# Design and Analysis of Power Distribution Networks with Accurate RLC Models

*Rajat Chaudhry, Rajendran Panda, Tim Edwards, David Blaauw*
Advanced Tools Group, Motorola, Austin, TX 78730.
**email**: rajat.chaudhry@motorola.com

**Abstract**

Due to higher power and faster switching frequencies a very robust power distribution network is required. To achieve this the power distribution network needs to be modeled accurately at different stages of the design cycle. We present a methodology for the design and analysis of power distribution networks. The methodology covers the need for power grid analysis across all stages of the design process. We present techniques to accurately model the effects on power distribution networks due to parasitic wire capacitances, parasitic capacitances of transistors, explicit decoupling capacitors and inductance from package leads.

## 1. Overview

The power distribution network distributes power and ground voltages from pad locations to all gates and devices in a design. Higher device densities and faster switching frequencies cause large switching currents to flow in the power and ground networks which degrade performance and reliability. Due to the resistance of the network, there is a voltage drop across the network, commonly referred to as the *IR-drop*. The package leads that supply voltage to the pads of the power grid have significant inductance. This inductance causes a voltage drop at the pad locations due to the time varying current drawn by connected devices. This voltage drop is referred to as the *di/dt-drop.* Therefore the voltage seen at the devices is the supply voltage minus the *IR-drop* and *di/dt-drop*. The parasitic capacitances between the power and ground grids and the parasitic device capacitances smoothen out the voltage fluctuations. In addition to the wire and device capacitances intrinsic to the circuit, the designer can add explicit decoupling capacitances in certain areas to improve the voltage drop.

Excessive voltage drops in the power grid reduce switching speeds and noise margins of circuits, and inject noise which might lead to functional failures. High average current densities lead to undesirable wearing out of metal wires due to electromigration[1]. Therefore, the challenge in the design of a power distribution network is in achieving excellent voltage regulation at the consumption points notwithstanding the wide fluctuations in power demand across the chip, and to build such a network using minimum area of the metal layers. These issues are prominent in high performance microprocessors as large amounts of power have to be distributed through a hierarchy of many metal layers. A robust power distribution network is vital in meeting performance guarantees and ensuring reliable operation of high performance microprocessors.

The crux of the problem in designing a power grid is that there are many unknowns until the very end of the design cycle. Nevertheless, decisions about the structure, size and layout of the power grid have to be made at very early stages when a large part of the chip design has not even begun. Unfortunately, most commercial tools focus on post-layout verification of the power grid when the entire chip design is complete and detailed information about the parasitics of the power and ground lines and the currents drawn by the transistors are known. Power grid problems revealed at this stage are usually very difficult or expensive to fix. A methodology was presented in [9][11] that helps to design an initial power grid and refine it progressively at various design stages. In this paper, we show how such a methodology can be extended to include the effects of package inductance in the analysis. Due to the growth in power consumption and switching speeds of modern high performance microprocessors, the *di/dt* effects are becoming a growing concern in these designs. Designers rely on the on-chip parasitic capacitances and intentionally added decoupling capacitors to counteract the *di/dt* variations in the voltage[10]. But it is necessary to model accurately the inductance and capacitance of the package and chip and analyze the grid with such models, as otherwise the amount of decoupling to be added might be underestimated or overestimated. Also it is necessary to maintain the efficiency of the analysis even when including these detailed models. In this paper we present techniques for efficient handling of these models in the analysis.

A critical issue in the analysis of power grids is the large size of the network (typically millions of nodes in a state-of-the-art microprocessor). Simulating all the non-linear devices in the chip together with the non-ideal power grid is computationally infeasible. To make the size manageable, the simulation is done in two steps. First, the non-linear devices are simulated assuming perfect supply voltages and the currents drawn by the devices are measured. Next, these devices are modeled as independent time-varying current sources for simulating the power grid and the voltage drops at the transistors are measured. Since voltage drops are typically less than 10% of the power supply voltage, the error incurred by ignoring the interaction between the device currents and the supply voltage is small. By doing these two steps, the power grid analysis problem reduces to solving a linear network which is still quite large. To further reduce the network size, we exploit the hierarchy in the chip design.

