\mathbf{y})$, $i=1,2,...,n$ does not attain all values in the $M$-dimensional space $\Re^M\equiv\Re^{mN}$. Instead, the excitation values at each sink $j=1,2,...,m$ and at each time instant $t=kh$, $k=1,2,...,N$ depend on the specific clock cycle where they are considered, i.e. on the pair of binary vectors $\{\mathbf{b}_p,\mathbf{b}_n\}$ being applied on the digital circuit before and after the clock edge. This means that the vector $\mathbf{y}$ is actually a vector-valued function $\mathbf{y}\equiv\mathbf{y}(\{\mathbf{b}_p,\mathbf{b}_n\})$ whose range of values $D\subset\Re^M$ constitutes the *domain* on which the linear function $\bar{v}_i(\mathbf{y})$ is defined. This domain $D$ (henceforth referred to as the "excitation space") is obviously bounded (since the drawn currents at every time instant are all finite) and closed (since it contains its boundary points), which means that it is a *compact* set of $\Re^M$. Due to the well-known Weierstrass theorem [9], a continuous function $f(\mathbf{y})$ defined on a compact set $D\subset\Re^M$ (i.e. $f:D\to\Re$) always attains a maximum at some point $\mathbf{y}^*\in D$, which is called *maximizing point* or *maximizer* of $f(\mathbf{y})$. We seek to find the

candidate maximizers $\mathbf{y}^*$ for the linear functions $\bar{v}_i(\mathbf{y})$, $i=1,2,...,n$ which will constitute the worst-case excitations. It is known that for R or RC grid models the linear functions $\bar{v}_i(\mathbf{y})$, $i=1,2,...,n$ have *nonnegative* coefficients in all components of the vector **y** (due to the matrix $\widetilde{\mathbf{G}}+\widetilde{\mathbf{C}}/h$ being inverse-nonnegative in the absence of inductive elements – for general RLC models see next section). For this class of linear functions the following hold with respect to the maximizing points:

**Definition 1.** A point $\mathbf{y}\in D$ is called a *maximal* (or *noninferior*) point of the set $D\subset\Re^M$ if there does not exist a $\mathbf{y}'\in D$ such that $\mathbf{y}'\geq\mathbf{y}$ (component-wise) with at least one component $i=1,\ldots,M$ being $y_i'>y_i$.

**Theorem 1.** [5] Let $f(\mathbf{y})=\sum_{i=1}^{M}a_i y_i=\mathbf{a}^T\cdot\mathbf{y}$ be a linear (or linear affine) function with nonnegative coefficient vector **a** (i.e. $\mathbf{a}\geq\mathbf{0}$ component-wise) which is defined on a compact set $D\subset\Re^M$. If $\mathbf{y}^*\in D$ is a maximizer of $f(\mathbf{y})$ [i.e. $f(\mathbf{y}^*)=\max_{\mathbf{y}\in D}f(\mathbf{y})$], then $\mathbf{y}^*$ is a maximal point of *D*.

The above theorem effectively states that the candidate maximizers of a linear function $f(\mathbf{y})=\mathbf{a}^T\cdot\mathbf{y}$ with $\mathbf{a}\geq\mathbf{0}$ which is defined on a closed and bounded set $D\subset\Re^M$ are the maximal points of its domain *D* (Fig. 2).

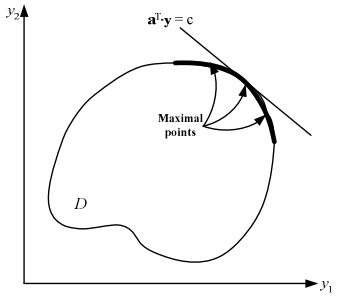


**Fig. 2.** Maximal points of a compact set $D\subset\Re^M$ and maximization of a linear function $f(\mathbf{y})=\mathbf{a}^T\cdot\mathbf{y}$ over it (only the level line of $f(\mathbf{y})$ at the maximizing point $\mathbf{y}^*$ is shown in the figure – the actual graph of $f(\mathbf{y})$ is a plane in 3 dimensions).

### *C. Nonnegativity of the coefficients of the voltage drop function in the case of general RLC grid models*

In contrast to the case of R or RC grid models, it is not at all obvious whether every mean voltage drop function $\bar{v}_i\equiv\bar{v}_i(\mathbf{y})$ in (7) has nonnegative coefficients, in order for the maximal waveform points to still represent the worst-case excitations.

At first, since all capacitances in the diagonal matrix $\mathbf{C}_n$ are positive, the matrix $\mathbf{C}_n^{-1}$ is also positive diagonal and thus it is sufficient to examine the $n\times n$ upper-left blocks $(\mathbf{B})_{n\times n}$, $(\mathbf{B}+\mathbf{B}^2)_{n\times n}$,..., $(\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N)_{n\times n}$ of the matrices **B**, $\mathbf{B}+\mathbf{B}^2$,..., $\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N$. The matrix **B** can be written:

$$(8)\qquad \mathbf{B}=(\widetilde{\mathbf{G}}+\widetilde{\mathbf{C}}/h)^{-1}(\widetilde{\mathbf{C}}/h)=(h\widetilde{\mathbf{C}}^{-1}(\widetilde{\mathbf{G}}+\widetilde{\mathbf{C}}/h))^{-1}=(h\widetilde{\mathbf{C}}^{-1}\widetilde{\mathbf{G}}+\mathbf{I})^{-1}$$
$$=\begin{bmatrix}\mathbf{I}_n & \mathbf{A}_C\\ -\mathbf{A}_L & \mathbf{I}_b+\mathbf{R}_L\end{bmatrix}^{-1}$$

where $\mathbf{A}_C=h\mathbf{C}_n^{-1}\mathbf{A}_{rl}$, $\mathbf{A}_L=h\mathbf{L}_b^{-1}\mathbf{A}_{rl}^T$, $\mathbf{R}_L=h\mathbf{L}_b^{-1}\mathbf{R}_b$

and $\mathbf{I}$, $\mathbf{I}_n$, and $\mathbf{I}_b$ are the $(n+b)\times(n+b)$, $n\times n$, and $b\times b$ identity matrices respectively.

By performing block matrix inversion [10] in (8) we have:

$$(9)\qquad \mathbf{B}=\begin{bmatrix}\mathbf{Q} & -\mathbf{Q}\mathbf{A}_C\mathbf{U}\\ \mathbf{U}\mathbf{A}_L\mathbf{Q} & \mathbf{U}-\mathbf{U}\mathbf{A}_L\mathbf{Q}\mathbf{A}_C\mathbf{U}\end{bmatrix}$$

where $\mathbf{U}=(\mathbf{I}_b+\mathbf{R}_L)^{-1}$ and $\mathbf{Q}=(\mathbf{I}_n+\mathbf{A}_C\mathbf{U}\mathbf{A}_L)^{-1}$

(just multiply (9) by $\begin{bmatrix}\mathbf{I}_n & \mathbf{A}_C\\ -\mathbf{A}_L & \mathbf{I}_b+\mathbf{R}_L\end{bmatrix}$ and verify that their product is the $(n+b)\times(n+b)$ identity matrix **I**).

