Purdue University Purdue e-Pubs

ECE Technical Reports

Electrical and Computer Engineering

1-1-2002

On-chip Interconnect Modeling by Wire Duplication

Guoan Zhong

Cheng-Kok Koh

Kaushik Roy

Follow this and additional works at: http://docs.lib.purdue.edu/ecetr

Zhong, Guoan ; Koh, Cheng-Kok ; and Roy, Kaushik , "On-chip Interconnect Modeling by Wire Duplication" (2002). *ECE Technical Reports*. Paper 164. http://docs.lib.purdue.edu/ecetr/164

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

On-chip Interconnect Modeling by Wire Duplication *

Guoan Zhong, Cheng-Kok Koh, and Kaushik Roy School of Electrical and Computer Engineering 1285 Electrical Engineering Building Purdue University West Lafayette, IN 47907-1285 {zhongg,chengkoh,kaushik}@ecn.purdue.edu

August 2, 2002

^{*}This research is supported in part by SRC (99-TJ-689), NSF (CAREER Award CCR-9984553), and a grant from Intel Corporation.

Abstract

In this paper, we present a novel wire duplication-based interconnect modeling technique. The proposed modeling technique uses the original inductance to exploit the sparsity of the L^{-1} matrix, where L is the inductance matrix, and construct a sparse and stable equivalent RLC circuit out of the windowed inductance matrices. The model avoids matrix inversions. Most important, it is more accurate and more efficient than many popular techniques in the literature.

1 Introduction

With the continual increasing of clock frequency and global interconnect length and decreasing of signal transition time, accurate modeling of inductance effects become increasingly more important. The partial inductance matrix L obtained from the PEEC model [8] is extremely large and dense. Direct simulation of the full L matrix is very time-consuming and memory-consuming. To make the simulation more efficient, sparsification of L and L^{-1} matrices has been considered in [7, 5, 3, 6, 1, 2].

One sparsification approach is to discard the mutual coupling terms that are below some threshold. However, the resulting inductance matrix may not be positive definite; that leads to an unstable circuit. The shift-truncate method proposed in [7, 5] can guarantee that the generated sparse inductance matrix is positive definite. However, the accuracy is not satisfactory [3, 6].

[3] demonstrates the locality of L^{-1} . It shows that the matrix L^{-1} has a similar locality property as the capacitance matrix. Hence, the L^{-1} matrix can be easily sparsified by dropping small entries while stability is guaranteed. Thus, modeling the inductance with the truncated L^{-1} matrix (denoted by L^{-1}) instead of the *L* matrix can reduce the coupling elements and speed up the simulation. In [6], a new circuit element *K*, which is defined as the inverse of inductance, is introduced and is incorporated in a simulation tool (known as *K* method). To avoid the *K* element in simulation, We can invert the truncated L^{-1} matrix to obtain a new inductance matrix (denoted by L). As L is also a dense matrix, direct simulation of L (referred to as L method) is not efficient. [1] performs sparsification on the L matrix(known as the double-inverse inductance model). Essentially, the double-inverse inductance model requires two approximation (sparsification) steps. [2] calculates the sparse inductance matrix directly by using exponential potentials and matrix inversions are avoided.

In this paper, we present a novel *wire duplication* interconnect modeling technique. This technique is motivated by the mathematical property that only a subset of the entries of the the double-inverse \pounds matrix is required to reconstruct the \pounds^{-1} matrix. Consequently, we can construct an circuit that is equivalent to the \pounds^{-1} matrix out of the subset of \pounds by *wire duplication*. It is stable, sparse and as accurate as the *K* method [3, 6]. Furthermore, we can apply the wire duplication technique to the original inductances directly. Thus, matrix inversions are avoided. Most important, the accuracy is improved.

The following notation is used in the paper:

- L: The original partial inductance matrix.
- L^{-1} : The inverse of L.
- L^{-1} : Truncated L^{-1} .
- L: The double-inverse inductance matrix; the inverse of truncated L^{-1} .
- Ł method: The method that uses Ł instead of *L* in the simulation.
- WD/Ł: The wire duplication model using the double-inverse inductance matrix Ł.

• WD/L: The wire duplication model using the original inductance matrix L.

