

上海交通大学硕士学位论文

基于模型降阶的信号完整性分析

姓名： 邓立群
申请学位级别： 硕士
专业： 电路与系统
指导老师： 施国勇

此研究由国家自然科学基金支持 (No. 60572028)

2007.1.19

上海交通大学
学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：

日期 2007 年 1 月 19 日

上海交通大学

学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权上海交通大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密，在__年解密后适用本授权书。

本学位论文属于

不保密。

(请在以上方框内打“ ”)

学位论文作者签名：

指导教师签名：

日期：2007 年 1 月 19 日

日期：2007 年 1 月 19 日

基于模型降阶的信号完整性分析

摘 要

深亚微米芯片设计的会带来许多新的挑战。在这一背景下，本文首先阐述了模型降阶的基本原理，以及信号完整性分析的意义。接着介绍了几种经典的模型降阶算法，并分析比较了各种算法的优缺点。文中提出了一种基于可观测标准型的模型降阶算法，具有形式简单，易于检验稳定性的特点。在论文的实验部分，应用多种方法对几组互连线模型进行降阶，实践证明模型降阶算法能够得到比原电路小得多的近似模型，但精度并没有多少损失，利用该方法进行信号完整性分析可行有效。

关键词：模型降阶，矩匹配，信号完整性

SIGNAL INTEGRITY ANALYSIS BASED ON MODEL ORDER REDUCTION

ABSTRACT

Deep submicron IC design will bring many new challenges. Under such background, this paper first introduces the principles of model order reduction and the significance of signal integrity analysis. Several classical model order reduction algorithms are introduced, and their advantages and disadvantages are compared. In this paper, a new algorithm based on observability canonical form is proposed. It has a simple structure and is easy to check the stability of reduced model. In the third part of this paper, model order reductions are performed on several interconnect models using various kinds of algorithms. The results prove that model order reduction techniques can obtain a much smaller model than the original circuit without losing much accuracy. Signal integrity analysis based on this technique is feasible and effective.

KEY WORDS: Model Order Reduction, Moment Matching, Signal Integrity

目录

1 绪论	1
1.1 引言	1
1.2 模型降阶基本原理	2
1.2.1 电路的改进节点法 (MNA) 表示	2
1.2.2 模型降阶的最初思路	5
2 信号完整性分析的内容和意义	8
2.1 信号完整性问题概述	8
2.1.1 概念	8
2.1.2 低功耗设计带来的SI新挑战	10
2.1.3 深亚微米SI分析挑战	11
2.2 信号完整性对电路性能的影响	12
2.2.1 串扰导致的逻辑错误	12
2.2.2 串扰对时序分析的影响	12
2.2.3 电源网络的IR压降	13
2.2.4 衬底耦合噪声	13
2.3 信号完整性的分析方法	14
2.3.1 毛刺噪声分析	14
2.3.2 噪声延迟分析	16
3 模型降阶算法及其对比	19
3.1 渐进波形求值算法 (AWE)	19
3.1.1 传输函数	20
3.1.2 迭代求矩	22
3.1.3 直流分析	25
3.1.4 Elmore延迟	27
3.1.5 AWE的性能	29
3.2 Pade Via Lanczos (PVL)	30
3.2.1 AWE的缺点	31
3.2.2 PVL算法	33
3.2.3 PVL的优点	36
3.2.4 PVL的性能	38

3.3 无源互连线宏模型降阶算法 (PRIMA)	39
3.3.1 电路的无源性	39
3.3.2 PRIMA算法	41
3.3.3 保持矩	44
3.3.4 保持无源性	44
3.4 保持结构的互连线降阶模型 (SPRIM)	45
3.4.1 变换矩阵的选取	45
3.4.2 SPRIM的性能	46
3.5 其他的降阶方法	46
3.5.1 节点消去法	46
3.6 基于可观测标准型的模型降阶算法	48
3.6.1 系统状态方程	48
3.6.2 系统的可观测性	49
3.6.3 可观测标准型	50
3.6.4 MOROC算法	51
3.6.5 矩匹配	52
3.6.6 α 和 r_q 的计算	53
3.6.7 检验稳定性	55
4 实验与分析	56
4.1 RLCG互连线	57
4.1.1 幅度频率响应	58
4.1.2 串扰的时域波形	67
4.2 5 根耦合导线 1	69
4.3 5 根耦合导线 2	70
5 结束语	73
参考文献	74
附录	77
致谢	78
攻读学位期间发表的学术论文	79

图片目录

图 1 电路网络及其对应的拓扑图	2
图 2 电路的改进节点方程	5
图 3 模型降阶的目的	5
图 4 降阶系统示意图	7
图 5 几种常见的信号完整性问题	9
图 6 低功耗设计中的信号完整性问题	10
图 7 相邻导线串扰对逻辑电平的影响	12
图 8 相邻导线串扰对时延的影响	13
图 9 衬底耦合噪声	14
图 10 级联反相器和它们的电压传输曲线	15
图 11 级联反相器的噪声容限	16
图 12 入侵者相对受害者的时延	17
图 13 被害者的理想时延	17
图 14 延迟增量计算	18
图 15 替换电感电容为特定的独立源	23
图 16 更新独立源的值	24
图 17 一个RLC电路与它的拓扑图（粗线支路为支撑树）	25
图 18 反向深度遍历	26
图 19 一个N阶RC阶梯	27
图 20 N阶RC阶梯的拓扑图	28
图 21 两种基本的RC互连线电路模型	29
图 22 RICE和PSPICE的仿真波形对比	30
图 23 AWE不同阶数的降阶结果对比	31
图 24 混合信号电路仿真结果，左图：幅度，右图：相位	38
图 25 线性元 件等效为一个降阶的宏模型	39
图 26 多口网络的表示式	39
图 27 一个由稳定模型组成的不稳定系统	40
图 28 稳定但非无源的宏模型	41
图 29 节点消去的过程	48
图 30 模型降阶实验流程	56

图 31	两根电容耦合的RLCG互连线	57
图 32	AWE对两根电容耦合的RLCG互连线的仿真结果	58
图 33	PVL对两根电容耦合的RLCG互连线的仿真结果	60
图 34	PRIMA对两根电容耦合的RLCG互连线的仿真结果	61
图 35	MOROC对两根电容耦合的RLCG互连线的仿真结果	62
图 36	MOROC降阶模型 (10 阶) A_q 的特征值分布	64
图 37	MOROC降阶模型 (13 阶) A_q 的特征值分布	64
图 38	MOROC降阶模型的无源性检验	65
图 38	端口 2 的幅度频率响应	66
图 39	串扰的时域波形, PVL 6 阶	67
图 40	串扰的时域波形, PVL 15 阶	68
图 41	5 根耦合互连线	69
图 42	各端口的串扰波形, PVL 20 阶	70
图 43	导线 3 上的瞬态波形 (有串扰和无串扰, 时延减少)	71
图 44	导线 3 上的瞬态波形 (有串扰和无串扰, 时延增加)	72

表格目录

表 1	RICE和PSPICE运行时间对比.....	30
表 2	图 23 例子中 M_q 的条件数随阶数的变化.....	32
表 3	AWE算法中M矩阵条件数随阶数 q 的变化.....	59
表 4	Hurwitz判据中的 q_i	62