For the remainder of the section we will assume only self-inductances in the power grid model – which is the most common case for present frequency ranges – so that the matrix $\mathbf{L}_b$ (and $\mathbf{L}_b^{-1}$) is positive diagonal (we will give, however, an analysis of mutual inductances in the next section). Under this condition we can readily show that the upper-left $n\times n$ block $(\mathbf{B})_{n\times n}=\mathbf{Q}$ of **B** has only nonnegative elements (i.e. $\mathbf{Q}\geq\mathbf{0}$ element-wise). Indeed the matrix $\mathbf{U}h\mathbf{L}_b^{-1}=(\mathbf{I}_b+h\mathbf{L}_b^{-1}\mathbf{R}_b)^{-1}h\mathbf{L}_b^{-1}$ will also be positive diagonal, and the product $\mathbf{V}=\mathbf{A}_{rl}(\mathbf{U}h\mathbf{L}_b^{-1})\mathbf{A}_{rl}^T$ where $\mathbf{A}_{rl}$ is an incidence matrix and $\mathbf{U}h\mathbf{L}_b^{-1}$ is a positive diagonal matrix is well known to have the following properties [8]: (i) positive diagonal elements ($v_{ii}>0$, $i=1,\ldots,n$), (ii) nonpositive off-diagonal elements ($v_{ij}\leq 0$, $i,j=1,\ldots,n$, $i\neq j$), (iii) symmetry ($v_{ij}=v_{ji}$, $i,j=1,\ldots,n$) (iv) diagonal dominance – in both rows and columns – i.e. $|v_{ii}|\geq\sum_{\substack{j=1\\ j\neq i}}^{n}|v_{ij}|$, $\forall i=1,\ldots,n$.

It is easy now to show (see [11] for a proof) that if **V** is a $n\times n$ symmetric diagonally dominant matrix with positive diagonal elements and nonpositive off-diagonal elements, and **C**, **D** are $n\times n$ positive and nonnegative diagonal matrices respectively (i.e. $c_i>0$ and $d_i\geq 0$, $i=1,\ldots,n$), then $\mathbf{W}=\mathbf{D}+\mathbf{C}\mathbf{V}$ is a *row* diagonally dominant matrix with positive diagonal elements and nonpositive off-diagonal elements. This result applies to the matrix $\mathbf{I}_n+\mathbf{A}_C\mathbf{U}\mathbf{A}_L=\mathbf{I}_n+h\mathbf{C}_n^{-1}\mathbf{A}_{rl}\mathbf{U}h\mathbf{L}_b^{-1}\mathbf{A}_{rl}^T$, which is hence row diagonally dominant with positive diagonal elements and nonpositive off-diagonal elements. These properties are sufficient for a matrix to be classified as an *M-matrix* [12], which by definition is inverse-nonnegative, i.e. $\mathbf{Q}=(\mathbf{I}_n+\mathbf{A}_C\mathbf{U}\mathbf{A}_L)^{-1}\geq\mathbf{0}$.

For the remaining blocks $(\mathbf{B}+\mathbf{B}^2)_{n\times n}$,..., $(\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N)_{n\times n}$ we have that $\mathbf{B}\equiv\mathbf{S}_1$, $\mathbf{B}+\mathbf{B}^2\equiv\mathbf{S}_2$,..., $\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N\equiv\mathbf{S}_N$ constitute partial sums of the series $\mathbf{B}+\mathbf{B}^2+\cdots=\mathbf{B}(\mathbf{I}+\mathbf{B}+\cdots)=\mathbf{B}\sum_{N=0}^{\infty}\mathbf{B}^N$ which converges to the matrix $\mathbf{B}(\mathbf{I}-\mathbf{B})^{-1}\equiv\mathbf{S}$, on condition that $\lim_{N\to\infty}\mathbf{B}^N=\mathbf{0}$ [13]. This means that the sequence of the *ij*th elements $s_{ij}^{(N)}$ of the partial sums $\mathbf{S}_N$ converges to the *ij*th element $s_{ij}$ of **S** for every $i,j=1,\ldots,n+b$ (i.e. $\lim_{N\to\infty}s_{ij}^{(N)}=s_{ij}$, $\forall i,j=1,\ldots,n+b$), and this of course is true for the upper-left $n\times n$ elements $s_{ij}^{(N)}$ and $s_{ij}$, $i,j=1,\ldots,n$. By writing the limit matrix **S** via (8) as:

$$(10)\qquad \mathbf{S}=\mathbf{B}(\mathbf{I}-\mathbf{B})^{-1}=(\mathbf{B}^{-1}-\mathbf{I})^{-1}=\begin{bmatrix}\mathbf{0} & \mathbf{A}_C\\ -\mathbf{A}_L & \mathbf{R}_L\end{bmatrix}^{-1}$$

we find by block matrix inversion that its $n\times n$ upper-left block is:

$$(11)\qquad (\mathbf{S})_{n\times n}=(\mathbf{B}(\mathbf{I}-\mathbf{B})^{-1})_{n\times n}=(\mathbf{A}_C\mathbf{R}_L^{-1}\mathbf{A}_L)^{-1}=(h\mathbf{C}_n^{-1}\mathbf{A}_{rl}\mathbf{R}_b^{-1}\mathbf{A}_{rl}^T)^{-1}$$

The matrix $h\mathbf{C}_n^{-1}\mathbf{A}_{rl}\mathbf{R}_b^{-1}\mathbf{A}_{rl}^T$ is row diagonally dominant with positive diagonal elements and nonpositive off-diagonal elements, and is thus an M-matrix which is inverse-nonnegative, i.e. $(\mathbf{S})_{n\times n}\geq\mathbf{0}$.

Overall, we have a series of matrices which starts off by a first term $\mathbf{S}_1=\mathbf{B}$ with a nonnegative upper-left $n\times n$ block, accepts additive terms $\mathbf{B}^2$, $\mathbf{B}^3$,... (to form the intermediate partial sums) which have gradually smaller elements than **B** (due to $\lim_{N\to\infty}\mathbf{B}^N=\mathbf{0}$, which is equivalent to $\lim_{N\to\infty}\|\mathbf{B}^N\|=0$ – see in a moment about this), and converges to a limit $\mathbf{S}=\mathbf{B}(\mathbf{I}-\mathbf{B})^{-1}$ with also a nonnegative upper-left $n\times n$ block. This ensures us that all intermediate partial sums $\mathbf{B}+\mathbf{B}^2$, $\mathbf{B}+\mathbf{B}^2+\mathbf{B}^3$,... will have nonnegative upper-left $n\times n$ blocks, i.e. $(\mathbf{B}+\mathbf{B}^2)_{n\times n}\geq\mathbf{0}$, $(\mathbf{B}+\mathbf{B}^2+\mathbf{B}^3)_{n\times n}\geq\mathbf{0}$,....