2 Mathematical Background

Here we present the mathematical property that the \pounds matrix contains redundant information. Consequently, we may use only the central band of \pounds to reconstruct the \pounds^{-1} matrix. That is the key to the proposed *wire duplication* method. The theorems and the proofs behind the mathematical property are given in the appendix.

Let *A* be a $N \times N$ band matrix with bandwidth equal to 2b + 1, and $B = A^{-1}$. We take rows i - b to i + b and columns i - b to i + b of *B* to form a sub-matrix. Then, the center row and center column of the inverse of the sub-matrix are identical to the *i*th row and *i*th column of the *A* matrix:

$$A(i,i-b:i+b) = (B(i-b:i+b,i-b:i+b))^{-1}(b+1,:),$$

$$A(i-b:i+b,i) = (B(i-b:i+b,i-b:i+b))^{-1}(:,b+1).$$
(1)

Here we use the following notation: A(i : j, m : n) refers to the sub-matrix at the intersection of rows *i* to *j* and columns *m* to *n* of A; A(:,m) refers to column *m* and A(i,:) refers to row *i*; A(:,m:n) refers to columns from *m* to *n* and A(i:j,:) refers to rows from *i* to *j*. The index of the matrix begins at 1. If i - b < 1 or i + b > N, we use 1 or *N* instead.

We illustrate the mathematical property using a layout of 7 parallel and aligned wires. The wire length is 100μ m, and the cross section is 0.5μ m×1 μ m. The separation between the wires is 0.5μ m. The *L* matrix and its inverse (L^{-1} matrix) are:

$$L = 10^{-11} \times$$

$$\begin{bmatrix} 10.8 & 8.51 & 7.22 & 6.45 & 5.90 & 5.47 & 5.13 \\ 8.51 & 10.8 & 8.51 & 7.22 & 6.45 & 5.90 & 5.47 \\ 7.22 & 8.51 & 10.8 & 8.51 & 7.22 & 6.45 & 5.90 \\ 6.45 & 7.22 & 8.51 & 10.8 & 8.51 & 7.22 & 6.45 \\ 5.90 & 6.45 & 7.22 & 8.51 & 10.8 & 8.51 & 7.22 \\ 5.47 & 5.90 & 6.45 & 7.22 & 8.51 & 10.8 & 8.51 \\ 5.13 & 5.47 & 5.90 & 6.45 & 7.22 & 8.51 & 10.8 \end{bmatrix}$$

$$L^{-1} = 10^{10} \times$$

2.54	-1.68	-0.13	-0.12	-0.08	-0.06	-0.11
-1.68	3.65	-1.60	-0.05	-0.07	-0.04	-0.06
-0.13	-1.60	3.65	-1.60	-0.05	-0.07	-0.08
-0.12	-0.05	-1.60	3.66	-1.60	-0.05	-0.12
-0.08	-0.07	-0.05	-1.60	3.65	-1.60	-0.13
-0.06	-0.04	-0.07	-0.05	-1.60	3.65	-1.68
-0.11	-0.06	-0.08	-0.12	-0.13	-1.68	2.54

If we drop the small off-diagonal terms in L^{-1} , we obtain a band matrix with bandwidth 3:

$$\begin{split} \mathbf{L}^{-1} &= 10^{10} \times \\ \begin{bmatrix} 2.54 & -1.68 \\ -1.68 & 3.65 & -1.60 \\ & -1.60 & 3.65 & -1.60 \\ & & -1.60 & 3.66 & -1.60 \\ & & & -1.60 & 3.65 & -1.60 \\ & & & & -1.60 & 3.65 & -1.68 \\ & & & & & -1.68 & 2.54 \\ \end{bmatrix} \end{split}$$

.

If we invert the L^{-1} matrix, we obtain the L matrix:

Note that \underline{k} is a full dense matrix, but from the theorem, we know that we can reconstruct the \underline{k}^{-1} matrix from only the boldface entries in the \underline{k} matrix. Here we illustrate how we can reconstruct the 1st and 3rd rows (columns) of the \underline{k}^{-1} matrix. In the rest of the paper, we drop the orders of \underline{k}^{-1} and \underline{k} for a more concise presentation.

To obtain the 1st row and column of the L^{-1} matrix, we use the (1:2,1:2) window of L:

$$l = \mathbb{E}(1:2,1:2) = \begin{bmatrix} 6.73 & 4.20 \\ 4.20 & 6.32 \end{bmatrix}.$$

Inverting l, we obtain

$$l^{-1} = \begin{bmatrix} 2.54 & -1.68 \\ -1.68 & 2.70 \end{bmatrix}$$

whose first row and column correspond to the first row and column of the L^{-1} matrix.

