

Global Harmony: Coupled Noise Analysis for Full-Chip RC Interconnect Networks

K. L. Shepard, V. Narayanan, P. C. Elmendorf¹, and Gutuan Zheng¹
IBM T. J. Watson Research Center, Yorktown Heights, NY 10598
¹IBM Microelectronics, Fishkill, NY

Abstract

Noise is becoming one of the most important metrics in the design of VLSI systems, certainly of comparable importance to area, timing, and power. In this paper, we describe Global Harmony, a methodology for the analysis of coupling noise in the global interconnect of large VLSI chips being developed for the design of high-performance microprocessors. The architecture of Global Harmony involves a careful combination of static noise analysis, static timing analysis, and reduced-order modelling techniques. We describe a reduced-order modelling approach that allows for passive multiport reduction of RC netlists as impedance macromodels while preserving the symmetry and sparsity of the state matrices for efficient storage. We describe how the macromodels are practically employed to perform coupling analysis and how timing constraints can be used to limit pessimism in the analysis.

1 Introduction

As CMOS technology scales into the deep submicron regime, noise immunity has become a metric for the analysis and design of VLSI systems of comparable importance to area, timing, and power. An essential component of the noise in VLSI chips comes from capacitive coupling between signal lines. The need to better understand and control coupling comes as scaling and performance requirements in general have created new demands for the analysis and design of interconnect networks.

Reduced-order modelling techniques have become an important method for analyzing linear interconnect networks. RLC analysis must be increasingly applied to analyzing clock trees, power buses, and off-chip interconnect to control clock skew and dI/dt noise. A more critical application for reduced-order modelling, however, remains the large RC netlists resulting from

the global chip interconnect extraction. These extractions are typically on the scale of tens of thousands of nets. AWE-based modelling of RC networks has been an essential part of timing analysis for more than five years[1]. Reduced-order modelling techniques must now be extended to consider the effects of coupling on noise and delay. There are three questions that need to be answered by this analysis, all potentially in the presence of coupled switching activity on other nets:

- What is the effective loading on each driver?
- What is the delay and slew at the receivers?
- What noise will be coupled at the receiver of an otherwise static net due to switching on other nets?

We describe a methodology, Global Harmony, which addresses these three issues and is being applied to the design of complex microprocessors.

The architecture we describe is a careful marriage of static timing analysis and static noise analysis[2]. Global Harmony consists of several important components – a reduced-order modelling algorithm that guarantees passive macromodelling of RC networks, a representation of this macromodel that allows for efficient storage and access, a timing rule structure that uses this macromodel with timing information to perform delay and noise analysis, and a noise analysis engine that considers temporal correlation of switching signals.

2 Overview of Global Harmony

The noise and timing methodology employed in the design of complex microprocessors is shown in Figure 1[3]. The shaded region in the figure denotes those pieces described in this paper as part of the architecture of Global Harmony.

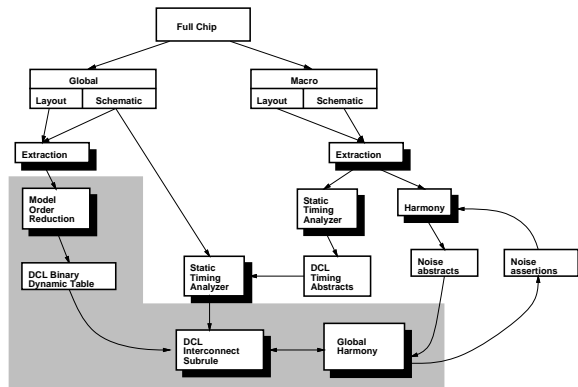


Figure 1: Global Harmony architecture

The timing and noise analysis methodology rely on a two-level hierarchical approach to practically manage the complexity of designs with tens of millions of transistors. This methodology involves identifying groups of 10,000 - 200,000 transistors as *macros*. Macros are individually laid out and floor-planned on the chip. They are timed using static timing analysis and abstracted in Delay Calculation Language (DCL)[4]. Similarly, static noise analysis (Harmony[2]) is performed on each macro and noise abstracts are generated for the global analysis. In some cases, noise assertions are returned back to Harmony analysis.

The global level consists of the interconnection network between the macros which contains all the long wire runs of the chip. The coupling capacitances between the macro and global levels are handled by treating them as worst-case hostile coupling sources in both the Harmony and Global Harmony analyses. The extraction of the global interconnect results in an RC networks that is reduced to a collection of multiport impedance macromodels, one for each net in the design, stored as a DCL binary dynamic table (BDT). The details of this reduction are described in Section 3. In Section 4, we describe how a DCL interconnect subrule model reads the BDT, folds in linearizations of the drivers and receivers and calculates the coupled noise. In Section 4, we describe the Global Harmony engine which computes the noise and flags violations, applying timing information obtained directly from the static timing environment. Section 5 describes in more detail the careful marriage of static timing and static noise analysis. In particular, we describe how timing information is used to reduce pessimism in coupled noise analysis and how the effects of noise on delay are considered in static timing analysis. Some results are briefly presented in Section 6.

3 Passive reduced-order modeling of coupled RC networks

The reduction part of Global Harmony begins with an extracted RC netlist for the global interconnect stored as a compressed binary file. The reduced-order modelling approach employed for Global Harmony guarantees passive, multiport macromodels with symmetry that allows for efficient storage of the results.