As a practical example, consider a grid with $n=6$ nodes and $b=8$ branches, and the following node capacitance, branch resistance, branch inductance, and incidence matrices:

$\mathbf{C}_n=\mathrm{diag}[52.50\quad 80.00\quad 52.50\quad 52.50\quad 70.00\quad 52.50]\,fF$

$$\mathbf{R}_b = \mathrm{diag}[17.5 \quad 17.5 \quad 17.5 \quad 17.5 \quad 35.0 \quad 35.0 \quad 35.0 \quad 50.0]\,\Omega$$

$$\mathbf{L}_b = \mathrm{diag}[17.53 \quad 17.53 \quad 17.53 \quad 17.53 \quad 35.07 \quad 35.07 \quad 35.07 \quad 10.00]\,pH$$

$$\mathbf{A}_{rl} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \end{bmatrix}$$

For this particular case the upper-left $6\times 6$ blocks of the first three partial sums, as well as the limit of the series are:

$$(\mathbf{B})_{n\times n} = \begin{bmatrix} 0.0291 & 0.0335 & 0.0222 & 0.0252 & 0.0314 & 0.0229 \\ 0.0220 & 0.0341 & 0.0220 & 0.0216 & 0.0289 & 0.0216 \\ 0.0222 & 0.0335 & 0.0291 & 0.0229 & 0.0314 & 0.0252 \\ 0.0252 & 0.0330 & 0.0229 & 0.0329 & 0.0370 & 0.0260 \\ 0.0236 & 0.0330 & 0.0236 & 0.0277 & 0.0401 & 0.0277 \\ 0.0229 & 0.0330 & 0.0252 & 0.0260 & 0.0370 & 0.0329 \end{bmatrix}$$

$$(\mathbf{B}+\mathbf{B}^2)_{n\times n} = \begin{bmatrix} 0.0329 & 0.0389 & 0.0261 & 0.0294 & 0.0370 & 0.0271 \\ 0.0255 & 0.0390 & 0.0255 & 0.0255 & 0.0340 & 0.0255 \\ 0.0261 & 0.0389 & 0.0329 & 0.0271 & 0.0370 & 0.0294 \\ 0.0294 & 0.0388 & 0.0271 & 0.0374 & 0.0430 & 0.0305 \\ 0.0277 & 0.0388 & 0.0277 & 0.0322 & 0.0461 & 0.0322 \\ 0.0271 & 0.0388 & 0.0294 & 0.0305 & 0.0430 & 0.0374 \end{bmatrix}$$

$$(\mathbf{B}+\mathbf{B}^2+\mathbf{B}^3)_{n\times n} = \begin{bmatrix} 0.0336 & 0.0398 & 0.0267 & 0.0301 & 0.0379 & 0.0278 \\ 0.0261 & 0.0398 & 0.0261 & 0.0261 & 0.0348 & 0.0261 \\ 0.0267 & 0.0398 & 0.0336 & 0.0278 & 0.0379 & 0.0301 \\ 0.0301 & 0.0398 & 0.0278 & 0.0382 & 0.0440 & 0.0313 \\ 0.0284 & 0.0398 & 0.0284 & 0.0330 & 0.0471 & 0.0330 \\ 0.0278 & 0.0398 & 0.0301 & 0.0313 & 0.0440 & 0.0382 \end{bmatrix}$$

$$(\mathbf{B}(\mathbf{I}-\mathbf{B})^{-1})_{n\times n} = \begin{bmatrix} 0.0337 & 0.0400 & 0.0268 & 0.0303 & 0.0381 & 0.0280 \\ 0.0262 & 0.0400 & 0.0262 & 0.0262 & 0.0350 & 0.0262 \\ 0.0268 & 0.0400 & 0.0337 & 0.0280 & 0.0381 & 0.0303 \\ 0.0303 & 0.0400 & 0.0280 & 0.0383 & 0.0442 & 0.0314 \\ 0.0285 & 0.0400 & 0.0285 & 0.0331 & 0.0472 & 0.0331 \\ 0.0280 & 0.0400 & 0.0303 & 0.0314 & 0.0442 & 0.0383 \end{bmatrix}$$

which are all nonnegative as expected (observe the remarkably fast convergence of the series to its limit after only three partial sums).

In order, now, to prove that $\lim_{N\to\infty}\mathbf{B}^N = \mathbf{0}$ (which is necessary for the above to hold) we first establish that $\mathbf{B}^N$ has the general form shown at the bottom of the page (just multiply $\mathbf{B}^N$ by $\mathbf{B}$ of (9) and verify that the resulting $\mathbf{B}^{N+1}$ has the same form). Therefore, in order to prove that $\lim_{N\to\infty}\mathbf{B}^N = \mathbf{0}$ we just need to prove that $\lim_{N\to\infty}\mathbf{U}^N = \mathbf{0}$ and $\lim_{N\to\infty}\mathbf{Q}^N = \mathbf{0}$. For this we will need the following theorem:

**Theorem 2.** If $\mathbf{V} = [v_{ij}]$ is a $n\times n$ – row or column – diagonally dominant matrix with positive diagonal elements, then for the matrix $\mathbf{W} = (\mathbf{I}+\mathbf{V})^{-1}$ it holds $\lim_{N\to\infty}\mathbf{W}^N = \mathbf{0}$.