Note that the circuit currents are not independent due to signal correlations between blocks. This is addressed by deriving the inputs for individual blocks of the chip from the results of logic simulation using a common set of chip-wide input patterns. An important issue in power grid analysis is to determine what these input patterns should be. For IR-drop analysis, patterns that produce maximum instantaneous currents are required, whereas for electromigration purposes, patterns producing large sustained (average) currents are of interest. Determining either set of patterns is a difficult problem and has been addressed by many, including [2][3][4]. This issue will not be addressed in this paper, and we will assume that suitable input patterns are available for simulation. Moreover, in this paper, we will not discuss electromigration issues, but concentrate on the issue of computing the worst voltage drops in the power network.

This paper is organized as follows. The next section describes the various analysis modes implemented in our methodology. Section 3 discusses linear solution techniques in the context of power distribution networks. Section 4 discusses package inductance and device capacitance modelling and the paper concludes with discussion of open issues.

## 2. Voltage Drop Analysis Modes

To apply the voltage drop analysis methodology described in this paper across all stages of the design of a complex microprocessor, we define several modes of operation of the tool. These modes are distinguished by different models of the power distribution network and of the currents being drawn by the functional blocks. In this paper, we describe three different *modes* of operation: *early* or *pre-floorplan* mode, *post-floorplan* mode and *post-layout* mode. As the design proceeds, voltage-drop analysis is run in these different modes using more and more accurate models of the power grid and the block currents.

### 2.1 Early Mode Analysis

At the very early stages of the design of a microprocessor, there are a number of issues related to the power distribution network that have to be addressed. These include locations of the clean VDD/GND pads, nominal pitches and widths of metal layers, via styles

(point or bar vias), and parameters of the chip package. Since at this early stage of the design, the power network has not yet been synthesized and the location and logic content of the blocks are not known, IR-drop analysis is performed using very simplistic models of the grid topology and the block currents.

1. A mock power grid down to the lowest metal layer is constructed using a simple uniform grid topology, where the metal lines in each layer have a user-specified pitch (separation) and width. At the areas where the metal lines of adjacent layers cross over, vias are placed according to user-specified via geometries and via styles. Other topologies such as rings can also be modeled. The clean VDD/GND pads can be placed at the periphery of the chip or on the surface of the chip using C4 pads (for flip-chip packages).
2. To model the currents drawn by the devices, a simple area-based DC estimate of the current is used. This is obtained by taking the current estimate of a previous chip and scaling it by the power supply voltage, operating frequency, complexity, size and technology variables. This estimate is inflated 3-7 times to account for differences between the average and maximum instantaneous currents and to obtain a robust grid. The current sinks are placed on the lines of the lowest metal layer at points midway between adjacent vias that connect the lowest layer to the upper layer. The value of the current is obtained by multiplying the per-unit-area current by the product of the pitches of the two metal layers.

Using simple length-based resistance formulae, a resistive electrical network is constructed from the mock grid topology. DC analysis of this network yields the IR-drops at various locations of the chip. This analysis is very fast and allows the designer to evaluate a large number of different topologies and to trade-off robustness and metal utilization in the power grid. This analysis is used to design the locations of the C4 pads and nominal pitches and widths of the metal layers. Moreover, if the processing technology allows different width and thickness combinations for some of the top metal layers, the user can determine the best values of these in terms of IR-drop. Even though the real power grid will not be as regular as the mock grid, and all the devices will not be drawing the estimated current simultaneously, important design decisions are made from the results of this simple analysis and an early picture of the robustness of the grid is obtained.

## 2.2 Post-Floorplan Analysis

In this mode, the global power distribution network has been designed and the blocks have been placed. The locations and geometries of the power lines and the blocks are read from the design database. Even though the blocks are placed, the power grids within them have not yet been wired. The *power service terminals* (PST's) of a block are the wires in the topmost metal layer within the block that connect the global and intra-block power networks. In this mode, the PST's for a block may or may not be known - if they are not known, mock PST's are constructed. Next, the block *ports* are determined by the intersection or overlays between the global lines and the block PST's. In our hierarchy, *blocks* are custom datapath components, synthesized random logic macros (RLM), and off-the-shelf (OTS) components (custom components that can be reused). Custom components are small but RLM's can be large. OTS components can range from small blocks (nands, nors, muxes, etc.) to large blocks (adders, comparators, etc.). A *functional block* (e.g., floating point unit, memory management unit) consists of several instances of custom, synthesized and off-the-shelf components.

Each block current can be independently described in one of the following ways, thus allowing a mixture of them.