Similarly, for the 3^{rd} row and column of the L^{-1} matrix, we use the (2:4,2:4) window of L:

$$l = \mathbf{L}(2:4,2:4) = \begin{bmatrix} 6.32 & 3.75 & 2.23 \\ 3.75 & 5.92 & 3.52 \\ 2.23 & 3.52 & 5.81 \end{bmatrix}$$
$$l^{-1} = \begin{bmatrix} 2.53 & -1.60 & -0.00 \\ -1.60 & 3.65 & -1.60 \\ -0.00 & -1.60 & 2.69 \end{bmatrix}.$$

,

Thus, the central band (with bandwidth 4b+1) of the Ł matrix contains all the information in the L^{-1} matrix.

3 Wire Duplication Model

In the previous section we demonstrated that the information of L^{-1} is contained in the central band of L. The next step is to build an equivalent circuit out of the entries in this band and ignore the remaining entries.

The following equation describes the magnetic couplings between the wires in the layout example in the previous section with the L^{-1} matrix:

$$\frac{d}{dt} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \\ I_5 \\ I_6 \\ I_7 \end{bmatrix} = \begin{bmatrix} 2.54 & -1.68 & & & & \\ -1.68 & 3.65 & -1.60 & & & \\ & -1.60 & 3.65 & -1.60 & & & \\ & & -1.60 & 3.65 & -1.60 & & \\ & & & -1.60 & 3.65 & -1.60 & & \\ & & & & -1.60 & 3.65 & -1.68 & \\ & & & & & -1.68 & 2.54 \end{bmatrix} \begin{bmatrix} V_{l1} \\ V_{l2} \\ V_{l3} \\ V_{l4} \\ V_{l5} \\ V_{l6} \\ V_{l7} \end{bmatrix}.$$
(2)

where V_{lk} and I_k refer to the voltage drop due to the inductance and the current in the wire k respectively.

5

We can rewrite the preceding equation in terms of the double-inverse inductance matrix Ł:

V_{l1}		6.73	4.20	2.49	1.48	0.90	0.57	0.38		I_1
V_{l2}		4.20	6.32	3.75	2.23	1.35	0.86	0.57		I_2
V_{l3}		2.49	3.75	5.92	3.52	2.13	1.35	0.90	,	I_3
V_{l4}	=	1.48	2.23	3.52	5.81	3.52	2.23	1.48	$\frac{d}{dt}$	I_4
V_{l5}		0.90	1.35	2.13	3.52	5.92	3.75	2.49		I_5
V_{l6}		0.57	0.86	1.35	2.23	3.75	6.32	4.20		I_6
V_{l7}		0.38	0.57	0.90	1.48	2.49	4.20	6.73		<i>I</i> 7

Now, we shall show how an equivalent circuit can be constructed out of the windows of the Ł matrix. For example, if we take the window corresponding to the (2:4,2:4) sub-matrix of the Ł matrix, and apply them to wires 2, 3 and 4, we have:

$$\begin{bmatrix} V_{l2} \\ V_{l3} \\ V_{l4} \end{bmatrix} = \begin{bmatrix} 6.32 & 3.75 & 2.23 \\ 3.75 & 5.92 & 3.52 \\ 2.23 & 3.52 & 5.81 \end{bmatrix} \frac{d}{dt} \begin{bmatrix} I_2 \\ I_3 \\ I_4 \end{bmatrix},$$
(3)

or

$$\frac{d}{dt} \begin{bmatrix} I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} 2.53 & -1.60 & -0.00 \\ -1.60 & 3.65 & -1.60 \\ -0.00 & -1.60 & 2.69 \end{bmatrix} \begin{bmatrix} V_{l_2} \\ V_{l_3} \\ V_{l_4} \end{bmatrix}.$$
(4)

Among the three circuit equations for I_2 , I_3 , and I_4 in Eqn.(4), only the following equation

$$\frac{dI_3}{dt} = -1.60V_{l2} + 3.65V_{l3} - 1.60V_{l4} \tag{5}$$

matches that in Eqn.(2). Hence, we can model wire 3 correctly, provided that V_{l2} and V_{l4} are correct. However, the equations for I_2 and I_4 in Eqn.(4) do not match those in Eqn.(2), i.e., wires 2 and 4 are not correctly modeled. Thus, their voltages V_{l2} and V_{l4} are incorrect. To provide a remedy to this problem, we can model these two wires correctly somewhere else and use the correct V_{l2} and V_{l4} values for the modeling of wire 3 here.