***Proof.*** It is well known [13] that $\lim_{N\to\infty}\mathbf{W}^N = \mathbf{0}$ if and only if $\rho(\mathbf{W}) = \max_{1\le i\le n}|\lambda_i(\mathbf{W})| < 1$, where $\lambda_i(\mathbf{W})$, $i = 1,2,...,n$ are the – not necessarily distinct – eigenvalues of $\mathbf{W}$ and $\rho(\mathbf{W})$ is the largest of their magnitudes (called the *spectral radius* of $\mathbf{W}$). Also, for $\mathbf{W} = (\mathbf{I}+\mathbf{V})^{-1}$ it is true that $\lambda_i(\mathbf{W}) = \dfrac{1}{1+\lambda_i(\mathbf{V})}$, $i = 1,2,...,n$ [10]. According, now, to the Gershgorin circle theorem [13], all eigenvalues of a matrix $\mathbf{V}$ are located in the set of $n$ disks in the complex plane defined by $\left\{ z : |z - v_{ii}| \le \sum_{\substack{j=1 \\ j\ne i}}^{n} |v_{ij}| \right\}$, $i = 1,2,...,n$ (i.e. the $n$ disks centered at $v_{ii}$ and having radius $\sum_{\substack{j=1 \\ j\ne i}}^{n} |v_{ij}|$, $i = 1,2,...,n$). If the matrix $\mathbf{V}$ is row diagonally dominant with positive diagonal elements, then $v_{ii} \ge \sum_{\substack{j=1 \\ j\ne i}}^{n} |v_{ij}|$, $\forall i = 1,2,...,n$, which means that all Gershgorin disks lie entirely in the positive real semi-plane and all eigenvalues of $\mathbf{V}$ have positive real parts, i.e. $\mathrm{Re}\,\lambda_i(\mathbf{V}) > 0$. This gives $|1+\lambda_i(\mathbf{V})| = \sqrt{(1+\mathrm{Re}\,\lambda_i(\mathbf{V}))^2 + (\mathrm{Im}\,\lambda_i(\mathbf{V}))^2} > 1$ and also $|\lambda_i(\mathbf{W})| = 1/|1+\lambda_i(\mathbf{V})| < 1 \quad \forall i = 1,2,...,n$, i.e. finally $\rho(\mathbf{W}) < 1$.

If $\mathbf{V}$ is column diagonally dominant with positive diagonal elements then $\mathbf{V}^T$ is row diagonally dominant with positive diagonal elements, and thus for the matrix $(\mathbf{I}+\mathbf{V}^T)^{-1} = [(\mathbf{I}+\mathbf{V})^T]^{-1} = [(\mathbf{I}+\mathbf{V})^{-1}]^T = \mathbf{W}^T$ it holds $|\lambda_i(\mathbf{W}^T)| < 1 \quad \forall i = 1,2,...,n$. However, $\mathbf{W}$ has the same eigenvalues as $\mathbf{W}^T$ and thus $|\lambda_i(\mathbf{W})| < 1 \quad \forall i = 1,2,...,n$, or finally $\rho(\mathbf{W}) < 1$. **Q.E.D.**

Since $\mathbf{Q} = (\mathbf{I}_n + \mathbf{A}_C\mathbf{U}\mathbf{A}_L)^{-1}$ where $\mathbf{A}_C\mathbf{U}\mathbf{A}_L = h\mathbf{C}_n^{-1}\mathbf{A}_{rl}\mathbf{U}h\mathbf{L}_b^{-1}\mathbf{A}_{rl}^T$ is a row diagonally dominant matrix with positive diagonal elements, the above theorem proves that $\lim_{N\to\infty}\mathbf{Q}^N = \mathbf{0}$. Also, as a special case, it proves for $\mathbf{U} = (\mathbf{I}_b + \mathbf{R}_L)^{-1}$ that $\lim_{N\to\infty}\mathbf{U}^N = \mathbf{0}$, since the positive diagonal matrix $\mathbf{R}_L = h\mathbf{L}_b^{-1}\mathbf{R}_b$ is obviously diagonally dominant with positive diagonal elements. Thus, we have finally established the nonnegativity of the upper-left blocks $(\mathbf{B})_{n\times n}$, $(\mathbf{B}+\mathbf{B}^2)_{n\times n}$, …, $(\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N)_{n\times n}$, along with the coefficients of the linear functions $\bar{v}_i \equiv \bar{v}_i(\mathbf{y})$, $i = 1,2,...,n$ in (7).

Before leaving this section, we note that the nonnegative coefficients of $\bar{\mathbf{v}}_n$ in (7) also establish that the mean – or the integral of – voltage drop within a time interval is *monotone* on the vector of excitations (meaning that increasing the current at any sink and at any time instant can only result in the increase of the mean voltage drop) which has long been known for R or RC grid models [14] but was an open problem for general RLC models (and, in fact, it does *not* hold for the instantaneous value of voltage drop).

### *D. Extension in the case of mutual inductances*

In the less common case where mutual inductances are also present in the power grid model, the matrix $\mathbf{L}_b$ of branch inductances is no longer positive diagonal and we cannot rigorously show that the blocks $(\mathbf{B})_{n\times n}$, $(\mathbf{B}+\mathbf{B}^2)_{n\times n}$, …, $(\mathbf{B}+\mathbf{B}^2+\cdots+\mathbf{B}^N)_{n\times n}$ are nonnegative. However, it is a known fact that the *inverse* $\mathbf{L}_b^{-1}$ is a symmetric diagonally dominant matrix with positive diagonal elements [15]. It can generally be proved (though the proof is rather lengthy to include here – please see [11]) that if $\mathbf{V}$ is a symmetric diagonally dominant matrix with positive diagonal elements and $\mathbf{D}$ is a positive diagonal matrix, then the matrix $(\mathbf{I}+\mathbf{V}\mathbf{D})^{-1}\mathbf{V}$ is also symmetric diagonally dominant with positive diagonal elements. This result applies to the matrix $\mathbf{U}h\mathbf{L}_b^{-1} = (\mathbf{I}_b + h\mathbf{L}_b^{-1}\mathbf{R}_b)^{-1}h\mathbf{L}_b^{-1}$ which is hence symmetric diagonally dominant with positive diagonal elements. Due to its diagonal dominance property, the matrix $\mathbf{U}h\mathbf{L}_b^{-1}$ (especially if it is a large one) is expected to behave a lot like a positive diagonal matrix within the product $\mathbf{A}_{rl}(\mathbf{U}h\mathbf{L}_b^{-1})\mathbf{A}_{rl}^T$ and produce a matrix where the entries originating from diagonal elements "dominate". Therefore the results of the previous section derived for a positive diagonal $\mathbf{L}_b^{-1}$ are expected to still hold because $\mathbf{L}_b^{-1}$ is now diagonally dominant with positive diagonal elements.

$$\mathbf{B}^N = \begin{bmatrix} \mathbf{Q}^N \pm \sum_i \left( \prod \mathbf{Q}^{N-a_i}\mathbf{A}_C\mathbf{U}^{N-b_i}\mathbf{A}_L\mathbf{Q}^{N-c_i} \right) & -\sum_j \left( \mathbf{Q}^{N-a_j}\mathbf{A}_C\mathbf{U}^{N-b_j} \right) \pm \sum_k \left( \mathbf{Q}^{N-a_k}\mathbf{A}_C\mathbf{U}^{N-a_k} \prod \mathbf{A}_L\mathbf{Q}^{N-c_k}\mathbf{A}_C\mathbf{U}^{N-d_k} \right) \\ \sum_l \left( \mathbf{U}^{N-a_l}\mathbf{A}_L\mathbf{Q}^{N-b_l} \right) \pm \sum_m \left( \mathbf{U}^{N-a_m}\mathbf{A}_L\mathbf{Q}^{N-b_m} \prod \mathbf{A}_C\mathbf{U}^{N-c_m}\mathbf{A}_L\mathbf{Q}^{N-d_m} \right) & \mathbf{U}^N \pm \sum_n \left( \prod \mathbf{U}^{N-a_n}\mathbf{A}_L\mathbf{Q}^{N-b_n}\mathbf{A}_C\mathbf{U}^{N-c_n} \right) \end{bmatrix}$$