1. If the logic content of a block is not yet defined, the current model is a DC estimate based on the block area. The total block current is divided equally among all the ports. Since the area-based numbers are calculated such that they reflect peak expected currents, the analysis where every block has area-based currents is likely to be pessimistic since it assumes that each block draws this current simultaneously.
2. The next more accurate block current model is derived from a full-chip gate-level power estimation tool. Given a set of chip-wide input vectors, this tool computes the average power consumed by each block over a cycle. From the average power consumed by a block, an average block current is computed and distributed equally among all of its ports. Hence in this mode, a multi-cycle DC current signature is used for a block. Since chip-wide vectors are used for the simulation, correlation among the blocks is preserved.
3. The most accurate current model comes from a detailed transistor-level simulation of a *functional block* using PowerMill [5]. The input vectors for the functional block are derived from the chip-wide vectors through logic simulation. This ensures that correlation across functional blocks is maintained. The transistor level netlist of the block is available and capacitances are extracted for the signal nets. However, since the power grid within the block has not been designed, it is considered to be ideal. The transient current waveform drawn by each custom, RLM or OTS block within the functional block is obtained from the PowerMill simulation and is divided equally among all of their individual ports. Since the blocks are not very large and the power grid within them have not yet been wired, this block current model is quite accurate for this stage of design.

If all block current models are derived from methods 1 and 2 above, then a resistance-only electrical network is extracted from the geometrical information using length-based resistance formulae. DC analysis is then performed to yield the IR-drop values at each of the block ports (multiple DC analyses if multi-cycle DC current signatures are used). If transient current signatures are used for some of the blocks, then an RC network is extracted from the global grid using length-based resistance extraction and statistical rule-based capacitance extraction (since global routing is not done, these statistical rules account for coupling between power and signal lines). Figure 1 shows an example of IR-drop analysis in this mode for the PowerPC$^{TM}$750 microprocessor. The worst case voltage drop is 170mV.

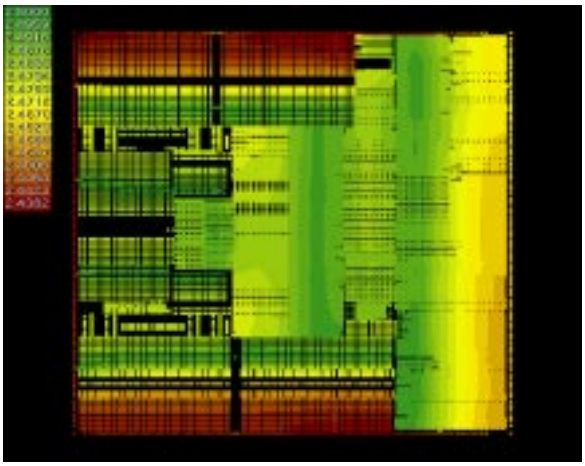

Figure 1.

### 2.3 Post-Layout Mode

This model is used when the global and block-level grids have been completely designed. Using the layout of the power grid, an accurate RC model of the power grid is extracted using a commercial extraction tool. In addition to the resistance of the network, the device capacitances and the inductance of the package can also be modeled. The models for inductance and capacitance are discussed in more detail in section 4. The tap points for each device are determined during extraction. The current profile for each gate is determined by using a fast transistor

level simulator.

In this mode the power network analysis produces time-varying voltages and currents at all points in the power grid. Figure 2 shows an example of the Post layout based analysis for the PowerPC$^{TM}$750 microprocessor. The worst drop in this case is 170mV which corresponds well with the floor plan analysis mode.

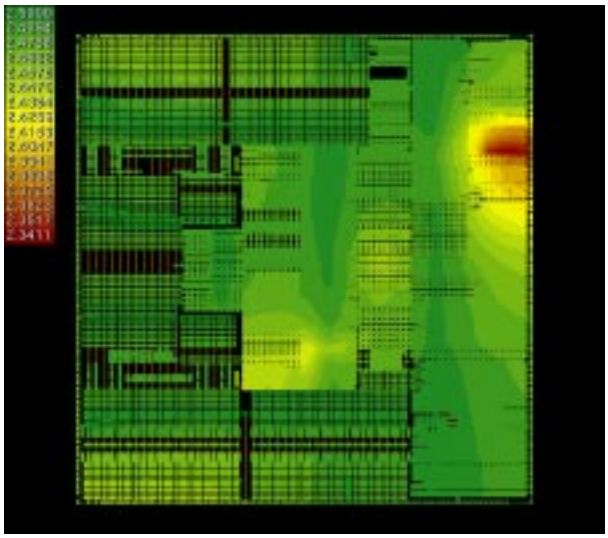

Figure 2.