符号说明

$H(s)$	原电路的传输函数
$H_q(s)$ 或 $\tilde{H}(s)$	降阶模型的传输函数
\mathbf{T}_q	Lanczos 过程的三对角阵
\mathbf{H}_q	Arnoldi 过程的 Hessenberg 阵
$K_q(A, r)$	矩阵 A 和向量 r 构成的 q 阶 Krylov 子空间
$[x]_k$	x 在第 k 次迭代中的值

1 绪论

1.1 引言

随着超大规模集成电路 (VLSI) 技术的不断发展, 芯片的工作频率越来越高, 而其特征尺寸却越来越小。在一块芯片上集成的器件数量以摩尔定律增加的同时, 芯片本身的大小却没有多大变化。这不仅增加了设计上的复杂度, 更对验证和仿真提出了巨大的挑战。

首先, 我们知道, 当工艺尺寸逐渐步入深亚微米后, 寄生参数效应将变得更加明显, 不仅是导线间的耦合电容, 由于工作频率的提高, 寄生电感, 甚至导线间的互感也体现了出来, 对芯片上的信号完整性提出了更加苛刻的要求。不考虑有源器件, 仅从互连线的角度来看, 传统的 RC 模型已经不能胜任了, 取而代之的是 RLC、RLCM 等更加精确的模型。除了最常见的耦合串扰问题, 影响信号完整性的还有 IR Drop, Ringing 等, 这都要求在芯片验证时提供更加复杂的模型。

其次, 除了对仿真模型的复杂性提出要求外, 还对仿真的规模提出了更高的要求。这不仅因为芯片集成的器件增多, 而且由于集成度增加后, 布线的长度会比从前更长, 从而需要用分布式互连线模型才能够描述。对于非常高的频率, 甚至需要用到传输线模型。由于寄生参数也是不得不考虑的因素, 这也增加了仿真的规模。这使得直接利用传统的仿真工具如 Spice, 无论是仿真时间还是内存开销都已经不能满足需要了。

为了解决以上这两个问题, 必须有一种方法, 能够既保证仿真的精度, 又能够把仿真的规模减小到能够接受的程度, 这就是近十几年来研究的一个热点: 模型降阶技术。模型降阶技术实际上在很早以前已经被应用在控制领域了, 为的是解决较大规模状态空间的求解问题。这一方法被引入到电路领域是自 1990 年 L. T. Pillage 首次提出 AWE[1]以后, 由于仿真的对象变成了大规模集成电路的互连线, 通常规模比控制领域更大, 而且体现出了新的电路所特有的问题, 因此在此之后, 又有来自包括电路、自动化、应用数学等领域的大量学者提出了多种算法, 逐渐解决了稳

定性[3]、无源性[7]、保持电路结构[8]等等难题，还发展出了模型降阶的几大不同的方向。其中的一些方法已经被用在了工业用 EDA 工具中，推动了微电子技术向更小的特征尺寸前进的步伐。

1.2 模型降阶基本原理

1.2.1 电路的改进节点法 (MNA) 表示

一般来说，不管电路的结构有多么复杂，我们都可以把它抽象成由线段和点组成的几何图形。在这种图中，主要是考虑点和线之间的连接关系，而点的位置、线的长度以及几何弯曲程度则不予考虑[26]。图 1 就是一个电网络 and 它对应的拓扑图。

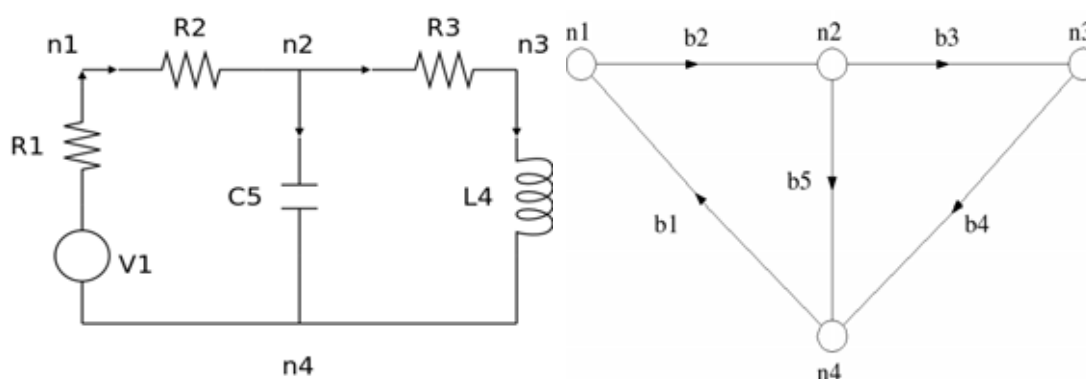


图 1 电网络及其对应的拓扑图

拓扑图有以下几个基本定义：

节点 (node)：拓扑图中各线段的端点。

支路 (branch)：连接两个节点之间的线段。

回路 (loop)：构成闭合路径的一组支路。

有向图 (directed graph)：支路具有方向的拓扑图。

虽然一个电网络可以用有向图来表示，但是把有向图输入计算机中去很不方便，需要有一种能使得计算机存储有向图的方法，这就用到了关联矩阵 (incidence

matrix) 。

设一个有向图有 $n+1$ 个节点, m 条支路, 则 E 是一个 $(n+1) \times m$ 阶矩阵, 它的行对应于节点, 列对应于支路, 其中元素 a_{ij} 的定义为:

$$\begin{aligned} a_{ij} &= 1, \text{ 如果支路 } j \text{ 与节点 } i \text{ 关联, 且支路 } j \text{ 的方向离开节点 } i \\ a_{ij} &= -1, \text{ 如果支路 } j \text{ 与节点 } i \text{ 关联, 且支路 } j \text{ 的方向指向节点 } i \\ a_{ij} &= 0, \text{ 如果支路 } j \text{ 与节点 } i \text{ 无关联} \end{aligned}$$

$$E = [a_{ij}] = \begin{matrix} & b_1 & b_2 & b_3 & b_4 & b_5 \\ \begin{matrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{matrix} & \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & -1 & -1 \end{bmatrix} \end{matrix} \quad (1)$$

由于矩阵中的任何一行都可以由其他行线性表出, 所以去掉一行, 把对应的节点作为参考节点, 仍然可以把这一行推算出来。简单起见, 通常就把去行后的矩阵称为关联矩阵, 而一般可以选取 0 点 (GND) 作为参考节点。

对电路网络的求解用的是 Kirchhoff 定律, 而 Kirchhoff 方程只取决于电路的连接关系, 因而可以用关联矩阵来表达:

$$\begin{aligned} \text{Kirchhoff 电流定律 (KCL):} & \quad E_{ib} = 0 \\ \text{Kirchhoff 电压定律 (KVL):} & \quad E^T v_n = v_b \end{aligned}$$