The first step in the reduction process is to identify a *net complex* for each net in the design. The *primary net* of the complex is the net on which we are trying to calculate the noise; that is, the net which should be statically quiet. The complex also includes *secondary nets* of significant coupling to the primary net. To determine which secondary nets to include in a complex, a coupling capacitance threshold ratio of total coupling to the secondary net in question to the self capacitance of the primary net, C_{coup}/C_{self} is used. Coupling capacitors with couplings below this threshold are treated as capacitors tied to ground. Couplings between the significant secondary nets and nets other than those already in the net complex are grounded. A representative net complex is shown in Figure 2(a).

We now describe the algorithm employed to reduce the state size of these RC net complexes. Modified nodal analysis (MNA) is used to stamp conductance and capacitance matrices according to the multi-input, multi-output, linear time-invariant differential equations:

$$\begin{aligned} C \dot{\mathbf{x}} &= -\mathbf{G}\mathbf{x} + \mathbf{B}\mathbf{i} \\ \mathbf{v} &= \mathbf{B}^T \mathbf{x} \end{aligned} \quad (1)$$

$\mathbf{x}, \mathbf{v}, \mathbf{i} \in \mathbb{R}^n$ are the state, output voltage, and input current vectors, respectively. For a system with n nodes and r ports, $\mathbf{G}, \mathbf{C} \in \mathbb{R}^{n \times n}$ are the symmetric, positive semidefinite conductance and capacitance matrices, respectively. The state vector is ordered so that the first r elements represent the port voltages. With this choice of ordering, the r -by- r matrix formed by the top r rows of the input-output matrix $\mathbf{B} \in \mathbb{R}^{n \times r}$ is the identity and the rest of the \mathbf{B} matrix is zero. Moving into the Laplace domain, Equations 1 lead to an expression for the r -by- r multiport impedance matrix for the net complex.

$$\mathbf{v}(s) = \mathbf{Z}(s)\mathbf{i}(s) \quad (2)$$

$$\mathbf{Z}(s) = \mathbf{B}^T(\mathbf{G} + s\mathbf{C})^{-1}\mathbf{B} \quad (3)$$

We choose impedance macromodelling over admittance macromodelling[5, 6, 7] because of the ease with

which we can fold linearized driver and receiver models into the analysis. Because a net complex in general does not have a DC path to ground, the impedance matrix is singular at $s = 0$. Because of this, the conductance matrix \mathbf{G} will in general be singular.

As a result, a nonzero expansion point s_o must be chosen for the moment matching. Using the change of variable $s = s_o + \tilde{s}$, Equation 3 becomes

$$\mathbf{Z}(s_o + \tilde{s}) = \mathbf{B}^T (\mathbf{G} + \tilde{s}\mathbf{C})^{-1} \mathbf{B} \quad (4)$$

where $\mathbf{G} = \mathbf{G} + s_o\mathbf{C}$. \mathbf{G} will be symmetric positive definite for a choice of real positive s_o . We then employ a multiport extension of the symmetric Lanczos process[8, 9] as shown in Algorithm 1 which is applicable to symmetric, positive-definite \mathbf{G} . The Cholesky factorizations can be performed because \mathbf{G} is positive definite and \mathbf{W} (after deflation) and \mathbf{B} have full column rank[10].

Algorithm 1 (Symmetric Lanczos)

```

lanczos(input  $\mathbf{C}, \mathbf{G}, \mathbf{B}, q$ ; output
 $\tilde{\mathbf{H}}_p, \tilde{\mathbf{B}}_p$ )
{
  Initialize:
  Set  $\mathbf{V}_0 = 0, \mathbf{T}_0 = 0, \beta_1 = \mathbf{I}$ 
  Cholesky factor:  $\rho^T \rho = \mathbf{B}^T \mathbf{G}^{-1} \mathbf{B}$ 
   $\mathbf{T}_1 = \mathbf{G}^{-1} \mathbf{B} \rho^{-1}, \mathbf{V}_1 = \mathbf{B} \rho^{-1}$ 
  for ( $j = 1; j \leq q; j++$ ) {
     $\mathbf{H}_{j,j} = \alpha_{j-1} = \mathbf{V}_j^T \mathbf{G}^{-1} \mathbf{C} \mathbf{T}_j$ 
     $\mathbf{W} = \mathbf{C} \mathbf{G}^{-1} \mathbf{V}_j - \mathbf{V}_j \alpha_{j-1} - \mathbf{V}_{j-1} \beta_{j-1}^T$ 
    Deflate  $\mathbf{W}$  to  $\mathbf{W}'$ 
    Cholesky factor:
     $\mathbf{R}^T \mathbf{R} = \mathbf{W}'^T \mathbf{G}^{-1} \mathbf{W}'$ 
     $\mathbf{V}_{j+1} = \mathbf{W}' \mathbf{R}^{-1}$ 
     $\mathbf{T}_{j+1} = \mathbf{G}^{-1} \mathbf{W}' \mathbf{R}^{-1}$ 
     $\mathbf{H}_{j,j+1} = \mathbf{H}_{j+1,j}^T = \beta_{j+1} = \mathbf{V}_{j+1}^T \mathbf{G} \mathbf{W}$ 
  }
   $\tilde{\mathbf{V}}_p = [\mathbf{V}_1 \mathbf{V}_2 \cdots \mathbf{V}_q]$ 
   $\tilde{\mathbf{B}}_p = [\rho; \mathbf{0}]$ 
   $\tilde{\mathbf{H}}_p = (\mathbf{H}_{i,j}), i, j = 1, \dots, q$ 
}