## 3. Linear System Solution Techniques

Several direct and iterative approaches are available [7] to solve linear systems. In this section, we analyze the relative merits and limitations of these methods as applied to solving large power networks. The size and structure of the conductance matrix of the power grid is important in determining the type of linear solution technique that should be used. Typically, the power grid contains millions of nodes, but the conductance matrix is sparse (typically, less than 5 entries per row/column). This matrix is also symmetric positive definite, but for a purely resistive network, it may be ill-conditioned. Sparsity favors the use of iterative methods, but convergence is slowed down by ill-conditioning and can be mitigated to some extent by preconditioning. Iterative methods do not suffer from size limitations so long as the (sparse) matrix and some iteration vectors can fit into the memory. The single-biggest problem with direct methods is the need for large amounts of memory to store the factors of the matrix. The number of fill-ins is of the order of O($N^2$), where $N$ is the number of rows/columns in the matrix. However, if fixed time steps are used for transient analysis, then the initial factorization can be reused with subsequent current vectors, thus amortizing the large decomposition time. Iterative methods do not have this feature of reusability.

## 4. Inductance and Capacitance models for power networks

A more detailed model including the network capacitances, device parasitic capacitances, explicit decoupling capacitors and the inductance of the package is often used when analyzing the network with time varying currents. Due to higher switching speeds the *di/dt-drop* is becoming a significant part of the total voltage drop. Therefore it is becoming very important to model the effect of package inductance. Analyzing the inductance effects without accurately modelling the capacitance will give meaningless results since the high fluctuation of currents at the pads will cause very large voltage swings.

### 4.1 Capacitance Models

Three sources of capacitances affect the voltages in the power grid: (1) parasitic wire capacitances between power wires and ground wires, substrate, or signal nets, (2) parasitic capacitance of transistors, and (3) explicitly placed decoupling capacitors.

Parasitic wire capacitances can be extracted using either approximate Chern equations, which use the width and spacing between wires, or by using a commercial extraction tool such as Star-RC[8]. A difficult issue is the treatment of the coupling capacitors to signal nets. The affect of these capacitors on the voltage in the power grid depends on the state of the signal net. For example coupling from a power network to a signal net that is high simply couples the power network to itself, with little or no effect on the voltage drop. Unfortunately, it is infeasible to model the signal nets and power grid simultaneously. Due to this reason a statistical approach needs to be taken while modelling the coupling capacitance between the power grids and signal nets. The switching activity is determined by calculating the average number of signal nets that switch in a clock cycle. Since the low and high switching probability of a signal is equal, it is reasonable to assume that their effects cancel each other out. Therefore switching nets can be ignored. Of the remaining non-switching nets half are considered in stable high state and half in stable low state. Each coupling capacitance is replaced by an effective capacitance to ground, in series with a resistor. The effective value of the coupling capacitor is

$$C_{eff} = \frac{1}{2} C_{coupling} \cdot (1 - P_{active})$$

where $P_{active}$ is the average switching activity. The resistor models the effective resistance of the gate holding the signal net in its stable state, and can be an average value over all gates or specific value for each particular gate.

Device capacitances attenuate the voltage drop in the power grid and have a larger effect on the voltage in the power grid as they are much larger than the wire capacitances. The effect of device capacitances also depends on the state of the signal. Again we propose a statistical approach to model these capacitances. Each transistor has 5 device capacitances, $C_{sb}$(source to bulk), $C_{db}$(drain to bulk), $C_{gs}$(gate to source), $C_{gd}$(gate to drain) and $C_{gb}$(gate to bulk). The $C_{sb}$ can be ignored since the source and bulk for both the pmos and nmos are always at the same potential. Figure 3 shows the remaining 4 device capacitances for an inverter. In figure 4 the capacitances are arranged across three inverters to make the analysis more convenient. Capacitances of switching devices contribute to the current drawn from the grid, which is already modeled by the time varying current source in the power grid analysis. Therefore in power grid analysis we need to consider the device capacitances of only those gates which do not switch.