这里向量 i_b 和 v_b 分别表示支路电流和支路电压。 v_n 则表示节点电压。我们还可以在关联矩阵中把同类元件的支路写在一起, 这样关联矩阵 E 就可以写成:

$$E = [E_g \quad E_c \quad E_l \quad E_v] \quad (2)$$

除了 Kirchhoff 定律, 电路还由支路关系约束, 对于 RLC 电路, 这里的支路约束关系分别是:

$$R: i_g = Gv_g$$

$$C: i_c = C \frac{dv_c}{dt}$$

$$L: v_l = L \frac{di_l}{dt}$$

G, C 是对角矩阵，对角线上的元素分别是每一个支路上电导和电容元件的值。如果不考虑互感效应的话， L 也是一个对角矩阵。很显然，这三个矩阵都是对称的，且其中的元素都是非负的。利用这些关系，消去尽可能多的未知电流，我们就可以把电路描述成改进节点法的形式：

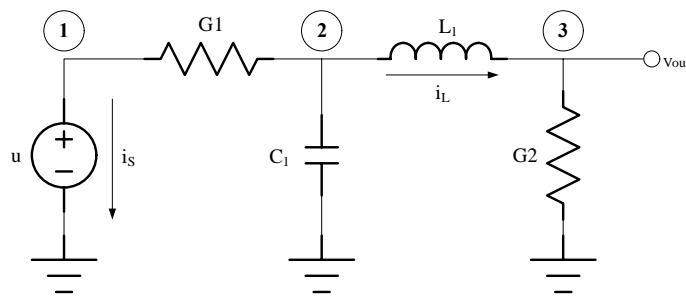
$$\begin{aligned} \mathbf{C}\dot{\mathbf{x}} &= -\mathbf{G}\mathbf{x} + \mathbf{B}u \\ y &= \mathbf{L}^T \mathbf{x} \end{aligned} \quad (3)$$

其中

$$\mathbf{C} \equiv \begin{bmatrix} E_c C E_c^T & 0 \\ 0 & H \end{bmatrix}, \quad \mathbf{G} \equiv \begin{bmatrix} E_g G E_g^T & E_l \\ -E_l^T & 0 \end{bmatrix}, \quad \mathbf{x} \equiv \begin{bmatrix} v \\ i \end{bmatrix} \quad (4)$$

C 表示储能元件电容和电感对电路方程的贡献， G 表示耗能元件电阻和电导对电路方程的贡献。 x 是状态变量，由所有节点的节点电压和一些未知电流变量组成。这些电流变量包括独立电压源的电流和电感支路的电流。

图 2 就是用改进节点方程表示一个电路的例子：



$$\begin{matrix}
 s \\
 \\
 \\
 \\
 \\
 \end{matrix}
 \underbrace{\begin{bmatrix}
 0 & 0 & 0 & 0 & 0 \\
 0 & C_1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 \\
 \hline
 0 & 0 & 0 & L_1 & 0 \\
 0 & 0 & 0 & 0 & 0
 \end{bmatrix}}_C
 \underbrace{\begin{bmatrix}
 v_1 \\
 v_2 \\
 v_3 \\
 i_L \\
 i_s
 \end{bmatrix}}_x
 = -
 \underbrace{\begin{bmatrix}
 G_1 & -G_1 & 0 & 0 & 1 \\
 -G_1 & G_1 & 0 & 1 & 0 \\
 0 & 0 & G_2 & -1 & 0 \\
 \hline
 0 & 1 & -1 & 0 & 0 \\
 1 & 0 & 0 & 0 & 0
 \end{bmatrix}}_G
 \underbrace{\begin{bmatrix}
 v_1 \\
 v_2 \\
 v_3 \\
 i_L \\
 i_s
 \end{bmatrix}}_x
 +
 \underbrace{\begin{bmatrix}
 0 \\
 0 \\
 0 \\
 0 \\
 1
 \end{bmatrix}}_b
 u$$

$$v_{out} = v_3 = [0 \quad 0 \quad 1 \quad 0 \quad 0] \mathbf{x}$$

图 2 电路的改进节点方程

1.2.2 模型降阶的最初思路

从改进节点方程我们可以看到，系数矩阵 C 和 G 的阶数等于电路中节点数加上电感支路数再加上端口数。这就说明当电路规模变得很大的时候，这个方程的系数矩阵将有很高的阶数，这对于求解带来很大的困难。

从另一方面来看，对于电路仿真来说，我们往往只关心输入和输出的关系是什么，即传输函数。而对于电路内部的状态变量 - - 节点的具体电压电流并不关心。因此如果能够找到另一个系统，它的状态变量比原来少很多，从而可以加快仿真的速度，但是它又能够很好得近似原系统的输入输出关系，那将带来很大的便利。

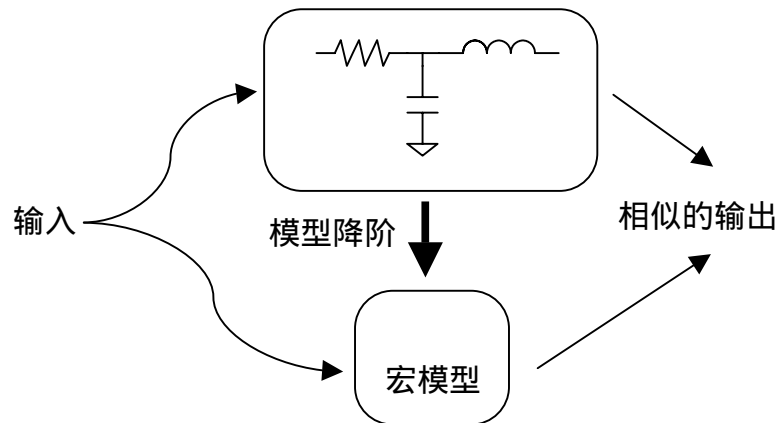


图 3 模型降阶的目的

这就是模型降阶的初衷，把原来的电路看成一个黑盒，这个黑盒的规模很大，希望找到另一个小规模模型，它的传输函数在一定程度上近似原电路的传输函数，那么就在这个模型称为降阶模型。这个降阶模型可以是一个等价电路，也可以

仅仅是一个对于仿真器来说等价的数学模型（图 3）。

出现最早的是一种基于矩匹配（Moment Matching）的降阶模型方法。它又分为两类：一类是利用 Pade 近似来进行瞬态匹配降阶[1][3][5][6][7][8]，包括著名的 AWE、PVL、PRIMA 等，另一类是基于近似的方法。Haifang Liao 等利用指数衰减的多项式函数来进行瞬态匹配降阶[9]。Xiaodong Yang 等利用 Hurwitz 多项式，提出了一个新的可实现的降阶模型技术。Bernard Sheehan 提出了一个基于时间常数平衡的可实现的降阶模型方法[10]。

我们已经知道，采用改进节点法 MNA 表示的一个集总、线性电路通常可以用一阶微分方程表示：

$$\begin{aligned} Cx' &= -Gx + Bu \\ y &= L^T x \end{aligned} \quad (5)$$

这个方程也可以改写成以下的形式：

$$\begin{aligned} x' &= Ax + Bu \\ y &= Cx \end{aligned} \quad (6)$$

其中矩阵 A, B, C 分别叫做系统矩阵，输入系数矩阵和输出系数矩阵。

如果系统矩阵 A 是稳定的，也就是说 A 的特征值全部在左开半平面内，我们可以定义 Controllability Grammian 为：

$$P = \int e^{At} B B^T e^{A^T t} dt \quad (7)$$

Observability Grammian 为：

$$Q = \int e^{A^T t} C^T C e^{At} dt \quad (8)$$

可以知道 P, Q 是对称的，并且满足如下的 Lyapunov 方程组：

$$\begin{aligned} AP + PA^T + BB^T &= 0 \\ A^T Q + QA + C^T C &= 0 \end{aligned} \quad (9)$$