In particular, the matrix $\mathbf{I}_n + \mathbf{A}_C\mathbf{U}\mathbf{A}_L = \mathbf{I}_n + h\mathbf{C}_n^{-1}\mathbf{A}_{rl}\mathbf{U}h\mathbf{L}_b^{-1}\mathbf{A}_{rl}^T$ is expected to be an M-matrix or quite like an M-matrix and still be inverse-nonnegative, i.e. $\mathbf{Q} = (\mathbf{B})_{n\times n} \geq \mathbf{0}$.

To demonstrate with a practical example, consider the same grid as in the previous section, but with the following branch inductance matrix which also has mutual inductances between branches (this matrix is an expansion of the matrix given in [15]):

$$\mathbf{L}_b = \begin{bmatrix} 11.4 & 4.2 & 2.5 & 1.8 & 1.4 & 1.1 & 0.9 & 0.7 \\ 4.2 & 11.4 & 4.2 & 2.5 & 1.8 & 1.4 & 1.1 & 0.9 \\ 2.5 & 4.2 & 11.4 & 4.2 & 2.5 & 1.8 & 1.4 & 1.1 \\ 1.8 & 2.5 & 4.2 & 11.4 & 4.2 & 2.5 & 1.8 & 1.4 \\ 1.4 & 1.8 & 2.5 & 4.2 & 11.4 & 4.2 & 2.5 & 1.8 \\ 1.1 & 1.4 & 1.8 & 2.5 & 4.2 & 11.4 & 4.2 & 2.5 \\ 0.9 & 1.1 & 1.4 & 1.8 & 2.5 & 4.2 & 11.4 & 4.2 \\ 0.7 & 0.9 & 1.1 & 1.4 & 1.8 & 2.5 & 4.2 & 11.4 \end{bmatrix} pH$$

The inverse of $\mathbf{L}_b$ is easy to verify that it is symmetric diagonally dominant with positive diagonal elements. The matrix $\mathbf{I}_n + \mathbf{A}_C\mathbf{U}\mathbf{A}_L$ is then the following:

$$\begin{bmatrix} 163.30 & -108.47 & 0.24 & -54.43 & 0.08 & 0.30 \\ -71.18 & 203.84 & -71.13 & -0.05 & -35.57 & 0.07 \\ 0.24 & -108.39 & 163.45 & 0.19 & -0.09 & -54.37 \\ -54.43 & -0.07 & 0.19 & 163.53 & -108.35 & 0.13 \\ 0.06 & -40.66 & -0.07 & -81.26 & 204.31 & -81.39 \\ 0.30 & 0.11 & -54.37 & 0.13 & -108.52 & 163.27 \end{bmatrix}$$

Observe that the entries not originating from diagonal elements are over two orders of magnitude smaller than those originating from diagonal elements, and normally do not play any role in the matrix properties. Indeed, this matrix is inverse-nonnegative, as can be easily verified.

For the sake of comparison, consider the same branch inductance matrix but without mutual inductances, i.e. the diagonal matrix:

$$\mathbf{L}_b = \text{diag}[11.4 \quad 11.4 \quad 11.4 \quad 11.4 \quad 11.4 \quad 11.4 \quad 11.4 \quad 11.4]\, pH$$

The matrix $\mathbf{I}_n + \mathbf{A}_C\mathbf{U}\mathbf{A}_L$ becomes now:

$$\begin{bmatrix} 163.38 & -108.13 & 0 & -54.24 & 0 & 0 \\ -70.96 & 203.47 & -70.96 & 0 & -35.59 & 0 \\ 0 & -108.13 & 163.38 & 0 & 0 & -54.24 \\ -54.24 & 0 & 0 & 163.38 & -108.13 & 0 \\ 0 & -40.68 & 0 & -81.10 & 203.89 & -81.10 \\ 0 & 0 & -54.24 & 0 & -108.13 & 163.38 \end{bmatrix}$$

which is a proper M-matrix.

For the same reason as above it is expected that $\lim_{N\to\infty} \mathbf{Q}^N = \mathbf{0}$. That also $\lim_{N\to\infty} \mathbf{U}^N = \mathbf{0}$ can be proved directly from Theorem 2 and the fact that $\mathbf{R}_L = h\mathbf{L}_b^{-1}\mathbf{R}_b$ is column diagonally dominant due to $\mathbf{L}_b^{-1}$ being symmetric diagonally dominant – both with positive diagonal elements (see [11]). Therefore the series $\mathbf{B} + \mathbf{B}^2 + \cdots$ is still expected to converge to the limit (10), whose upper-left $n\times n$ block, given by (11), is independent of $\mathbf{L}_b$ and is always nonnegative. This establishes the nonnegativity of the upper-left $n\times n$ blocks of all intermediate partial sums $\mathbf{B} + \mathbf{B}^2$, $\mathbf{B} + \mathbf{B}^2 + \mathbf{B}^3$,….

## IV. DEVELOPMENT OF A PRACTICAL POWER GRID VERIFICATION METHODOLOGY

There can possibly be a variety of ways to estimate the maximal subset of the excitation space of a digital circuit in order to employ it for verification of the power grid. In this paper, like in [5], we have adopted a statistical estimation framework, which consists of acquiring a sample of discretized current waveforms drawn from the sinks for a number of binary input vectors, computing the sample's own set of maximal points, and then statistically projecting this set to the expected global position of the maximal subset of the excitation space.