Figure 1 shows the modeling of signal 3. In this figure, the diamond sign \diamond stands for voltage controlled voltage source (VCVS) element. The two VCVSs provide the correct voltages for L_2 and L_4 . The inputs of the VCVSs come from the correct modelings of L_2 and L_4 . Since L_2 and L_4 here are controlled by their corresponding correct modelings, they are just dummy copies. We call such copies *dummy wires* and draw them in dashed lines. In contrast, if a wire



Figure 1: Modeling of wire 3.

is correctly modeled, we call it a *real wire* and draw it in solid lines. Here wire 3 is a real wire and wires 2 and 4 are dummy wires. The real wire and the dummy wires form a *group*. The total number of wires in a group is called the *group size* or *window size*. Figure 1 shows such a group, which models wire 3 correctly. Similarly, we can construct a group that includes dummy copies of wire 1 and wire 3 to model wire 2 correctly. In the group that models wire 4 correctly, dummy copies of wire 3 and wire 5 are included. Real wires 2 and 4 in these two groups provide the correct voltages V_{12} and V_{14} for the VCVSs in Figure 1.

In general, only one wire is correctly modeled in one group; so we need N groups for N wires in the simulation. There are one real wire and 2b dummy wires in each group if the bandwidth of L^{-1} is 2b + 1 (the groups at two ends are special cases).

In each group, every pair of wires (including both real and dummy ones) are inductively coupled, and there is no inductive coupling among groups. Let \tilde{L} be the partial inductance matrix for the wire duplication model, then \tilde{L} is block diagonal and each block corresponds to one group. \tilde{L}^{-1} is also block diagonal. If we remove the rows for dummy wires and utilize the fact that dummy wires have the same voltages as the corresponding real wires, we get back the L^{-1} matrix, which is positive-definite [6, 1, 2]. Thus, the circuit obtained by wire duplication is stable.

We use HSPICE to verify the correctness of this *wire duplication* model. We refer to the wire duplication model using Ł matrix as the WD/Ł model. Two sets of simulation are carried out: one set uses Ł method, the other uses the WD/Ł model. They produce *exactly* the same results (See Section 6 for details). It indicates that the WD/Ł model is also numerically equivalent to the Ł model. This is expected, as they are physically and mathematically equivalent. Since Ł model and *K* model [3] are equivalent, the wire duplication model is also equivalent to the *K* model [3].

4 Optimize the Group Size

In the wire duplication model described in the previous section, there are 2b + 1 wires in each group (less than 2b at the two ends). There are altogether about $N \cdot (2b+1) \cdot b$ inductive couplings, whereas the full inductance matrix contains



Figure 2: Modeling of wires 1 and 2.

 $N \cdot (N-1)/2$ couplings. If $b \ll N$, the wire duplication technique will produce an equivalent circuit of a smaller size. There are two methods to reduce the circuit size even further. The first method merges the groups at the ends. The following window captures the modeling of wire 2:

$$l = \pounds(1:3,1:3) = \begin{bmatrix} 6.73 & 4.20 & 2.49 \\ 4.20 & 6.32 & 3.75 \\ 2.49 & 3.75 & 5.92 \end{bmatrix},$$
$$l^{-1} = \begin{bmatrix} 2.54 & -1.68 & 0.00 \\ -1.68 & 3.65 & -1.60 \\ 0.00 & -1.60 & 2.70 \end{bmatrix}.$$

We can see that wire 1 is also correctly modeled. It means that wire 1 and 2 can share one group, as shown in Figure 2. Similarly, wires N - 1 and N can share one group.