下面我们来定义降阶模型的概念。所谓降阶模型就是为原来的线性系统（电网也可以看成是一个线性系统）的输入输出行为产生一个比较好的近似，至少在一定的激励范围内，例如在一定的频率范围内。衡量这种近似的对象就是系统的传输函数。

以上 MNA 方程组所描述系统的传输函数为：

$$\begin{aligned} H(s) &= C(sI - A)^{-1} B \\ Y(s) &= H(s)U(s) \end{aligned} \quad (10)$$

所以降阶模型就是要去寻找一个新的系统：

$$\begin{aligned} x_r' &= A_r x_r + B_r u \\ y_r &= C_r x_r \end{aligned} \quad (11)$$

以至于这个新的系统的状态变量远远小于原来系统的状态变量，而新系统的传输函数：

$$\begin{aligned} G_r(s) &= C_r(sI - A_r)^{-1} B_r \\ Y_r(s) &= G_r(s)U_r(s) \end{aligned} \quad (12)$$

在可接受的误差范围内与原系统很接近。

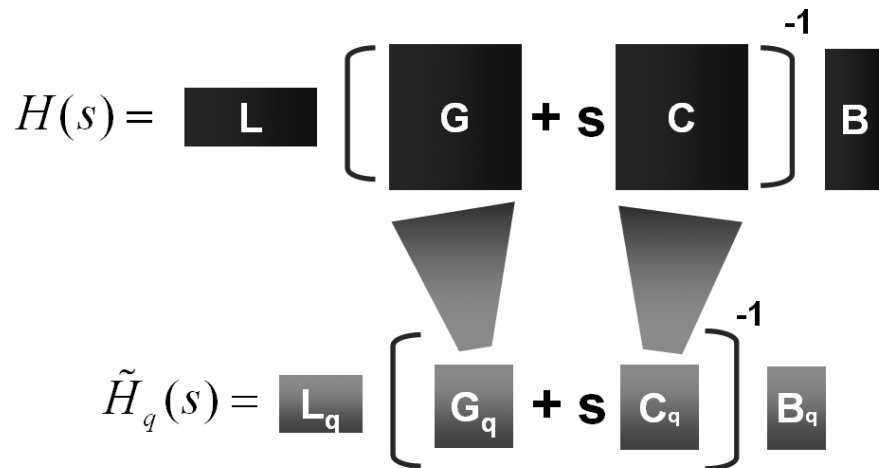


图 4 降阶系统示意图

2 信号完整性分析的内容和意义

超大规模集成电路技术已经遵循 Moore 定律发展了三十年，趋势是越来越复杂、特征尺寸越来越小，并且这个趋势在未来的十年也不会停止。然而随着技术发展的不仅仅是性能的提高，过去没有引起人们注意的各种物理现象在尺寸减小、频率提高后逐渐体现了出来，成为了阻碍技术发展的主要问题。这些物理现象包括：

- 更小的几何尺寸使得导线和过孔的电阻增大、相同电压下的电场强度增大。
- 更多的金属层使得线间耦合电容与对地耦合电容的比例增大。
- 更低的电源电压使得对于给定的功率需要更多的电流。
- 更低的器件阈值使得噪声容限变小。

因此，信号完整性因此被作为一个重要议题被提出来，我们要解决的就是如何克服电路寄生参数对信号完整性破坏的挑战，而使得性能和功能继续提高。

2.1 信号完整性问题概述

2.1.1 概念

信号完整性(SI)是指信号在电路中以正确的时序和电压作出响应的能力。如果电路中信号能够以要求的时序、持续时间和电压幅度到达 IC，则该电路具有较好的信号完整性。反之，当信号不能正常响应时，就出现了信号完整性问题。从广义上讲，信号完整性问题主要表现为 5 个方面：延迟、反射、串扰、同步切换噪声(SSN)和电磁兼容性(EMI)。

延迟是指信号在导线上以有限的速度传输，信号从发送端发出到达接收端，其间存在一个传输延迟。信号的延迟会对系统的时序产生影响，在高速数字系统中，传输延迟主要取决于导线的长度和导线周围介质的介电常数。

另外，当导线(高速数字系统中称为传输线)的特征阻抗与负载阻抗不匹配时，

信号到达接收端后有一部分能量将沿着传输线反射回去，使信号波形发生畸变，甚至出现信号的过冲和下冲。信号如果在传输线上来回反射，就会产生振铃和环绕振荡。

由于任何两个器件或导线之间都存在互容(mutual capacitance)和互感，当一个器件或一根导线上的信号发生变化时，其变化会通过互容和互感影响其它器件或导线，即串扰。串扰的强度取决于器件及导线的几何尺寸和相互距离。

当众多数字信号同步进行切换时(如 CPU 的数据总线、地址总线等)，由于电源线和地线上存在阻抗，会产生同步切换噪声，在地线上还会出现地平面反弹噪声(简称地弹)。另外，还有电磁兼容性问题，其产生主要与导线和器件的布局布线方式有关。

信号完整性分析是数字系统设计过程中保证完全符合要求规范的一个环节，包括诸如最优化时序和噪声容限以及理解波形、串扰、电磁干扰等种种模拟效应。随着系统性能的提高，这些二、三级效应逐渐变为主导因素，能够控制好它们直接决定了项目的成功与否。

数字系统和模拟系统在过去有着很大的差异，数字设计不需要考虑诸多的模拟问题因为时钟都工作在几十兆的频率下。而现在的数字系统往往工作在几百兆，甚至吉赫兹的频率下，工艺节点也走向了纳米级，信号通路中传播的信号再也不能看成是理想的方波了，来自各方面的噪声混杂其中，因此信号从源端到目的端的传输过程中能否保持完整性就成了不得不考虑的问题。