We first look at the case where net N in Figure 3(a) is in a low state. The device capacitances shown in Figure 3(a) can be modeled with the equivalent RC circuit shown in Figure 3(b). The resistance Rp corresponds to the effective pull-up resistance of inverter 1, the resistance $R_n$ corresponds to the effective pull-down resistance of inverter 2, and the resistance $R_{wp}$ and $R_{wn}$ correspond to the P and N well resistances, respectively. Since net N is low, capacitances $C_{dbn}$, $C_{gsn}$, and $C_{gbn}$ are discharged and do not contribute to the decoupling between the power and ground grid. Furthermore, since $R_{wp}$ is a relatively high resistance and since $C_{dbp}$ and $C_{gbp}$ are small, they can be ignored without significant loss in accuracy. An analogous analysis can be made for the decoupling capacitances when the state of signal N is high. This state is assumed to have equal probability of being high or low when the gate is not switching, although a different ratio of high to low signal states could easily be incorporated in the analysis. An approximate model of the device capacitances for power grid analysis is shown in figure 3(c). The effective decoupling

capacitance in this simplified model is the sum of the effective high and low decoupling capacitances weighted by the probability of the gate being in either state:

$$C_{eff} = (1 - P_{active})(C_{gdp} + C_{gdn}) + \frac{1}{2}(1 - P_{active})(C_{gsn} + C_{gsp})$$

Similarly, the effective resistance is approximated by the sum of the high and low resistance of the gates:

$$R_{eff} = R_p + R_n$$

It is important that the resistance and capacitance network that is extracted for the power grid is not reduced using standard reduction techniques, since these techniques are designed for reducing signal nets. Such techniques preserve the delay characteristics of a signal net with high fidelity, but may not preserve the voltage drop in a power grid correctly.

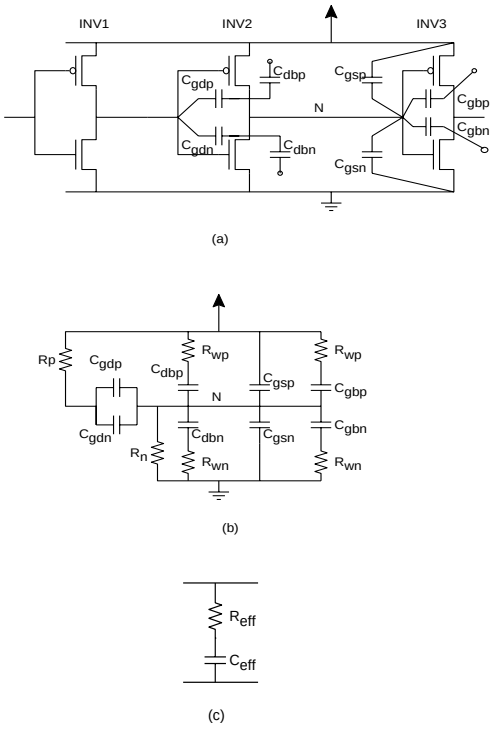


Figure 3 Capacitance model for device capacitance in power grid analysis

## 4.2 Inductance Model

Package leads have resistance, self-inductance and mutual-inductance to other package leads. The package model forms a dense symmetric partial inductance matrix, where there is mutual inductance between each of the entries. To simplify the computation the package leads are grouped into equivalent leads which supply power to a group of pads in close proximity. Figure 4 shows a package model and the corresponding reduced model. If the package includes decoupling capacitances they should also be modelled.

Since there is mutual inductance between all the package leads the current through both Vdd and Gnd leads needs to be modeled simultaneously. This can be done by simultaneously analyzing both the Vdd and Gnd grids, but this doubles the size of the network to be analyzed. The alternative is to make a simplifying assumption that the current through each power lead is equal to a an associated ground lead. This is a good assumption if the power and ground leads are in close proximity. With this assumption we can take into account the mutual inductance of the ground leads on the power leads or vice versa, without modelling both Gnd and Vdd networks separately.

Certain computational advantages are lost when the RC network model is expanded to include the model for the package and off-chip effects. Note that an RC model of the power network produces a symmetric positive definite left-hand side(LHS) matrix for a modified nodal analysis of the system Ax = b. The symmetry and positive definiteness of A makes it possible to solve the system very efficiently using either Cholesky factorization or Incomplete Cholesky Conjugate Gradient approach. However, when the package model is included in the analysis, the self and mutual inductors of the package now require the currents through the inductors to be declared as variables as well, resulting in a LHS matrix whose positive definiteness is not guaranteed. One could use general solution techniques, such as LU-decomposition, to solve the RLC model, but the large size of the network makes such general solution infeasible. We overcome this difficulty using the following partitioning approach.