Such an improvement is marginal; the second method, which uses larger windows, can achieve more reduction. For example, if we use a window of size 4, for the (1:4,1:4) window, the corresponding matrices are:

$$l = \pounds(1:4,1:4) = \begin{bmatrix} 6.73 & 4.20 & 2.49 & 1.48 \\ 4.20 & 6.32 & 3.75 & 2.23 \\ 2.49 & 3.75 & 5.92 & 3.52 \\ 1.48 & 2.23 & 3.52 & 5.81 \end{bmatrix}$$



Figure 3: Simulation times for different window sizes.

$$l^{-1} = \begin{bmatrix} 2.54 & -1.68 & 0.00 & -0.00 \\ -1.68 & 3.65 & -1.60 & -0.00 \\ 0.00 & -1.60 & 3.65 & -1.60 \\ -0.00 & -0.00 & -1.60 & 2.69 \end{bmatrix}$$

In this case, wires 1, 2, and 3 are correctly modeled in this group with a dummy copy of wire 4.

With larger windows, fewer groups are needed to model all the wires. However, this reduction is at the expense of more couplings within each group; the number of couplings in each group increases quadratically with the window size. We discuss the trade-off in the remainder of this section.

For simplicity, we assume that all the groups are of the same size B. The number of wires commonly found in two adjacent groups should be 2b. Let n be the number of groups needed. Then,

$$n \cdot B - 2b(n-1) = N \Rightarrow n = \frac{N - 2b}{B - 2b}$$
(6)

The number of total wires used is $B \cdot n$; the number of total couplings is $\frac{B \cdot (B-1)}{2}n$. The purpose of this study is to build an equivalent circuit with a smaller size. It includes both wires and the coupling elements. As a rough estimate, we use $B^2 \cdot n/2$ as the circuit size and try to minimize it. We can easily conclude that $\frac{B^2 \cdot n}{2} = \frac{B^2(N-2b)}{2(B-2b)}$ is minimized when

$$B = 4b, \tag{7}$$

and the minimal value of $B^2 \cdot n/2$ is

$$(B^2 \cdot n/2)_{min} = 4b(N - 2b).$$
(8)

For the circuit example in Section 6, we set the bandwidth of L^{-1} to be 5, i.e., b = 2, and perform wire duplication simulations for different window sizes (from 5 through 12). The run times are shown in Figure 3. We can see that the



Figure 4: Eigenvalues for full matrix, wire duplication using Ł matrix, and wire duplication using L matrix.

circuit obtained with a window size of 8, which is 4b, has the smallest run time. That coincides with our estimation.

Although the circuit sizes at different window sizes are different, all the resulting circuits are physically and mathematically equivalent. The simulation results for all of them are *exactly* the same.

When B = 4b, there are about (4b - 1)(N - 2b) inductive couplings and 2(N - 2b) wires (N - 4b of them are dummy wires). Note that the *K* method has about $b \cdot N$ couplings and requires no dummy wires. Therefore, the wire duplication technique uses about four times couplings required in the K method and an additional N - 2b dummy wires. This is the price that we pay for RLC simulation instead of RKC simulation.

5 Wire Duplication Using *L* Matrix

In the previous sections, we use the double-inverse inductance \pounds in the wire duplication model. It turns out that we can use the original inductance *L* directly in the modeling to avoid matrix inversions. An additional and more significant benefit is that the accuracy is also improved. We refer to the wire duplication model using *L* matrix as the WD/L model.

The use of windowed Ł matrix is based on strict mathematical property. That mathematical property can also explain the validity of using windowed original matrix *L* to a certain extent, as L^{-1} is almost a band matrix. From an engineering point of view, we can use windowed *L* matrix to calculate the L^{-1} [6, 2]. It means that a window of the *L* matrix can also capture the magnetic couplings as the Ł matrix. We have shown that WD/Ł is stable. For WD/L, we can also prove its stability in a similar way.

There is another benefit of using the *L* matrix instead of the Ł matrix: it is more accurate. If we use the 4*b* window size (as suggested in Section 4), we can capture about 4b - 1 inductive couplings between each wire and its neighbors. In contrast, the WD/Ł model can only capture 2*b*. Although some of the extra coupling captured by WD/L may not be very accurate, it is better than ignoring them. Figure 4 plots the eigenvalues for the full matrix,

the equivalent inductance matrices for WD/Ł, and WD/L. The bandwidth of L^{-1} and window size used in the wire duplication methods are 5 and 8, respectively. We can see that the eigenvalues of the equivalent inductance matrix for WD/Ł diverges earlier from that of the full matrix. In contrast, the equivalent matrix for WD/L matches better. Simulation results validate this conclusion (see Section 6 for details).