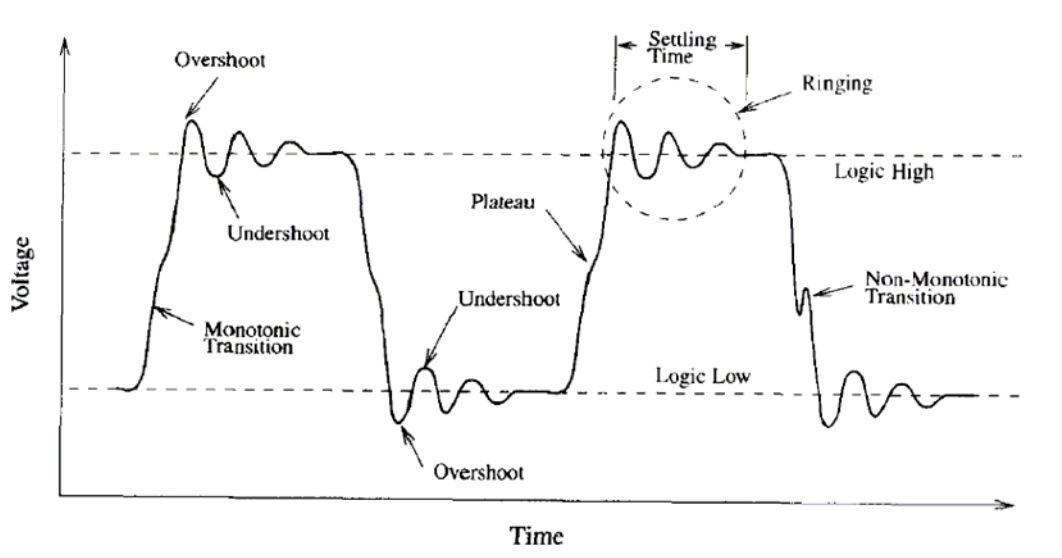


图 5 几种常见的信号完整性问题

2.1.2 低功耗设计带来的 SI 新挑战

业界在向 65 纳米节点变迁过程中，迫切需要低功耗设计，但随之而来的是新的 SI 挑战，如图 6 所示。究其原因，首先可归结于低功耗设计一般都采用多种电压，可能导致不同电压信号之间产生耦合。较之相同电压信号之间的耦合，从较高电压向较低电压的耦合要强得多。多电源电压(Vdd)设计还需要使用电平转换器。由于电路的复杂性，要验证电平转换器不受噪声问题的影响更加困难。

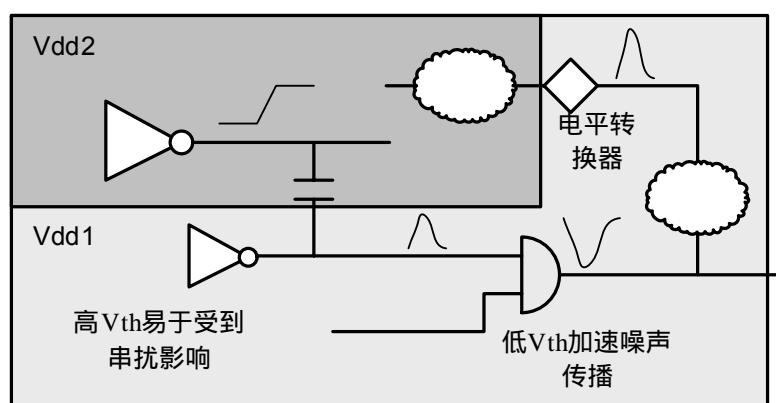


图 6 低功耗设计中的信号完整性问题

此外，低功耗设计使用了多阈值电压值(V_t)门电路。高 V_t 的器件往往具有更高的保持阻抗，从而更易受串扰的影响。另一方面， V_t 较低器件作为干扰传递者情形更糟糕，因为它们的转换速度较快。而且，它们在输入端还往往对波形影响更敏感。总体而言低功耗设计中，需要谨慎处理时序、功率和噪声的权衡取舍问题。

低功耗设计的另一个要素，是为了节省电力消耗，需要对芯片不同部分进行导通和关断。导通和关断过程在电源轨中产生瞬态效应，这可能对电路中仍在工作的其它部分造成影响。

动态电压和频率调整(DVFS)是一种旨在降低功耗、延长电池寿命的技术。利用这种技术，设计人员能够在芯片运行的同时动态地为不同模块设置不同的电压或不同的频率。当采用 DVFS 技术时，设计人员需要针对每个模块所允许的电压和频率的所有可能组合，对芯片性能进行验证，这大大增加了电气分析的复杂性和时间要求。而使用传统分析解决方案的另一项考虑因素，是每一个电压点都需要不同的库。

2.1.3 深亚微米 SI 分析挑战

在向深亚微米工艺变迁过程中，除了低功率设计方面的考虑，其它的一些 SI 问题也开始涌现。 V_t 和 V_{dd} 间的差距随电压调整在不断减小。由于 $(V_{dd}-V_t)^2$ 对性能的影响重大，故电压变化对门电路的延迟和抗扰性有着显著的影响。

电压变化也以非线性的方式影响延迟和波形。对门电路性能造成显著影响的还有温度。在高密度芯片中，芯片上的温度变化可达 50 摄氏度。这种变化对门电路的性能和信号完整性产生负面影响。

目前，设计人员是采用芯片波动(OCV)因数来计算 V_{dd} 变化和芯片上的温度变化。然而，在 65 纳米及更小节点时，由于这种变量加大，OCV 因数及其相关保护间隔(guard banding)极可能非常大，这时，需要一种能够提供电压和温度处理特定实例的新方案。

尽管一直以来大量注意力都放在工艺变异(process variations)及其对纯时序的影响上面，但实际上，工艺变异对芯片上所有信号的完整性都有影响，不同的 V_{th} 导致的噪声传播差别很大，因此信号的完整性必须在这些工艺波动的环境中进行分析。

信号完整性分析中的另一个重要的问题是如何使分析结果不要过分悲观。一般而言，分析工具都确立了一种假设，即设计的全部路径中皆存在最坏情况。它们分析设计路径中每一个网格可能被相邻网格最严重干扰的情况。当然，在大多数电路的工作情形中，最坏的情况实际上不会发生。较高级别的算法考虑和门级逻辑约束可以防止许多干扰源聚集在一起相互交换的情形发生。

减轻设计人员在 SI 分析中的悲观情绪的方法之一，是在任何可能的地方都要考虑到逻辑关系，这样一来，那些明显不能聚集在一起交换的信号就不被视为同时干扰者。另一个方法是利用随机分析技术来模拟较为实际的环境而非绝对的最坏情况。尽管这种方法本身存在风险，对建立(setup)分析仍不失为相当安全。

还有一种减轻 SI 分析悲观情绪的方法，即采用更实际的模型来计算串扰引起的延迟变化。考虑到某一特定噪声对路径的总体影响而不单单是对出现串扰的网格的影响，这一点十分重要。这种基于路径的对准(path-based alignment)方法通过产生最坏情况下的路径(而非网格)延迟变化，可显著降低悲观情绪。

在分析 IR 压降的影响时，分析工具考虑的也是最坏情况。当执行噪声分析或时序分析时，采用动态 IR 压降分析，并考虑到 IR 压降事件和噪声事件之间的暂时关系，可有助于利用唯一的实际 IR 压降值来减低悲观情绪。

2.2 信号完整性对电路性能的影响

信号完整性对性能的影响表现在以下几个方面：

2.2.1 串扰导致的逻辑错误

相邻导线（进攻者）的信号变化会通过耦合电容导致受害者导线的电压发生本不该有的变化。如果这种变化大到超过了逻辑阈值，就会带来错误的逻辑翻转。这种串扰的症状就是芯片在某种逻辑操作的时候会重复性地失败。