```

Algorithm 1 results in a reduced-order model of order p :

$$\mathbf{Z}_p(s_o + \tilde{s}) = \tilde{\mathbf{B}}_p^T \left(\mathbf{I} + \tilde{s} \tilde{\mathbf{H}}_p \right)^{-1} \tilde{\mathbf{B}}_p \quad (5)$$

where $\mathbf{Z}_p \in \mathbb{R}^{r \times r}$ and $p = p_1 + p_2 + \cdots + p_q$ where p_j is the rank of \mathbf{V}_j . $\mathbf{Z}_p(s)$ matches the first $2q$ coefficients

of $\mathbf{Z}(s)$ [11]. $\tilde{\mathbf{H}}_p \in \mathbb{R}^{p \times p}$ is a block tridiagonal matrix such that:

$$\tilde{\mathbf{H}}_p = \tilde{\mathbf{V}}_p^T \mathbf{G}^{-1} \mathbf{C} \mathbf{G}^{-1} \tilde{\mathbf{V}}_p \quad (6)$$

The deflation of \mathbf{W} performed as part of Algorithm 1 deserves additional comment. As the Lanczos iteration approaches convergence, the matrix \mathbf{W} will lose column rank[12]. This can happen easily for those net complexes in which it is not true that the number of nodes n is much greater than the number of ports r ($n \gg r$). In the Golub-Underwood block Lanczos procedure[13], the linearly independent columns of the \mathbf{W} matrix are replaced with randomly generated vectors orthogonalized to all the preceding columns of \mathbf{V}_j . We instead employ a deflation procedure similar to Cullum and Donath[14]. This means that the block size on successive iterations will decrease as the \mathbf{W} matrix loses rank.

The computational cost of Algorithm 1 is dominated by the factorization of the symmetric positive definite \mathbf{G} matrix. To perform this factorization efficiently, we make use of a modified version of the multifrontal algorithm[15] for sparse Cholesky factorization embodied in the Watson Symmetric Sparse Matrix Package (WSSMP)[16].

We choose an expansion frequency s_o for the reduction which is associated with a typical net Elmore delay. We generally use the same expansion frequency for all nets. We then monitor the convergence of the Frobenius norm of the matrix:

$$\tilde{\mathbf{B}}_p^T \left(\mathbf{I} + \tilde{\mathbf{B}}_p \mathbf{C}_{port} \tilde{\mathbf{B}}_p^T - s_o \tilde{\mathbf{H}}_p \right)^{-1} \tilde{\mathbf{B}}_p \mathbf{G}_{port} \quad (7)$$

to determine the order of the reduction to be used. \mathbf{G}_{port} is a diagonal conductance matrix with diagonal elements chosen to be a large default conductance. This equation follows directly from the results of Section 4. This is in lieu of other schemes that monitor the convergence of the eigenvalues[17] or the magnitude of the residue[18].

Passivity is the property that the network always dissipates energy and has been of considerable current interest[19]. It is a well-known result of network theory that the combination of two passive networks is guaranteed to be passive, while the combination of two asymptotically stable networks is not guaranteed to be asymptotically stable. Therefore, ensuring passivity of the macromodel is essential for ensuring the stability[20] of the final system used for noise analysis, which comes from combining the interconnect with the driver and receiver models. In fact, the driving point impedance \mathbf{Z}_p in Equation 5 is already passive. To prove this, we need to develop a few results. The first is a test for passivity:

Theorem 1 *A multiport linear interconnect network is passive if and only if the impedance matrix $\mathbf{Z}(s)$ satisfies:*

1. $\mathbf{Z}(s^*) = \mathbf{Z}^*(s)$, $\forall s \in \mathbb{C}$
2. $\mathbf{Z}(s)$ is positive; that is, $\mathbf{z}^{*T}(\mathbf{Z}(s) + \mathbf{Z}^T(s^*))\mathbf{z} \geq 0$, $\forall \mathbf{z} \in \mathbb{C}^p$ and $\forall s \in \mathbb{C}$ with $\text{Re}(s) > 0$.

The proof of Theorem 1 can be found in several reference on linear network theory[21, 22, 23]. In addition, to prove passivity, we will need the following result:

Theorem 2 *The matrix $\tilde{\mathbf{H}}_q$ generated by the symmetric Lanczos process is symmetric, positive semidefinite.*

Proof. Let \mathbf{x} be an arbitrary non-zero vector in \mathbb{R}^p . Then

$$\begin{aligned} \mathbf{x}^T \tilde{\mathbf{H}}_q \mathbf{x} &= \mathbf{x}^T \tilde{\mathbf{V}}_q^T \mathbf{G}^{-1} \mathbf{C} \mathbf{G}^{-1} \tilde{\mathbf{V}}_q \mathbf{x} \\ &= (\mathbf{G}^{-1} \tilde{\mathbf{V}}_q \mathbf{x})^T \mathbf{C} (\mathbf{G}^{-1} \tilde{\mathbf{V}}_q \mathbf{x}) \\ &= (\tilde{\mathbf{T}}_q \mathbf{x})^T \mathbf{C} (\tilde{\mathbf{T}}_q \mathbf{x}) \\ &\geq 0 \end{aligned}$$

The inequality results from the fact that \mathbf{C} is positive semidefinite. \square

We are now finally in the position to claim the following.