In the wire duplication method, using original inductance is more accurate than using double-inverse matrix only when the window size *B* is larger than its minimum value 2b + 1. If B = 2b + 1, using *L* does not capture more couplings. Indeed, it may not capture the correct L^{-1} values due to the small window size. Since using minimum window size (B = 2b + 1) is not efficient, a large window size is always preferred. Thus, we should always use the original inductance matrix *L* instead of *L*.

Recall that the simple truncation method also uses a portion of the original inductance matrix to model the interconnects. But the two methods are radically different. The key difference is that in the WD/L method, dummy wires are added. Thus, stability is guaranteed and inductance effect is captured accurately. It exploits the sparsity of L^{-1} , whereas simple truncation tries to exploit the sparsity of L matrix, it is not accurate and may not be stable. In the previous example, the simple truncation method that captures the same inductance terms as the wire duplication method has negative eigenvalues. It is not as accurate as the wire duplication model even when it is stable.

6 Experiment Results

We demonstrate the wire duplication technique on a bus with 128 signals. Shields are inserted after every four signals. The wire length is 1mm, the cross-section is $1 \times 1\mu$ m, and the separation between wires is 1μ m. The wires are divided into 5 segments along the length. The driver resistance is 30Ω and the load capacitance is 40 fF.

Different modelings are studied: the full matrix method, the Ł method, WD/Ł, WD/L, simple truncation, shifttruncate [7, 5], and the double-inverse inductance model [1, 2]. Because Ksim [6] is not available to us, we cannot perform *K* method directly. Instead, we use the Ł method, which is mathematically equivalent to the *K* method. (Note that Ł method is different from the double-inverse inductance model). The bandwidth of L^{-1} is 5, and the window size for wire duplication is 8.

A 1V 20*ps* ramp input is applied to the first signal, and the rest are quiet. The waveforms for the first, second that third signals are shown in Figures 5, 6, and 7 respectively. In each figure, results from the Ł method and the wire duplication methods are shown on the left; and results from the shift-truncate and double-inverse methods are shown on the right. As a reference, full matrix modeling appears at both sides. The truncation only method is not stable and not shown.

The first conclusion we can draw from these figures is that WD/Ł is equivalent to the Ł method. These two methods match *exactly* on all figures. The second conclusion is that WD/L is more accurate than WD/Ł. Although both of them match very well with the full matrix method, the results of using original matrix are closer. The difference between these two methods for the aggressor response (Figure 5) is insignificant, but we can still see that WD/L (which overlaps with the full matrix) is more accurate than WD/Ł. The differences between them in the victim responses (Figure 6 and



Figure 5: Simulation results for signal 1.



Figure 6: Simulation results for signal 2.



Figure 7: Simulation results for signal 3.

	-	-
Method	Memory(MB)	Run Time(s)
full L matrix	812.3	4.74×10^{5}
Ł method	230.2	2.30×10^{4}
double-inverse (1% cutoff)	97.1	897
double-inverse (2.2% cutoff)	44.3	357
shift-truncate	38.5	283
wire duplication	15.3	58

Table 1: Run time and memory usage.

7) are more pronounced.

In the double-inverse method [1, 2], the same cutoff percentage is used for the L^{-1} matrix and the Ł matrix. Both the double-inverse method with 2.2% cutoff and the shift-truncate method have similar number of mutual inductances as the wire duplication methods, while the double-inverse method with 1% cutoff has about twice as many mutual inductances. We can see that they perform worse than WD/L.

Table 1 lists the memory and time usage for different methods. Both wire duplication methods use the same time and memory. We can see that although there are additional dummy wires in wire duplication method, it uses much less memory and runs much faster. Although the number of wires and couplings contribute directly to run time of simulation, the convergence rate, which depends on how well-conditioned the matrices are, is also very important. We can see that the wire duplication methods are faster and require less memory even than the shift-truncate method and the double-inverse method (2.2% cutoff), despite the fact that their circuit sizes are actually larger (due to the additional dummy wires). We believe that the reason is that the inductance matrices of the wire duplication model are well-conditioned and they are block-diagonal, contributing to faster simulation times.