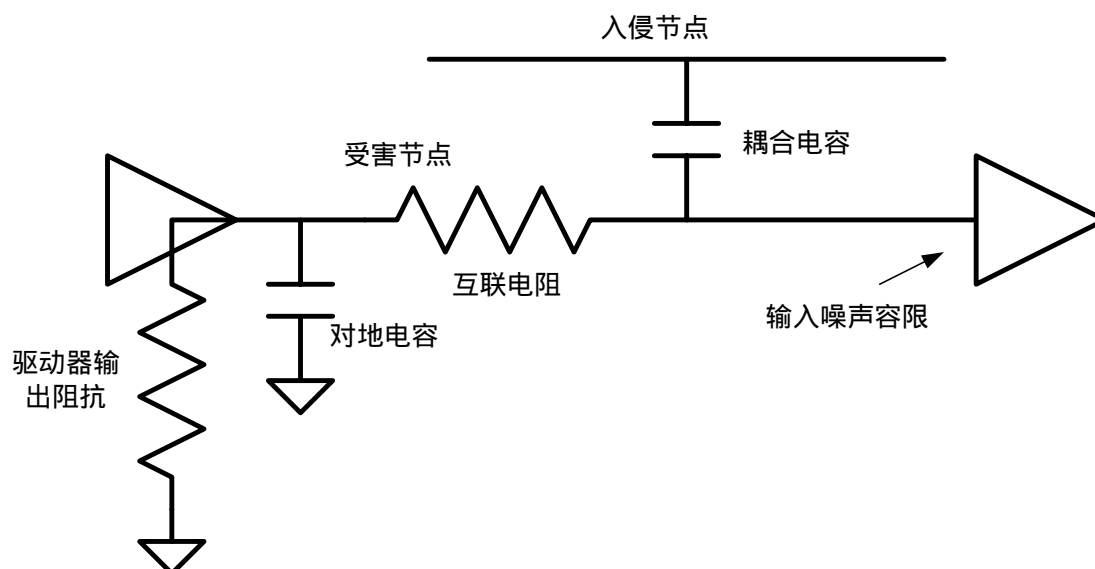


图 7 相邻导线串扰对逻辑电平的影响

2.2.2 串扰对时序分析的影响

通常我们采用静态时序分析（STA）来估计电路中可能存在的时序隐患，这种方法是基于 RC 模型对导线的时延进行估算。但由于是静态分析，所以并没有考虑到如果相邻导线如果处在不同的状态会对时序分析的结果带来什么影响。而实际上由于耦合电容的关系，时序分析取决于相邻导线的状态，得到的时延可能会完全不

同。

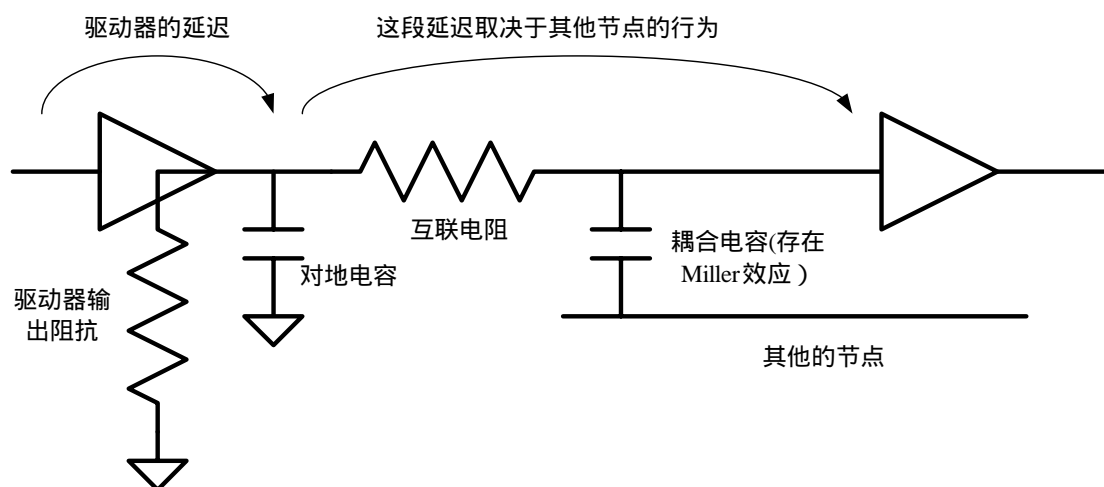


图 8 相邻导线串扰对时延的影响

2.2.3 电源网络的 IR 压降

电源电压的降低、器件噪声容限的减小使得 IR 压降问题日益突出。一般来说设计电源网络在布局布线之前完成，所以此时还并不知道最后所有的器件是如何连接到电源网络上的，只能是对需要的电流进行一种估算。如果在布线以后发现电源线不够宽再去修改将会是很麻烦的一件事。而且系统工作在不同的状态时，电流的分配会非常不同，如何找到最差情形的状态矢量也是一个很大的问题。

2.2.4 衬底耦合噪声

在 CMOS 工艺下，器件都是工作在同一个衬底上。当一个器件高速翻转的时候，电流就会通过衬底影响到邻近的其他器件。此外，由于每个 MOS 器件的衬底一般都是连到 VDD 或 VSS，所以电源的波动电流也会影响器件的功能。

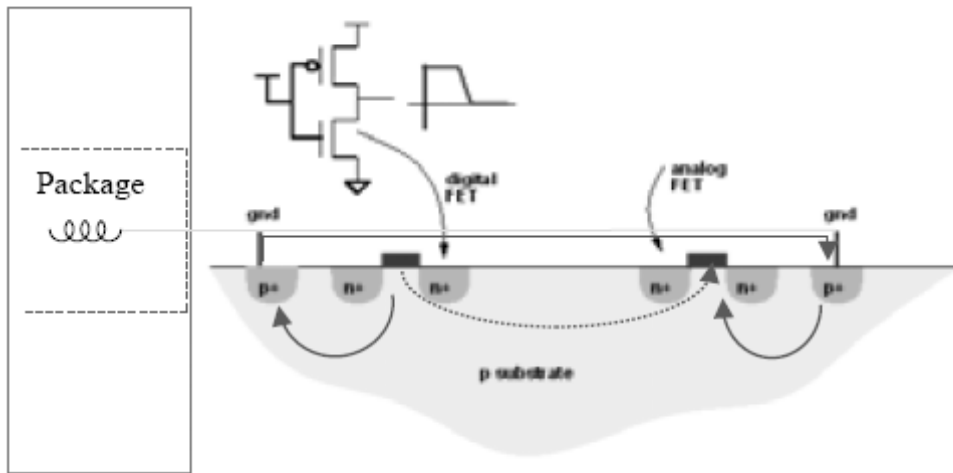


图 9 衬底耦合噪声

2.3 信号完整性的分析方法

2.3.1 毛刺噪声分析

毛刺噪声分析是检测在有噪声环境下（如串扰导致的毛刺），信号在接收端是否可能会导致逻辑功能错误的一种分析。它的一个假设就是如果电路中的每一个逻辑门都能抑制噪声的话，噪声就不会产生逻辑错误。

为了保证数字电路能够正常工作，必须保证在噪声存在的环境中，锁存器结构所保持的逻辑状态不会错误的改变。正常状态下，锁存器是处于稳定状态的，要引起它的翻转，就需要使锁存器的回路构成正反馈，从而处于不稳定的状态。噪声引起的电路的偏置发生变化，就有可能引起正反馈。