Theorem 3 *The reduced-order impedance macromodel \mathbf{Z}_p produced by the symmetric Lanczos process for RC circuits is passive.*

Proof. To prove this, it is necessary to prove the two conditions in Theorem 1. The first follows from the fact that $\tilde{\mathbf{B}}_p$ and $\tilde{\mathbf{H}}_p$ are all real. To prove the second condition of Theorem 1, it is necessary to show that:

$$\mathbf{z}^{*T}(\mathbf{Z}_p(s) + \mathbf{Z}_p(s^*))\mathbf{z} \geq 0 \quad \forall \mathbf{z} \in \mathbb{C}^p$$

for any s such that $\text{Re}(s) \geq 0$. From Equation 5 and taking $\tilde{s} = \sigma + i\omega$,

$$\begin{aligned} \mathbf{z}^{*T}(\mathbf{Z}_p(s) + \mathbf{Z}_p(s^*))\mathbf{z} &= \mathbf{z}^{*T} \tilde{\mathbf{B}}_p^T \left[(\mathbf{I} + \tilde{s} \tilde{\mathbf{H}}_p) + (\mathbf{I} + \tilde{s}^* \tilde{\mathbf{H}}_p)^T \right] \tilde{\mathbf{B}}_p \mathbf{z} \\ &= 2\mathbf{w}^{*T} (\mathbf{I} + \sigma \tilde{\mathbf{H}}_p) \mathbf{w} \\ &\geq 0 \end{aligned}$$

where $\mathbf{w} = (\mathbf{I} + \tilde{s}^* \tilde{\mathbf{H}}_p)^{-T} \tilde{\mathbf{B}}_p \mathbf{z}$. Because the matrix $\tilde{\mathbf{H}}_p$ is positive semidefinite by Theorem 2, $\mathbf{I} + \sigma \tilde{\mathbf{H}}_p$ is symmetric, positive definite for all $\sigma \geq 0$. \square

These interconnect macromodels are stored as DCL binary dynamic tables (BDT) which are subsequently utilized by a DCL interconnect subrule for noise analysis. We take advantage of the sparsity in storing the $\tilde{\mathbf{H}}_p$ and $\tilde{\mathbf{B}}_p$ matrices. The $\tilde{\mathbf{H}}_p$ matrix is symmetric and block tridiagonal. The $\tilde{\mathbf{B}}_p$ matrix is zero except for the top r -by- r which is upper triangular.

4 Interconnect modelling for noise

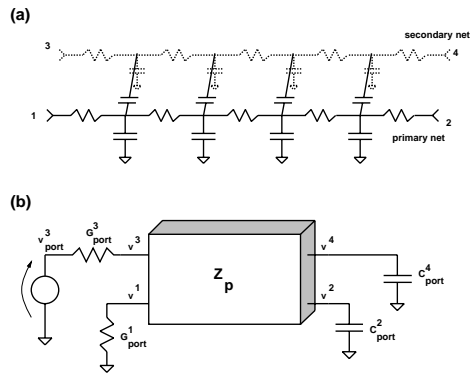


Figure 2: Multiport modelling of global interconnect. (a) A typical net complex consisting of a primary net coupled (in this case) to a single secondary net. (b) The driver resistances and receiver capacitances are folded into the multiport impedance macromodel.

Following reference [2], we classify noise by type. Noise that reduces the node voltage below the supply level is denoted V_H , while noise that increases a node voltage above the ground level is denoted V_L . Noise may also be bootstrapping if it increases a node voltage above the supply level (V_H^*) or below the ground level (V_L^*). As part of the macro-level static noise analysis, noise abstracts are generated for macro blocks. These noise abstracts contain linearized capacitances for each primary input and pull-up and pull-down driver resistances for each primary output. Drive resistances are determined by the linear region of the FET $I_D - V_{DS}$ characteristic. As part of the static noise analysis of the macro, the maximum amount of noise for each output that can be propagated for each of V_L and V_H noise types is calculated. For each input or output that is sensitive to noise, a *noise tolerance* is stored in the abstract by noise type. This is calculated using a sensitivity analysis at the first restoring logic gate from the pin in the macro-level static noise analysis[2].