To be more specific, we first acquire a sample $S = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l\}$ of $mN$-dimensional waveform super-vectors (henceforth referred to as the "sample space") by simulating the digital circuit for $l$ random binary vector pairs $\{\mathbf{b}_p, \mathbf{b}_n\}$. This multivariate sample is made up of an assortment of $mN$ univariate samples $S_i = \{y_{i,1}, y_{i,2}, \ldots, y_{i,l}\}$, $i = 1,2,\ldots,mN$, each one representing the current observed at one sink and at a particular time instant for the $l$ random vector pairs $\{\mathbf{b}_p, \mathbf{b}_n\}$. In each univariate sample $S_i$ we can estimate the expected maximum $\omega(y_i)$ of the random variable $y_i$ sampled by $S_i$ by results from statistical extreme value theory (EVT). Specifically, if $S_i$ is partitioned into $l/r$ sub-samples of size $r$ from which the maxima units $z_{i,j} = \max\{y_{i,(j-1)r+1}, \ldots, y_{i,jr}\}$, $j = 1,2,\ldots,l/r$ are taken out to create a new sample $Z_i = \{z_{i,1}, z_{i,2}, \ldots, z_{i,l/r}\}$ of size $l/r$, then an estimate for the expected maximum $\omega(y_i)$ of $y_i$ can be computed as follows [16]:

$$\hat{\omega}(y_i) = \hat{\mu}_i + \frac{\hat{\sigma}_i}{1 + r\sqrt{\pi \log r}\left(\text{erf}(\sqrt{\log r}) - 1\right)} \tag{12}$$

where $\text{erf}(x) = (2/\sqrt{\pi})\int_0^x \exp(-t^2)dt$ is the "error function" and $\hat{\mu}_i$, $\hat{\sigma}_i$ are estimates of the location-scale parameters of the asymptotic extreme value distribution (not related to the corresponding parameters of the normal distribution), which are usually obtained by Maximum Likelihood (ML) estimation on the sample $Z_i$. However, due to the large dimension ($mN$) of the space of current waveforms it is somewhat impractical (though not entirely prohibitive) to perform ML estimation (meaning the solution of a nonlinear optimization program) for all $i = 1,2,\ldots,mN$, and we have instead used the method of matching the first and second moments (i.e. mean and standard deviation) of the sample $Z_i$ with those of the extreme value distribution, by which we have [16]:

$$\hat{\sigma}_i = (\sqrt{6}/\pi)std(Z_i) \tag{13a}$$

$$\hat{\mu}_i = mean(Z_i) - \gamma\hat{\sigma}_i \tag{13b}$$

where $\gamma \approx 0.5772\ldots$ is the "Euler gamma" constant. Experiments have shown that the above approximations found by moment matching are remarkably close to the actual ML estimates.

We pause here to stress that the units of $S$ are completely random and do not have to be close to the (unknown) worst-case vectors, since the maximum $\omega(y_i)$ for each $i = 1,2,\ldots,mN$ is actually estimated via the parameters of the extreme value distribution which is followed asymptotically (i.e. for large enough $r$) by the sample of maxima $Z_i$. This case is very similar to the estimation of the mean $\mu(y_i)$ of an r.v. $y_i$ as parameter of the normal distribution which is followed asymptotically by the sample of averages. In that case too, the units in the samples of $y_i$ from which the averages are taken out are completely random and need not necessarily be close to the unknown mean of $y_i$.

Returning to our main discussion, the sample space $S = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l\}$ has a set of maximal points of its own, which will be scaled *down* in each individual coordinate $i = 1,2,\ldots,mN$ (Fig. 3) with respect to the maximal subset of the excitation space $D$ (since there will always be points $\mathbf{y} \in D$ lying outside the outermost boundary of $S$). A reasonable approximation for this down-scaling of the maximal subset as a *whole* in each $i = 1,2,\ldots,mN$ is $\omega(y_i) - \max\{y_{i,1}, y_{i,2}, \ldots, y_{i,l}\}$, where $\max\{y_{i,1}, y_{i,2}, \ldots, y_{i,l}\}$ is the maximum value of each univariate sample $S_i$ (i.e. the maximum of the sample space $S$ in each coordinate axis). Writing this succinctly in vector form for all $i = 1,2,\ldots,mN$ as:

$$\mathbf{d} \equiv \boldsymbol{\omega}(\mathbf{y}) - \max\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l\} \tag{14}$$

[where the max operator in interpreted component-wise] we have a difference vector by which we can shift the maximal subset of $S$ in order to move it to the expected location of the maximal subset of $D$ in $\Re^{mN}$. It must, of course, be mentioned that the maximal subset of $D$ will have much different structure and include many more points than the maximal subset of $S$, but the maximum value of a linear function is fairly insensitive to the local structure of the maximal subset and instead depends predominantly on its global position in $\Re^{mN}$ (Fig. 4). In order, finally, to compute the maximal points of the space $S$ consisting of $l$ points, we have to 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 $l^2$ comparisons. It can be shown [17], however, that the necessary comparisons can be reduced to at most $O\left(l(\log_2 l)^{mN-2}\right)$, where $mN$ is the dimension of the space and its constituent points.

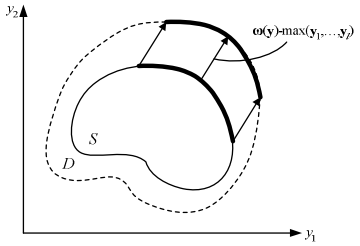


**Fig. 3.** Sample space *S* and shift of its maximal points towards the expected position of the maximal points of the excitation space *D*.

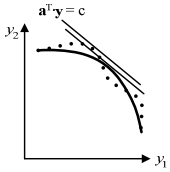


**Fig. 4.** Insensitivity of the maximum of a linear function to the local structure of the subset of maximal points.

## V. 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 G*n*-*p*, where *n* stands for the number of grid nodes and *p* for the number of power supply pads (e.g. the label G75-6 denotes a grid with 75 nodes and 6 supply pads). All test grids were uniform rectangular meshes and had equal widths for all branches in every straight line (horizontal or vertical). For the digital circuits supplied by the grids we have implemented the traditional ISCAS85 benchmarks in 90*nm* technology, and partitioned each one of them to a number of functional modules (representing the *m* current sinks). The placements of the current sinks and the power pads across the grid area were chosen in random.

For each digital circuit the process of estimation of the worst-case current excitations, involving creation of the sample space *S* (by circuit simulation), univariate extreme value estimation in each of the coordinate axes $i = 1,2,\ldots,mN$, and shifting of the maximal points of *S* (towards the global position of the maximal points of the excitation space) is *independent* of the supplying power grid and needs to be carried out only once. The main steps of this process are summarized hereafter along with some brief remarks on their implementation and computational complexity:

1) *Generate a total of* $l = 2500$ *random pairs of binary vectors* $\{\mathbf{b}_p, \mathbf{b}_n\}$ *for the circuit under consideration*. This step can be performed by any standard random number generator producing uniform numbers. The selection $l = 2500$ for the number of input pairs is discussed below.