比如考虑如图 10 所示的由两个反相器构成的双稳态结构以及电压传输曲线。容易证明，不管反相器的具体电路组成如何，只要反相器在传输曲线过渡区的小信号增益大于 1，只有 A 点和 B 点是稳定的，C 点是一个亚稳态点。只要在 C 点有一个小的扰动，由于反相器环路的放大作用，电路马上会回到 A 点或者 B 点。而在 A 点和 B 点，环路增益远小于 1，所以即使有一个比较大的扰动也不能改变原

来的状态。

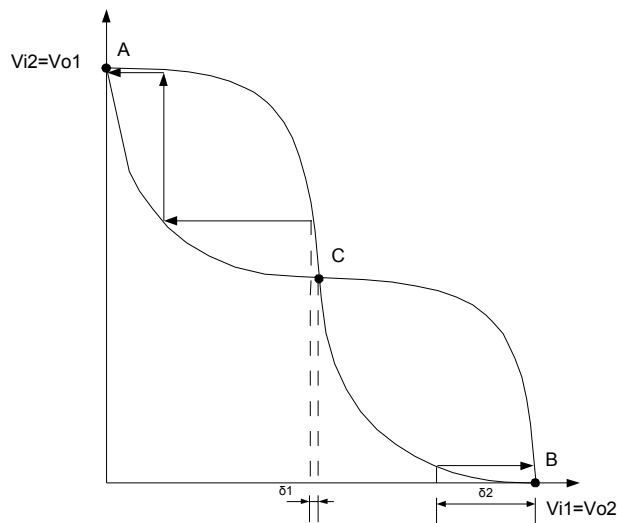
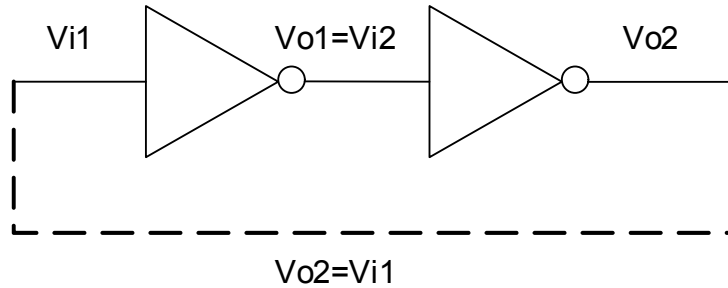


图 10 级联反相器和它们的电压传输曲线

如果在反向器输入端的噪声足够大，结果改变了反相器的偏置，使之处在 C 点的附近，那么这个双稳态结构就发生了翻转。所以保证不会发生翻转的充分条件就是：

$$\left| \frac{\partial V_{o1}}{\partial V_{i1}} \right| < 1, \quad \left| \frac{\partial V_{o2}}{\partial V_{i2}} \right| < 1 \quad (13)$$

在信号完整性分析的时候，可以检查噪声的幅度是否会达到使增益大于 1 的偏置来确保电路是可靠的。但是这样的检查依然过于保守，报出的很多的 Violation 并不会在实际流片后遇到。这时因为没有考虑逻辑门本身的低通滤波器的特性，取决于噪声波形的形状和逻辑门的性质，噪声的峰值完全可能超过以上条件所规定的阈值。这种动态的噪声阈值是依赖于噪声的时域特性的，所以为了尽量减少不必要的误报，就需要得到准确的时域仿真信息，这也正是需要应用到模型降阶的地方。

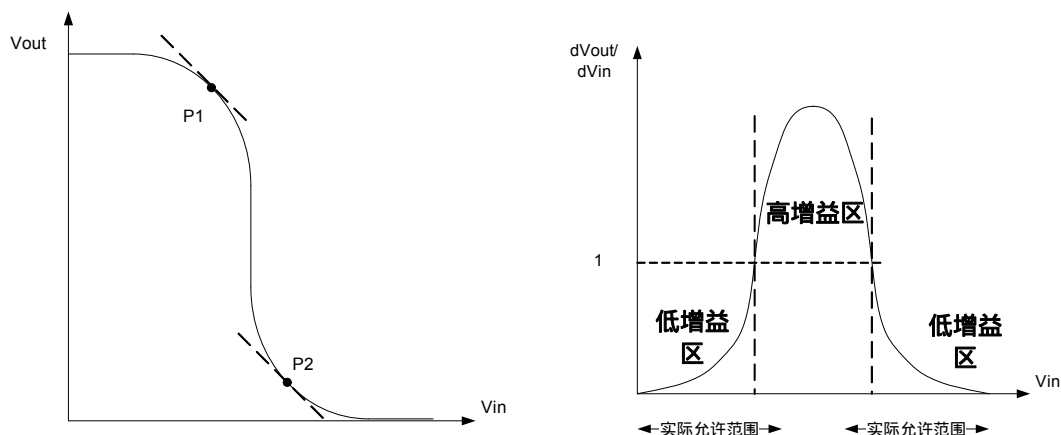


图 11 级联反相器的噪声容限

2.3.2 噪声延迟分析

在有串扰的情况下，信号通路上的延迟会发生变化。噪声延迟分析试图分析最坏情况下延迟的变化。这种延迟的变化不仅取决于导线间寄生参数的大小，还与噪声源的位置，噪声到达的时间等有关。比如考虑耦合互连线的例子，一般称受到串扰影响的导线为受害者（Victim），称施加串扰的导线为入侵者（Aggressor）。噪声延迟分析主要有以下几个步骤[30]：

入侵者时延计算

对于受害者来说，当导线上的电平发生变化时，并不知道入侵者产生的串扰什么时候会到达，取决于到达的时间点，有可能对延迟有影响，也可能没有影响。为了作最保守的估计，就必须考虑影响最大的一种情况，就是所有入侵者的毛刺噪声在受害者电平变化到 50% 的时候达到最大值。

由于不同的入侵者本身的时间常数、和受害者耦合的位置等都不同，所以要计算出毛刺到达被害者的时延。

入侵者时延定义为从入侵者输入端发生 50% 电平变化 (t_{a1}) 到被害者的噪声达到最大 (t_{va1}) 的延迟。在计算一个入侵者 (Agg1) 时延的时候，其他入侵者 (Agg2) 的输入端都接地，被害者的输入端接驱动器的输出阻抗。入侵者的耦合电容除了连接到被害者的以外都接地。被害者的终端接的是下一级接收器的输入端电容作为负载。