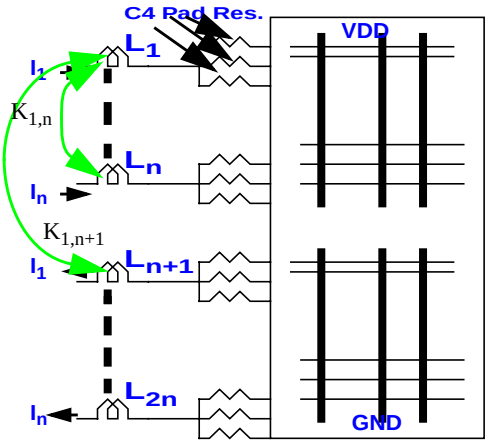


Figure 4.

The network is partitioned at the interface between the package network and the power grid network. Then the power grid network (consisting only of R's, C's, and the current sources) is reduced to equivalent admittance and currents at the interface points, and the package network can now be solved including the reduced representation of the power grid. This approach works well since the reduction of the power grid network involves solving a symmetric positive definite system. The package network with the reduced power grid model is much smaller and hence can be solved using a general solution technique, such as LU-decomposition. The idea behind this approach is illustrated in Figure.5

Suppose Ax = b is the system representing the combined network, where

$$A = \begin{bmatrix} A11 & A12 \\ A12^T & A22 \end{bmatrix} \qquad x = \begin{bmatrix} x1 \\ x2 \end{bmatrix} \qquad b = \begin{bmatrix} b1 \\ b2 \end{bmatrix}$$

*x1* is the vector of voltages at the nodes in the power grid, and *x2* is the vector of voltages and inductor currents in the package network.

The above is solved in two steps as follows.

$$x2 = (A22 - A12^T \cdot A11^{-1} \cdot A12)^{-1} \cdot (b2 - g12^T \cdot g11^{-1} \cdot b1)$$

$$x1 = g11^{-1} \cdot (g22 - g12 \cdot x2)$$

Note that the pre-multiplications with the inverse of A11 are efficiently done indirectly using Cholesky factors or Conjugate gradient method. The model reduction requires solving the power network *p+1* times, where p is the number of interface nodes.

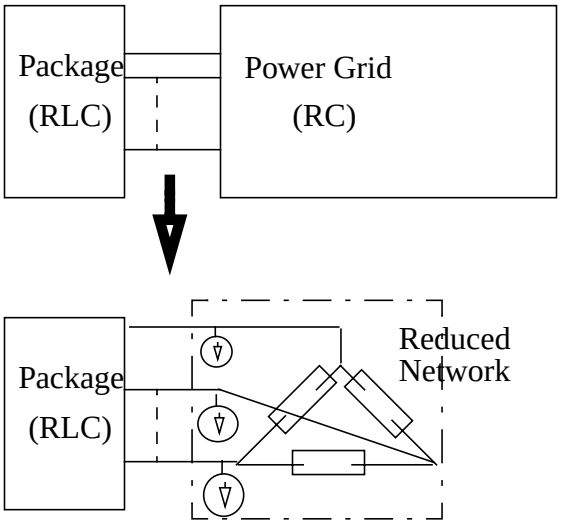


Figure 5.

### 4.2 Comparison between R, RC and RLC analysis

Figure 6 shows the analysis of the power grid of the PowerPC$^{TM}$750 processor with three types of models. The y-axis in figure 6 represents the voltage in volts and the x-axis is time in nano seconds. The supply voltage was 2.5V and the cycle time is 5 nano seconds.The first curve (R) shows the analysis when only the resistance of the power grid was modeled. In this case the worst drop is 199mV. The second curve (RC) shows the analysis when the power grid resistance and the intrinsic decoupling cap of the grid is modelled. In this case the worst drop improved to 170mV from 199mV for the R-only case. The total intrinsic decoupling cap for the whole chip was 30nF. The third curve shows the analysis with power grid resistance, device decoupling caps and package inductance. Due to the package inductance the worst drop has increased to 233mV from 170mV for the RC case. We can also notice the change in the voltage at the pad at points where the voltage is higher compared to the RC curve.