2) *Simulate the circuit under all generated pairs and record the discretized current waveforms in each sink* $j = 1,2,...,m$ *and for each time instant* $t = kh$, $k = 1,2,...,N$ *within an interval of interest (e.g. a clock period)*. The recorded data $S_i = \{y_{i,1}, y_{i,2}, \ldots, y_{i,l}\}$, $i = 1,2,\ldots,mN$, taken jointly as $mN$-dimensional vectors will constitute the sample space $S = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l\}$. The number *N* of time instants within the interval can be kept small, as seen in the examples of Section III-C (a number $N = 10$ should be enough). 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 2500 binary input pairs is sufficient to produce a reasonable statistical estimate *independently* of the circuit size or sink size, as is further explained below.

3) *Arrange each univariate sample* $S_i$, $i = 1,2,\ldots,mN$ *into* $l/r = 100$ *sub-samples of size* $r = 25$. Here the size *r* only needs to be adequate so that the sample of the maxima units from the sub-samples follows an asymptotic extreme value distribution. We have found by experimentation that $r = 25$ is a fair value. The number $l/r = 100$ of sub-samples (leading to a total of $l = 2500$ 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 [16]. 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 $l/r$ of sub-samples will have to be increased (together with the total number *l* of input pairs).

4) *For each* $i = 1,2,\ldots,mN$ *construct the sample* $Z_i$ *of the maxima units from the* $l/r$ *sub-samples of* $S_i$.

5) *For each* $i = 1,2,\ldots,mN$ *calculate the estimates* $\hat{\mu}_i$, $\hat{\sigma}_i$ *of the extreme value distribution parameters from* (13), *and the estimate* $\hat{\omega}(y_i)$ *of the expected maximum* $\omega(y_i)$ *from* (12).

6) *Determine the maxima* $\max\{y_{i,1}, y_{i,2}, \ldots, y_{i,l}\}$ *of all univariate samples* $S_i$, $i = 1,2,\ldots,mN$, *and in conjunction with the estimates* $\hat{\omega}(y_i)$, $i = 1,2,\ldots,mN$ *for the expected maxima*, *compute the mN -dimensional difference vector* **d** *from* (14).

7) *Locate the maximal points among the* $l = 2500$ *points of the sample space S*. As already mentioned, this step has complexity of $O\left(l(\log_2 l)^{mN-2}\right)$ comparisons.

8) *Shift the maximal points of the sample space S by the computed difference vector* **d**. This step is performed by plain component-wise addition of the vector **d** to the maximal points of *S*, and is a trivial one.

For the practical implementation of the above procedure, a custom simulator program [18] was employed for the creation of the sample space in step 2 (although any other transistor-level simulator could have been used instead), while a program in C/C++ was developed for the execution of steps 3 to 8. Experiments were run on an Intel Core2 2.7GHz with 2GB RAM. The running times for some benchmark circuits, along with information concerning the number of sinks and the computed number of maximal points (which coincides with the final number of the worst-case excitations) are shown in Table I. It is easily observed that the running time of the whole procedure is almost exclusively determined by the running time of step 2.

**Table I.** Results on the estimation of the worst-case current excitations (CPU times and number of excitations) for various benchmark circuits.

| Circuit | Num. of sinks | CPU time for step 2 | CPU time for steps 3-8 | Total CPU time | Num. of maximal points |
|---|---|---|---|---|---|
| c1355 | 5 | 2164.76s | 2.07s | 2166.83s | 44 |
| c2670 | 13 | 2433.22s | 8.45s | 2441.67s | 181 |
| c6288 | 32 | 2674.16s | 20.70s | 2694.86s | 231 |
| c7552 | 42 | 2804.07s | 25.49s | 2829.56s | 502 |

The output of steps 1 to 8 is a set of shifted sample maximal points for a particular circuit which approximate the position of the maximals of its excitation space, and thus constitute worst-case waveform excitations for any grid supplying the circuit. The verification of any given grid is then performed by the following steps:

9) *Apply the shifted maximal points as excitation waveforms in a linear network simulator to perform an analogous number of transient analyses for the given power grid*. This step relies exclusively on a linear network simulator, and its execution time is determined by the capability of the simulator to carry out the required analyses for the given grid.

10) *For each sink* $j = 1,2,...,m$ *compute the mean voltage drop for each transient analysis and determine the maximum value among the computed mean voltage drops*. The resulting value for each sink finally constitutes an estimate of the worst-case cycle-mean voltage drop over all possible cycles and corresponding binary vector pairs.

In our implementation we have employed a direct solver of the transient linear network equations in (6), based on factorization of the system matrix and without utilizing any sparse matrix techniques (which could further speed up the solution). The results for the maximum voltage drops in various test grids supplying some of the ISCAS85 benchmark circuits are shown in Table II. All computed worst-case voltage drops are compared to

statistical estimates obtained by applying the univariate extreme value estimation procedure of [16] on direct samples of voltage drops for each circuit/grid combination. These direct statistical estimates can be considered to be as close as possible to the actual maximum voltage drop values (within a certain confidence interval). From the table it can be readily verified that the direct statistical estimates and those obtained by estimation of the worst-case excitations (as being proposed here) are remarkably close to each other. A slight pessimism which is observed for the proposed method is not a matter of concern (since it will not lead to any grid underdesign), and can possibly be attributed to the deviation of the shifted maximal points of the sample space compared to the maximal points of the excitation space, which eventually seems to lie on the pessimistic side (i.e. the vector (14) slightly overestimates their relative positions). The runtimes for the direct estimation are much larger because they include all necessary setup (sample acquisition etc.) which, for the proposed approach, was reported separately in Table I. However, the runtime overhead of Table I is a "one-time cost" that needs to be endured only once for a specific digital circuit, while the setup (and its associated execution time) for the direct estimation has to be repeated every time the grid changes. Since, at the signoff stage, the grid typically undergoes many iterations of redesign and verification (with the same underlying circuit) until deemed robust, the calculation of maximum voltage drop via estimation of the worst-case excitations becomes the only practical solution for power grid verification.