The noise abstracts are then used along with the interconnect macromodels to check the noise on global interconnect. A DCL interconnect subrule performs the noise calculation from the macromodels loaded with the BDT. We first fold the driver resistance and receiver capacitances into the multiport impedance as shown in Figure 2(b). The Z_p macromodel is given by:

$$\mathbf{v} = \tilde{\mathbf{B}}_p^T (\mathbf{I} + \tilde{s} \tilde{\mathbf{H}}_p)^{-1} \tilde{\mathbf{B}}_p \mathbf{i} \quad (8)$$

where $\tilde{s} = s - s_o$. If we let \mathbf{x} denote the new state variable set, then the equations become:

$$\begin{aligned} \mathbf{x} &= \tilde{\mathbf{B}}_p \mathbf{i} - \tilde{s} \tilde{\mathbf{H}}_p \mathbf{x} \\ \mathbf{v} &= \tilde{\mathbf{B}}_p^T \mathbf{x} \end{aligned}$$

When receiver capacitances are added at the ports then the port currents become:

$$\mathbf{i} \rightarrow \mathbf{i} - s C_{port} \tilde{\mathbf{B}}_p^T \mathbf{x}$$

We now apply voltage sources \mathbf{v}_{port} at the ports through a diagonal port conductance matrix \mathbf{G}_{port} .

$$\mathbf{i} = \mathbf{G}_{port} (\mathbf{v}_{port} - \tilde{\mathbf{B}}_p^T \mathbf{x})$$

The conductance of the primary net driver is obtained from the drive resistance stored in the noise abstract. Diagonal conductance elements associated with capacitive receivers are zero, while the conductance associated with the secondary net drivers are set to a large default value. Using these expressions, one finds that the transfer function between the voltage sources and the ports of the interconnect are given by:

$$\mathbf{v} = \tilde{\mathbf{B}}_p^T (\mathbf{I} + s\mathbf{A})^{-1} (\mathbf{I} + \tilde{\mathbf{B}}_p \mathbf{G}_{port} \tilde{\mathbf{B}}_p^T - s_o \mathbf{H})^{-1} \tilde{\mathbf{B}}_p \mathbf{G}_{port} \mathbf{v}_{port} \quad (9)$$

where

$$\mathbf{A} = \left(\mathbf{I} + \tilde{\mathbf{B}}_p \mathbf{G}_{port} \tilde{\mathbf{B}}_p^T - s_o \mathbf{H} \right)^{-1} \left(\tilde{\mathbf{B}}_p C_{port} \tilde{\mathbf{B}}_p^T + \mathbf{H} \right).$$

\mathbf{A} is diagonalized as $\mathbf{S} \mathbf{\Lambda} \mathbf{S}^{-1}$, resulting in

$$\begin{aligned} \mathbf{v} &= \tilde{\mathbf{B}}_p^T \mathbf{S} (\mathbf{I} + s\mathbf{\Lambda})^{-1} \mathbf{S}^{-1} \left(\mathbf{I} + \tilde{\mathbf{B}}_p \mathbf{G}_{port} \tilde{\mathbf{B}}_p^T - s_o \mathbf{H} \right)^{-1} \\ &\quad \times \tilde{\mathbf{B}}_p \mathbf{G}_{port} \mathbf{v}_{port} \end{aligned} \quad (10)$$

Let $\boldsymbol{\mu}_i$ denote the i th column of $\tilde{\mathbf{B}}_p^T \mathbf{S}$ and $\boldsymbol{\nu}_i$ the i th row of

$$\mathbf{S}^{-1} \left(\mathbf{I} + \tilde{\mathbf{B}}_p \mathbf{G}_{port} \tilde{\mathbf{B}}_p^T - s_o \mathbf{H} \right)^{-1} \tilde{\mathbf{B}}_p \mathbf{G}_{port} \mathbf{v}_{port}.$$

Then

$$\mathbf{v} = \sum_i \frac{\boldsymbol{\mu}_i \boldsymbol{\nu}_i}{1 + s \lambda_i} \mathbf{v}_{port} \quad (11)$$

where $\boldsymbol{\mu}_i \boldsymbol{\nu}_i \in \mathbb{R}^{r \times r}$ and the λ_i are the diagonal elements of $\mathbf{\Lambda}$. The sum is over all the eigenvalues of \mathbf{A} . For a particular transfer function, one of the residues k_i is chosen from each $\boldsymbol{\mu}_i \boldsymbol{\nu}_i$ matrix. The time-domain response to a saturate ramp waveform of slew t_r such as that assumed to be at each switching secondary net driver, is given by:

$$v(t) = \begin{cases} m \sum_i [k_i \lambda_i (e^{-t/\lambda_i} - 1) - k_i t] & \text{for } 0 \leq t \leq t_r \\ \sum_i [m k_i \lambda_i (e^{-t/\lambda_i} - e^{-(t-t_r)/\lambda_i}) - k_i] & \text{for } t \geq t_r \end{cases} \quad (12)$$

where $m = V_{dd}/t_r$, the slope of the switching secondary net. Because of the *driver linearization* approximation, superposition applies.

5 The interaction of static timing and static noise analysis

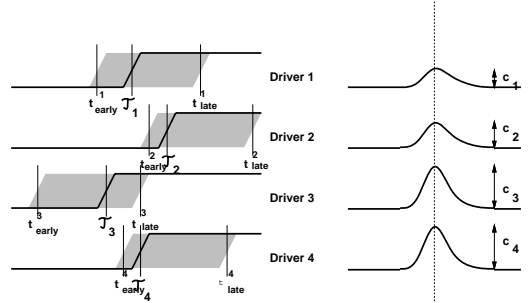


Figure 3: Timing orthogonality. The switching times τ_i are chosen so that the peak noises align.