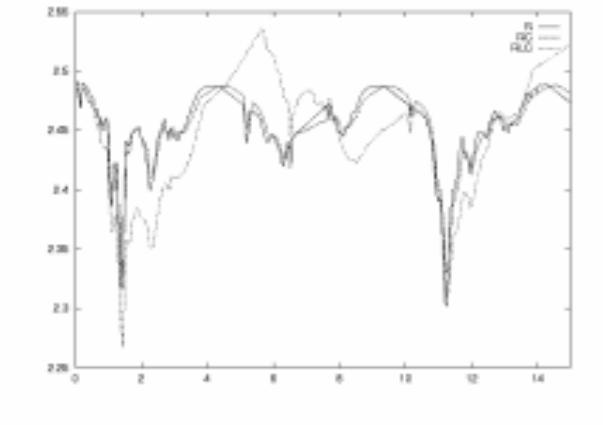

Figure 6.

Figure 7 shows the analysis with the power grid resistance and the package inductance, excluding the capacitance. The curve shows that without capacitance the di/dt noise is very significant. Therefore any meaningful analysis with package inductance needs to include the capacitance model also.

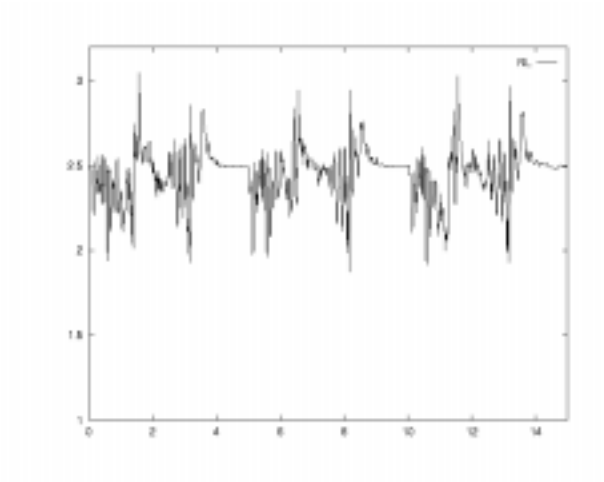

Figure 7.

### 5. Conclusions

We presented a coherent design and analysis flow for designing the power grids of large, high performance processors. The usefulness of the multi-mode analysis capability in progressively refining the design of a power grid design was demonstrated through examples. Several issues related to power grid analysis were discussed and examples of the PowerPC$^{TM}$750 processor was presented.

Areas of future work involve investigating fast solution techniques for power and ground networks, calibrating and reducing the error to hierarchical analysis, incorporating inductance and capacitance effects into the voltage-drop analysis at the early stage.

### References


[1] Black, J.R. Electromigration failure modes in aluminum metallization for semiconductor devices. *Proc. IEEE*, pp. 1587-1594, Sept 1969.

[2] S. Chowdhry et al. Estimation of maximum currents in MOS IC logic circuits. *IEEE Trans. CAD*, pp. 642-654, June 1990.

[3] H. Kriplani et al. Pattern independent maximum current estimation in power and ground buses of CMOS VLSI circuits: algorithms, signal correlations, and their resolution. *IEEE Trans. CAD*, pp. 998-1012, Aug 1995.

[4] A. Krstic et al. Vector generation for maximum instantaneous current through supply lines for CMOS circuits. *Proc. 34th DAC*, pp. 383-388, June 1997.

[5] *PowerMill User Guide*, Synopsys Inc., 1997.

[6] *RailMill User Guide*, Synopsys Inc., 1997.

[7] Golub, G and Van Loan, C. *Matrix Computations*. The Johns Hopkins Univ. Press, 1989.

[8] Star-RC Reference Manual, Avant! Corp, 1997.

[9] A. Dharchoudhury, R. Panda, D. Blaauw, R. Vaidyanathan, B. Tutuianu and D. Bearden, “Design and Analysis of power distribution networks in PowerPC microprocessors,” Proc. ACM/IEEE Design Automation Conf., pp. 738-743, 1998.

[10] H. Chen, D. Ling, “Power supply noise analysis methodolology for deep- submicron VLSI chip design”, Proc. Design Automation Conference, 1997.

[11] D. Blaauw, “Industrial perspectives on emerging CAD tools for low power processor design”, Invited Talk, ISLPED, 98.