7 Summary and Conclusion

In this paper, we propose a new interconnect modeling technique—wire duplication. With this technique, we can generate stable, sparse and yet accurate inductance models for on-chip interconnects. It uses a very small portion of the original inductance matrix L and discards the rest. The wire duplication model using the double-inverse inductance matrix is physically and mathematically equivalent to K method. However, we exploit the sparsity of L^{-1} by using inductances in the simulation. Although it is not as sparse as the K simulation, it avoids using the new circuit element K. Using original inductance instead of double-inverse inductance makes the wire duplication model even more accurate. Moreover, it avoids matrix inversions.

A common theme in the *K* method, double-inverse inductance model, and the wire duplication model are that they all exploit the sparsity of L^{-1} by windowed inductance matrix. The *K* method and the double-inverse inductance model do it at the inductance extraction level, while the wire duplication model does it at the simulation level. The wire duplication model is more efficient than the double-inverse inductance model. We have no chance to compare its efficiency to *K* method. It has several other advantages:

• It is compatible with most of the existing simulators; we use inductance in the simulation.

- It is compatible with existing inductance extractors; we use the original inductance.
- It requires no matrix inverse.
- It is potentially more accurate.

8 Appendix

Theorem 1 Suppose A is a $n \times n$ non-singular matrix, and B is the inverse of A. M_m is a minor of A with order m formed by rows i_1, i_2, \dots, i_m and columns j_1, j_2, \dots, j_m . N_{n-m} is the matrix remained in matrix B after deleting rows i_1, i_2, \dots, i_m and columns j_1, j_2, \dots, j_m . Then,

$$|N_{n-m}| = (-1)^{\sum (i_k + j_k)} |M_m| / |A|.$$
(9)

Proof: We first consider the special case that $i_k = j_k = k$ $(1 \le k \le m)$. Thus, M_m is at the top left corner of A and N_{n-m} is at the bottom right corner of B. We can re-write A and B in block format:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

where A_{11} and B_{11} are $m \times m$ matrixes; A_{22} and B_{22} are $(n-m) \times (n-m)$ matrixes.

Suppose $|A_{11}| \neq 0$, we have [4]:

$$|B_{22}| = 1/|A_{22} - A_{21}A_{11}^{-1}A_{12}|$$
(10)

and,

$$|A| = |A_{11}||A_{22} - A_{21}A_{11}^{-1}A_{12}|$$
(11)

Therefore,

$$|B_{22}| = |A_{11}|/|A| \tag{12}$$

If $|A_{11}| = 0$, then $|B_{22}|$ should also be zero. Otherwise, if $|B_{22}| \neq 0$, since $A = B^{-1}$, following similar step, we can also argue that $|A_{11}| = |B_{22}|/|B| \neq 0$, a contradiction. So $|B_{22}| = |A_{11}|/|A|$ for all cases.

Now we extend the conclusion to general case. By making $(i_1 - 1) + (i_2 - 2) + \dots + (i_m - m) = (i_1 + i_2 + \dots + i_m) - m(m+1)/2$ successive interchanges of adjacent rows, and $(j_1 - 1) + (j_2 - 2) + \dots + (j_m - m) = (j_1 + j_2 + \dots + j_m) - m(m+1)/2$ successive interchanges of adjacent columns in *A*, we can shift M_k into the top left corner of *A*. We should also interchanges the corresponding columns and rows in *B*, in order to maintain the inverse relation between A and B. As a result, N_{n-k} is shifted to the bottom right corner of matrix *B*.

Let A' and B' be the resulting matrixes (here A' does not stand for the transpose of A, same for B').

$$|A'| = (-1)^{i_1 + i_2 + \dots + i_m + j_1 + j_2 + \dots + j_m - m(m+1)} |A|$$
(13)

The relative positions of the elements in M_k and N_{n-k} are preserved, so we have

$$A'_{11} = M_m, and B'_{22} = N_{n-m}$$
 (14)

Then,

$$|N_{n-m}| = |B'_{22}| = |A'_{11}|/|A'| = (-1)^{\sum (i_k + j_k)} M_m/|A|$$
(15)

Theorem 2 Suppose A is a $n \times n$ non-singular matrix, and B is the inverse of A. For the r^{th} row of A matrix, only the entries at columns i_1, i_2, \dots, i_m are non-zero, the rest are zeroes. Let N_m be the sub-matrix of B formed by rows i_1, i_2, \dots, i_m and columns j_1, j_2, \dots, j_m . If there exists $p, 1 \le p \le m$ such that $j_p = r$, then the p^{th} row of the inverse of N_m is equal to the r^{th} row of A (omitting the zero entries in the latter):

$$N_m^{-1}(p,:) = A(r,:).$$
(16)

Similar conclusion holds for the columns of A.