The Global Harmony architecture shown in Figure 1 includes a tight coupling with the static timing analysis of the same design. This enables timing information to be used in the calculation of noise and noise information to be used in the calculation of timing. For the former case, we obtain secondary net driver slews from the timer model. Timing windows, as defined by the earliest and latest possible arrival time, are determined for each secondary net driver. This allows us to calculate the worst possible noise in the presence of arrival time constraints, reducing pessimism in the analysis. The problem can be formally stated as follows. Let c_i be the peak noise on a given primary receiver associated with driver i . Let t_{early}^i be the earliest arrival time associated with secondary driver i and let t_{late}^i be the latest arrival time associated with secondary driver i . In addition, let τ_i be the

switching time associated with secondary net driver i , such that all the noise peaks align for the primary receiver in question. Let x_i be the binary variable indicating whether the given secondary net driver is switching, and let n be the number of secondary nets. The problem is then to maximize:

$$\sum_{i=1}^n c_i x_i \quad (13)$$

such that the following $2n$ constraints can be satisfied for all i :

$$(t_{late}^i - \tau_i - t_{ref})x_i \geq 0 \quad (14)$$

$$(\tau_i + t_{ref} - t_{early}^i)x_i \geq 0 \quad (15)$$

where t_{ref} is a continuous variable determining the absolute time reference for the τ_i . In this form, the problem takes the form of a mixed integer programming problem. Alternately, the constraints can be reformulated to remove t_{ref} and consider only relative times. For all $i \neq j$,

$$(t_{late}^i - t_{early}^j - \tau_i + \tau_j)x_i x_j \leq 0 \quad (16)$$

$$(t_{late}^j - t_{early}^i + \tau_i - \tau_j)x_i x_j \leq 0 \quad (17)$$

We refer to these constraints as *timing orthogonality*. Because t_{early}^i and t_{late}^i result from early and late path propagation in static timing analysis, the timing windows incorporate the switching of the secondary-net drivers due to hazards.

This formulation assumes a certain “sharpness” to the noise peaks, since when the peak falls outside the windows, its contribution is taken as zero. We utilize a branch-and-bound algorithm[24, 25] to solve this problem since the noise on each subtree can be easily bounded by the assumption that each node in that subtree is contributing. The maximum noise of Equation 13 is added to the propagate noise from the noise abstracts for each receiver and compared against the noise margins also contained in the noise abstracts. A noise “slack” report results. These slacks are based on pessimistic dc noise margins at the macro inputs. To eliminate this pessimism, one can perform a Harmony run on the macro using assertions of the actual input noise generated from Global Harmony.

Even when the entire design satisfies the condition of *noise stability*[2] as verified by static noise analysis, coupled noise in the interconnect can still have a significant effect on timing. The coupled noise calculation is generally performed only after a converged timing solution is calculated. The algorithm employed in Global Harmony to handle the effects of noise on delay is essentially an iterative one:

1. Perform an initial timing with all secondary nets grounded.
2. Freeze the arrival time windows and slews at each driver.
3. Reset the timing model and recalculate all delays including the effects of coupling in the interconnect analysis. The *frozen* secondary net driver information is used for this analysis.
4. Go to step 2 and repeat until convergence.

We now consider the calculations performed as part of step 3. Following the usual practice in static timing analysis, we first find the waveform at the driver, abstract it as a saturate ramp, and propagate it forward to each of the net receivers. To find the driver waveform, we calculate:

$$i_{port} = G_{port} \left[\mathbf{I} - \tilde{\mathbf{B}}_p^T (\mathbf{I} + s\mathbf{A})^{-1} (\mathbf{I} + \tilde{\mathbf{B}}_p G_{port} \tilde{\mathbf{B}}_p^T - s_0 \mathbf{H}_q)^{-1} \tilde{\mathbf{B}}_p G_{port} \right] v_{source}$$

G_{port} in this case differs from that used in Equation 9 in that a very small primary driver resistance is used to effectively attach the driver output voltage waveform directly to the multiport. Diagonalizing \mathbf{A} as in Equations 10 and 11 yields the following expression for the driver current:

$$i_{driver}(t) = \sum_n \sum_i m_n \left[-k_i^n \lambda_i (e^{-(t-t_i)/\lambda_i} - 1) u(t - t_i) + k_i^n \lambda_i (e^{-(t-t'_i)/\lambda_i} - 1) u(t - t'_i) \right]$$

where the outer sum is over the primary and secondary drivers and the inner sum is over the eigenvalues of \mathbf{A} . The waveforms at the primary and secondary drivers are assumed to be saturate ramps with slopes m_i which begin at time t_i and end at time t'_i (i. e. $t'_i - t_i = |V_{dd}/m_i|$). k_i^n is the residue due to the i th pole with the n th driver contributing. Opposite direction switching (negative m_i) produces the worst-case load, while same-direction switching (positive m_i) produces the best-case load. The driver waveform is first calculated for the case of grounded secondary nets. Secondary net arrival times (i. e. the t_i) are chosen so that the peak noise at the primary driver, as determined by the techniques of Section 4 arrives at the 50% response point of this driver waveform. This heuristic seems to work well in determining the worst (or best) response. We then use an effective-C[26] iteration to find the driver waveform, because of the use of driver delay characterizations based on lumped capacitance. In the presence of a more general nonlinear

driver model, recursive convolution techniques[5, 27] can be used.