**Table II.** Comparative results of max. voltage drop estimation [direct estimation vs. estimation via worst-case excitations (proposed solution) vs. pessimistic estimation] at two current sinks for various benchmark circuits and test power grids.

| Case study | Grid | Circuit | Max. voltage drop to sink-A (mV) | | | | Max. voltage drop to sink-B (mV) | | | | CPU times | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Direct statistical estimation | Our method | Pessim. analysis (MEC) | % diff. | Direct statistical estimation | Our method | Pessim. analysis (MEC) | % diff. | Direct statistical estimation | Our method | Pessim. analysis (MEC) |
| #1 | G25-2 | c1355 | 522.0 | 539.2 | 732.5 | 35.9 | 535.7 | 548.5 | 743.8 | 35.6 | 2165.59s | 0.014s | <1ms |
| #2 | G50-3 | c1355 | 299.5 | 303.7 | 416.7 | 37.2 | 288.3 | 292.1 | 397.8 | 36.2 | 2168.05s | 0.057s | <1ms |
| #3 | G75-6 | c1355 | 141.1 | 141.7 | 190.1 | 34.1 | 125.8 | 126.0 | 168.8 | 33.9 | 2172.70s | 0.157s | 0.016s |
| #4 | G100-10 | c1355 | 56.9 | 57.1 | 77.8 | 36.2 | 53.8 | 54.1 | 73.7 | 36.1 | 2178.85s | 0.294s | 0.051s |
| #5 | G100-6 | c2670 | 196.2 | 230.8 | 341.5 | 48.0 | 191.5 | 224.7 | 332.4 | 47.9 | 2446.71s | 1.00s | 0.044s |
| #6 | G150-10 | c2670 | 106.9 | 119.4 | 178.7 | 49.6 | 110.0 | 119.5 | 178.6 | 49.4 | 2465.44s | 2.61s | 0.161s |
| #7 | G400-10 | c2670 | 74.8 | 82.9 | 126.1 | 52.1 | 72.9 | 81.5 | 122.2 | 49.9 | 2698.25s | 21.06s | 1.28s |
| #8 | G625-15 | c2670 | 58.0 | 66.6 | 98.6 | 48.0 | 57.1 | 65.0 | 96.0 | 47.8 | 3153.22s | 55.98s | 4.82s |
| #9 | G400-10 | c6288 | 130.4 | 160.7 | 247.2 | 53.8 | 128.0 | 157.6 | 242.0 | 53.6 | 2937.36s | 25.32s | 1.35s |
| #10 | G400-15 | c6288 | 85.6 | 108.0 | 166.5 | 54.1 | 88.0 | 112.0 | 172.5 | 54.0 | 2948.47s | 26.54s | 1.30s |
| #11 | G900-15 | c6288 | 84.3 | 107.5 | 165.7 | 54.2 | 86.1 | 104.7 | 161.4 | 54.1 | 4346.45s | 166.56s | 13.74s |
| #12 | G1369-20 | c6288 | 56.2 | 69.0 | 106.0 | 53.5 | 55.2 | 67.9 | 104.3 | 53.5 | 6331.49s | 383.32s | 47.54s |
| #13 | G400-10 | c7552 | 419.7 | 425.7 | 590.8 | 38.8 | 427.6 | 433.4 | 601.1 | 38.7 | 3093.84s | 54.11s | 1.37s |
| #14 | G900-15 | c7552 | 283.7 | 290.7 | 400.8 | 37.9 | 292.2 | 299.3 | 412.4 | 37.8 | 4506.87s | 334.22s | 14.08s |
| #15 | G1369-20 | c7552 | 186.9 | 191.9 | 265.6 | 38.4 | 189.2 | 194.3 | 268.5 | 38.2 | 6549.67s | 758.98s | 47.41s |

For every tested case a pessimistic analysis has also been carried out by forming a fictitious waveform consisting of the estimates of the expected maxima $\hat{\omega}(y_i)$, $i = 1,2,\ldots,mN$ for each sink and each time instant. This is effectively a construction of the Maximum Envelope Current (MEC) waveform that was introduced in [14] and which was subsequently used in a number of papers as a (pessimistic) upper bound waveform for power grid verification. We can clearly see the overestimation incurred by this pessimistic analysis which was in the range 35-55% for all tested circuit/grid combinations. Admittedly, the use of a MEC-type waveform for grid verification is faster because it only entails simulation for one (unrealistic) waveform excitation instead of a set of (realistic) worst-case ones – the latter being typically in the order of a few hundreds, as seen in Table I (the one-time setup cost of Table I is the same for both methods). However, we believe that the additional computational cost of the proposed method is justified by the capability to accurately predict the maximum voltage drop, since modern power grid design involves an extremely tight trade-off between robustness and area resources, and increasing the grid overhead by nearly 50% can prove to be completely impractical. Since the ISCAS85 benchmarks are actually small circuits compared to today's standards, the differences between the proposed method and the MEC-based analysis are expected to be even more pronounced in the case of larger designs with several current sinks and more complex interdependencies between them.

## VI. CONCLUSION

A method for general RLC-model power grid verification has been developed, which is based on characterization of the worst-case current waveform excitations, and their subsequent estimation by sampling the excitation space and statistically projecting the sample's own worst-case excitations to their expected position in the excitation space. Experimental results have demonstrated the potential of the method to accurately predict the worst-case voltage drop at any node of the power distribution network.

## REFERENCES


[1] S. Sapatnekar and H. Su, "Analysis and optimization of power grids", *IEEE Design and Test of Computers*, May 2003.

[2] M. Nizam, F. Najm, and A. Devgan, "Power grid voltage integrity verification", *ACM/IEEE Int. Conf. Low-Power Electron. & Design*, 2005.

[3] H. Qian, S. Nassif, and S. Sapatnekar, "Early-stage power grid analysis for uncertain working modes", *IEEE Trans. Computer-Aided Design*, vol. 24, pp. 676-682, 2005.

[4] N. Ghani and F. Najm, "Handling inductance in early power grid verification", *ACM/IEEE Int. Conf. Computer-Aided Design*, 2006.

[5] N. Evmorfopoulos, D. Karampatzakis, and G. Stamoulis, "Precise identification of the worst-case voltage drop conditions in power grid verification", *ACM/IEEE Int. Conf. Computer-Aided Design*, 2006.

[6] I. Ferzli, F. Najm, and L. Kruse, "A geometric approach for early power grid verification using current constraints", *ACM/IEEE Int. Conf. Computer-Aided Design*, 2007.

[7] N. Srivastava, X. Qi, and K. Banerjee, "Impact of on-chip inductance on power distribution network design for nanometer scale integrated circuits", *IEEE Int. Symp. Quality Electronic Design*, 2005.

[8] A. Ruehli (ed.), *Circuit Analysis, Simulation and Design*, North-Holland, 1986.

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

[10] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge, 1990.

[11] N. Evmorfopoulos, "Some results on diagonally dominant matrices with positive diagonal elements", UTH Technical Report (http://www.inf.uth.gr/images/stories/eggrafa/research/inf-001.pdf)

[12] A. Berman and R. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, 1979.

[13] Y. Saad, *Iterative Methods for Sparse Linear Systems,* SIAM, 2003.

[14] 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.

[15] H. Ji, A. Devgan, and W. Dai, "KSim: A stable and efficient RKC simulator for capturing on-chip inductance effect", *ACM/IEEE Asia & South Pacific Design Automation Conf.*, 2001.

[16] 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.

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

[18] D. Bountas, G. Stamoulis, and N. Evmorfopoulos, "A macromodel technique for VLSI dynamic simulation by mapping pre-characterized transitions", *IEEE Int. Conf. Computer Design*, 2008.