Proof: For $1 \le q \le m$,

$$N_m^{-1}(p,q) = (-1)^{p+q} |N_{m(q,p)}| / |N_m|$$
(17)

Here $|N_{m(q,p)}|$ means the minor of element (q, p) of N_m . From Theorem 1, we know that

$$|N_{m(q,p)}| = (-1)^{\sum_{k \neq q} i_k + \sum_{k \neq p} j_k} |M_{n-m+1}| / |A|,$$
(18)

and

$$|N_m| = (-1)^{\sum i_k + \sum j_k} |M_{n-m}| / |A|,$$
(19)

where M_{n-m+1} refers to the sub-matrix of A formed by deleting columns $i_1, \dots, i_{q-1}, i_{q+1}, \dots, i_m$ and rows $j_1, \dots, j_{p-1}, j_{p+1}, \dots, j_m; M_{n-m}$ refers to the sub-matrix of A formed by deleting columns i_1, \dots, i_m and rows j_1, \dots, j_m . The $(r - (p-1))^{th}$ row of M_{n-m+1} comes from the r^{th} row of A, but there is only one non-zero element in it. It is the $(r - (p-1), i_q - (q-1))$ entry (or the (r, i_q) entry of A). Using Laplace Expansion [4] on this row:

$$|M_{n-m+1}| = (-1)^{r-(p-1)+i_q-(q-1)}A(r,i_q)|M_{n-m}|.$$
(20)

We obtain,

$$N_m^{-1}(p,q) = A(r,i_q).$$
(21)

Theorem 3 Suppose A is a $n \times n$ non-singular band matrix with bandwidth 2b + 1, and B is the inverse of A. Let N be the sub-matrix of B formed by rows i - b to i + b and columns i - b to i + b. Then the central row and central column of the N^{-1} are identical to the *i*th row and *i*th column of matrix A, respectively:

$$A(i, i-b: i+b) = N^{-1}(b+1, :),$$

$$A(i-b: i+b, i) = N^{-1}(:, b+1).$$
(22)

If i - b < 1 or i + b > n, use 1 or n instead.

Proof: If we apply Theorem 2 to the i^{th} row and i^{th} column of A at the same time, we will reach this conclusion immediately.

Similar conclusion holds for the case that $i \le b$ or $i + b \ge n$. For example, if $i \le b$, the *N* matrix should be formed by rows 1 to i + b and columns 1 to i + b of *B*, and

$$A(i, 1: i+b) = N^{-1}(i, :),$$

$$A(1: i+b, i) = N^{-1}(:, i).$$
(23)

References

- M. Beattie and L. Pileggi. Efficient inductance extraction via windowing. In Proc. Design Automation and Test in Europe Conf., pages 430–436, 2001.
- [2] M. Beattie and L. Pileggi. Modeling magnetic coupling for on-chip interconnect. In Proc. Design Automation Conf, pages 335–340, 2001.
- [3] A. Devgan, H. Ji, and W. Dai. How to efficiently capture on-chip inductance effects: introducing a new circuit element K. In *Proc. Int. Conf. on Computer Aided Design*, pages 150–155, 2000.
- [4] Howard Eves. Elementary matrix theory. Allyn and Bacon, Boston, 1966.
- [5] Zhijiang He, Mustafa Celik, and Lawrence T. Pileggi. SPIE: Sparse partial inductance extraction. In Proc. Design Automation Conf, pages 137–140, 1997.
- [6] H. Ji, A. Devgan, and W. Dai. KSim: a stable and efficient RKC simulator for capturing on-chip inductance effect. In Proc. Asia South Pacific Design Automation Conf., pages 379–384, 2001.
- [7] B. Krauter and T. L. Pileggi. Generating sparse partial inductance matrices with guaranteed stability. In *Proc. Int. Conf. on Computer Aided Design*, pages 45–52, 1995.
- [8] A. E. Ruehli. Inductance calculation in a complex integrated circuit environment. *IBM Journal of Research and Development*, pages 470–481, September 1972.