Calculating the delay and slew follows similarly. In this case, each receiver waveform is calculated for the case of grounded secondary nets and then the secondary net arrival times are chosen so that the peak noise arrives at the 50% response point of this receiver waveform.

6 Results

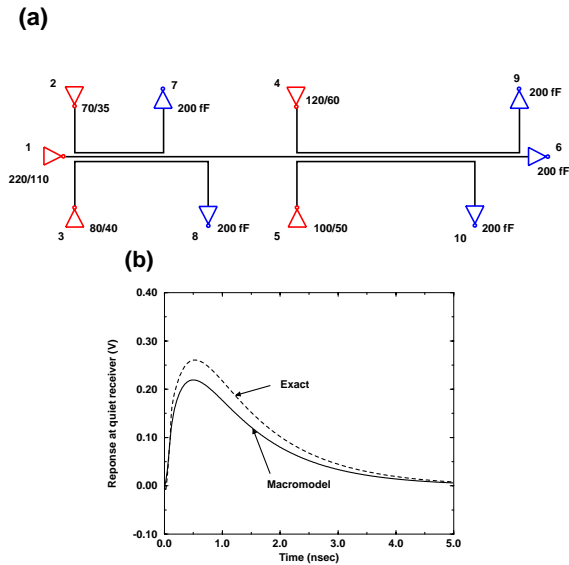


Figure 4: Test interconnect structure (a) The net complex consisting of a primary net and four secondary nets. (b) The response on the primary net receiver 6 due to the switching of driver 4. Both the exact and macromodel responses are shown.

Figure 4(a) shows one of the test structure used to verify our interconnect analysis. This is a net complex consisting of four secondary nets. The interconnect is $0.8\mu\text{m}$ -thick aluminum of $0.45\mu\text{m}$ width and $0.45\mu\text{m}$ spacing. The primary net is 5mm long. Driver sizes (as widths in microns) and loads are also indicated in the figure. There are 129 state variables in the original network. The Z_p macromodel used in this design has $p = 10$ and $s_o = 10^8\text{sec}^{-1}$. Figure 4(b) shows the exact and approximate noise waveshapes at the primary net driver due to a switching driver 4 with a 100-psec slew.

In Figure 5, we shows typical “noise slack” results in histogram form for a Global Harmony run on a

section of a high-performance CMOS microprocessor with 7714 receivers. The noise tolerance at each input is set to zero for this run so that the full spectrum of the coupling noise can be observed. The supply voltage is 1.7 V.

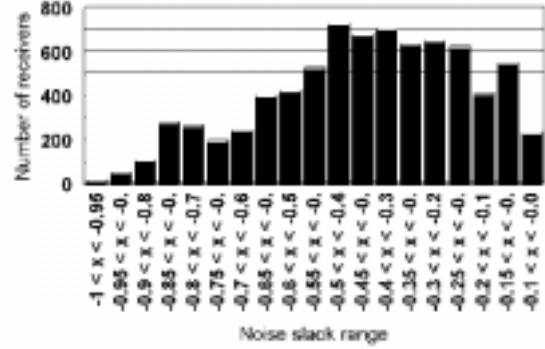


Figure 5: Noise slack histogram for a section of a high-performance CMOS microprocessor. Noise slacks are in volts. The supply voltage is 1.7 V.

Figure 6 shows a noise-on-delay calculation for a reasonably typical point-to-point global net, 1.3 mm long, coupled to eleven secondary nets. With worst-case simultaneous switching of the secondary nets, the effective capacitance increases from 170fF to 250fF . The net delay increases from about 130ps to 200ps .

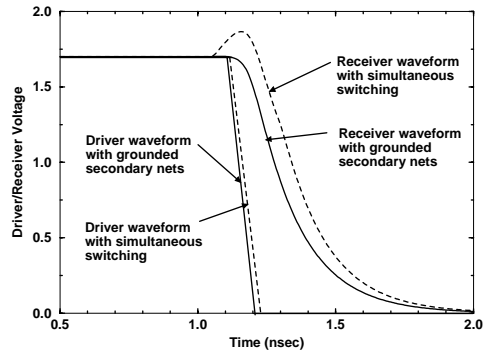


Figure 6: Effect of coupling on interconnect response of typical global net

7 Conclusions

In this paper, we have described a methodology for performing global noise analysis as part of a sta-

tic noise analysis methodology. This incorporates a unique combination of timing and noise analysis and employs a reduced-order modelling algorithm that allows for passive interconnect macromodelling and efficient storage of the macromodel result.

8 Acknowledgements

The authors would like to acknowledge many useful discussion with Abe Elfadel, Eli Chiprout, Anshul Gupta, Chandu Visweswariah, Joe Rahmeh, and Paul Villarrubia. The authors also acknowledge the invaluable contributions of John Beatty to DCL and Alex Sues to the static timing environment. The authors also acknowledge the contributions of the microprocessor design teams in Poughkeepsie and Yorktown. We also thank Abe Elfadel for a careful reading of the manuscript.

References

- [1] L. T. Pillage and R. A. Rohrer. Asymptotic Waveform Evaluation for Timing Analysis. *IEEE Trans. CAD*, 9(4):352–366, April 1990.
- [2] K. L. Shepard and V. Narayanan. Noise in deep submicron digital design. In *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pages 524–531, San Jose, CA, November 1996.
- [3] K. L. Shepard, S. Carey, E. Cho, B. Curran, R. Hatch, D. Hoffman, S. McCabe, G. Northrop, and R. Seigler. Design Methodology for the G4 S/390 Microprocessors. *IBM Journal of Research and Development*, 1997. To be published.
- [4] Delay Calculation Language (DCL) and Procedural Interface (PI). Technical report, CAD Framework Initiative. Version 1.0.1.
- [5] Vivek Raghavan, J. Eric Bracken, and Ronald A. Rohrer. AWESpice: A General Tool for the Accurate and Efficient Simulation of Interconnect Problems. In *29th ACM/IEEE Design Automation Conference*, pages 87–92, Anaheim, California, June 1992.
- [6] Seok Yoon Kim, Nanda Gopal, and Lawrence T. Pillage. AWE Macromodels of VLSI Interconnect for Circuit Simulation. In *Proceedings of the IEEE/ACM International Conference on Computer Aided-Design*, pages 64–70, Santa Clara, California, November 1992.
- [7] Seok-Yoon Kim, Nanda Gopal, and Lawrence T. Pillage. Time-domain macromodels for VLSI interconnect analysis. *IEEE Trans. CAD*, 13:1257 – 1270, 1994.
- [8] R. W. Freund and P. Feldmann. Reduced-order modeling of large passive linear circuits by means of the SyPVL algorithm. In *ICCAD'96*, pages 280–287, San Jose, California, November 1996.
- [9] Roland W. Freund and Noël M. Nachtigal. Software for simplified Lanczos and QMR algorithms. *Applied Numerical Mathematics*, 19:319–341, 1995.
- [10] G. H. Golub and C. F. Van Loan. *Matrix Computation*. The Johns Hopkins University Press, third edition, 1996.
- [11] Peter Feldmann and Roland W. Freund. Reduced-order modeling of large linear subcircuits via a block lanczos algorithm. In *32nd ACM/IEEE Design Automation Conference*, pages 474–479, San Francisco, California, June 1995.
- [12] Jane K. Cullum and Ralph A. Willoughby. *Lanczos Algorithms for Large Symmetric Eigenvalue Computations, Volume I: Theory*. Birkhaeuser, 1985.
- [13] G. H. Golub and R. Underwood. The block lanczos method for computing eigenvalues. In J. R. Rice, editor, *Mathematical Software III*, pages 361–377. Academic Press, New York, 1977.
- [14] J. Cullum and W. E. Donath. A block lanczos algorithm for computing the q algebraically largest eigenvalues and a corresponding eigenspace for large, sparse symmetric matrices. In *Proceeding of the IEEE Conference on Decision and Control*, pages 505–509, 1974.
- [15] Joseph W. H. Liu. The multifrontal method for sparse matrix solution: Theory and practice. *SIAM Review*, 34:82–109, 1992.
- [16] Anshul Gupta and Fred Gustavson. WSSMP: Watson symmetric sparse matrix package. Technical report, IBM T.J. Watson Research Center, 1997. Research Report RC 20645.
- [17] P. Feldmann and R. W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos process. *IEEE Trans. CAD*, 14:639–649, May 1995.
- [18] I. J. Jaimoukha and E. M. Kasenally. Oblique projection methods for large scale model reduction. *SIAM J. Matrix Anal. Appl.*, 16:602 – 627, 1995.
- [19] I. Elfadel and D. D. Ling. Zeros and passivity of arnoldi-reduced-order transfer functions for interconnect networks. In *34th DAC*, 1997. to appear.
- [20] L. Miguel Silveira, Mattan Kamon, Ibrahim Elfadel, and Jacob White. A Coordinate-Transformed Arnoldi Algorithm for Generating Guaranteed Stable Reduced-Order Models of RLC Circuits. In *IEEE/ACM International Conference on Computer-Aided Design*, pages 288 – 294, San Jose, CA, November 1996.
- [21] E. S. Kuh and R. A. Rohrer. *Theory of Linear Active Networks*. Holden-Day, 1967.
- [22] E. A. Guillemin. *Synthesis of Passive Networks*. John Wiley and Sons, 1957.
- [23] K. S. Narendra and J. H. Taylor. *Frequency Domain Criterion for Absolute Stability*. Academic Press, 1973.
- [24] G. Nemhauser and L. Wolsey. *Integer and Combinatorial Optimization*. Wiley, 1988.
- [25] C. Papadimitriou and K. Steiglitz. *Combinatorial Optimization*. Prentice-Hall, 1982.
- [26] Curtis L. Ratzlaff, Satyamurthy Pullela, and Lawrence T. Pillage. Modelling the RC-interconnect effects in a hierarchical timing analyzer. In *Custom Integrated Circuits Conference*, pages 15.6.1 – 15.6.4, 1992.
- [27] J. E. Bracken, V. Raghavan, and R. A. Rohrer. Interconnect Simulation with Asymptotic Waveform Evaluation. *IEEE Trans. Circuits Syst.*, 39(11):869–878, November 1992.