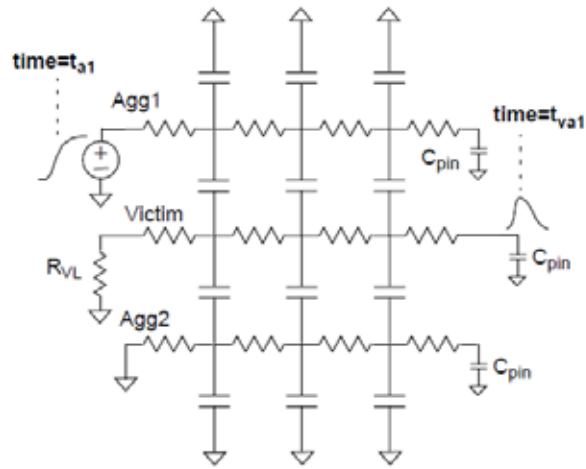


图 12 入侵者相对受害者的时延

受害者理想时延计算

受害者时延指的是在没有串扰的理想情况下，从导线的驱动端到接收器引脚的延迟，如图 13 所示。由于假设没有入侵者的影响，因此入侵者的输入端都接上驱动器的输出阻抗，模拟驱动器在稳态时上拉或者下拉网路的阻抗。对于受害者则把驱动器和导线一起仿真，驱动器的输入端加上适当斜率的阶跃输入，计算 50%-50%的延迟 ($t_{v2} - t_{v1}$)。

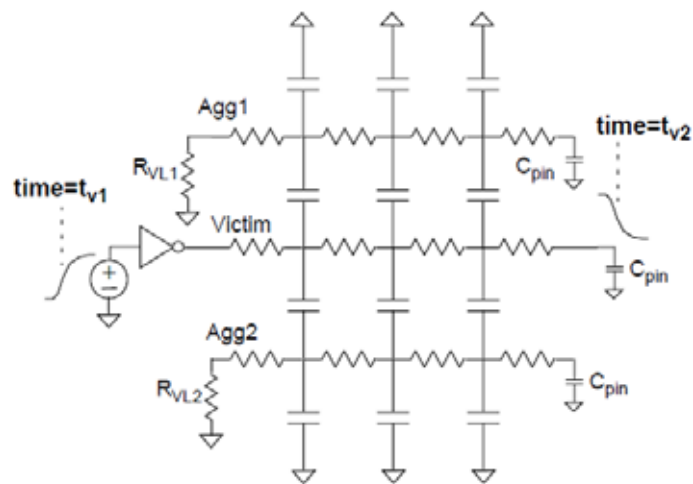


图 13 受害者的理想时延

延迟增量计算

结合前两步的计算结果，为了使入侵者的噪声同时到达，且到达的时候正好是受害者电平变化到 50% 的时刻，在入侵者的输入端加上一个分段线性的驱动信号，其偏移量等于受害者理想时延减去入侵者的时延，如图 14 所示。现在就可以求出延迟变化的最大值。这时对于延迟增大的情形，如果噪声和信号变化的方向是相同的话，延迟就会减小。计算延迟减小时，一般使噪声最大点和信号变化 5% 的时刻对齐。

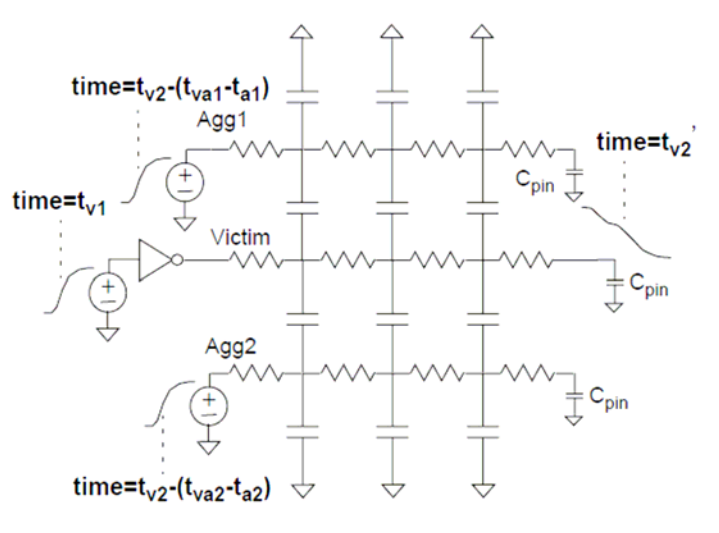


图 14 延迟增量计算

3 模型降阶算法及其对比

90 年代以来，模型降阶技术得到了充分的发展。互连电路可以采用 RLC 模型，并被描述为一个一阶线性微分代数方程的形式。对于该一阶系统，模拟降阶的主流技术是基于 Krylov 投影、矩匹配和合同变换的降阶方法。线性投影降阶的终极目标是同时保持降阶系统的结构、稳定性、无源性以及矩匹配的数目。这方面的先驱性工作是由 1990 年美国德州大学奥斯汀分校的 Lawrence T. Pillage 提出 AWE 方法[1]，但该方法由于显示计算矩而失去了数值稳定性，无法获得高阶模型。1995 年 AT&T Bell 实验室的 Roland W. Freund 提出的 PVL[3]、MPVL[4]类方法能够精确匹配矩，也能保证降阶电路的稳定性，但不能保证电路方程的无源性。1998 年 L. Pillage 等人提出的 PRIMA[7]可以达到 MPVL 的精度同时能够保持降阶电路的稳定性和无源性，但不能保持电路的结构，以至于很难作为实际电路综合，而且 PRIMA 只匹配了最优数量一半的矩。到 2004 年，R. Freund 提出了可以保持数值稳定性、无源性，以及保证降阶系统结构的 SPRIM 算法[8]，且精度高于 PRIMA，全面解决了基于一阶系统的 Krylov 投影降阶问题。

与此同时，模型降阶领域还衍生出了很多不同的方向。比如基于测量数据的降阶算法[19][20]，多点匹配法[21][22]，针对 RCS 电路的二阶 Arnoldi 算法[23]，节点消去法[10][24]。

3.1 渐进波形求值算法 (AWE)

Lawrence T. Pillage 是把模型降阶算法引入到电路仿真中的先驱，他在 1990 年提出的 Asymptotic Waveform Evaluation (AWE) 方法是模型降阶算法在电路仿真领域的首次应用。虽然在现在看来，当时的方法比较原始，而且精度不高，电路的很多其他特性也无法保证，但是这却是一个开创性的工作，使得当时对互连线时延的计算的速度大大提高，并开辟了 EDA 方向的又一个研究热点。下面简要介绍一下 AWE 的思想，这些概念将是我们研究模型降阶的基础。

3.1.1 传输函数

由电路的 MNA 方程，它的传输函数可以表示为：

$$H(s) = C(sI - A)^{-1}B \quad (14)$$

显然，这是 s 的一个函数。AWE 的基本过程是用电路传输函数 $H(s)$ 的极点和余式的一个子集去近似原来的 $H(s)$ ：

$$\tilde{H}(s) = \frac{k_1}{s-p_1} + \frac{k_2}{s-p_2} + \dots + \frac{k_q}{s-p_q} \quad (15)$$

其中 p_i 和 k_i 是分别是极点和余式。 q 是近似的阶数。这个式子很容易写成时域对应的响应：

$$\tilde{h}(t) = k_1 e^{p_1 t} + k_2 e^{p_2 t} + \dots + k_q e^{p_q t} \quad (16)$$

这就是我们时序分析所希望得到的。

既然 $H(s)$ 和 $\tilde{H}(s)$ 都是 s 的函数，就可以用在某一频率点的泰勒展开来衡量它们之间的近似程度：

$$H(s_0 + \sigma) = \tilde{H}(s_0 + \sigma) + O(\sigma^q) \quad (17)$$

这表示在 s_0 点，降阶前后的传输函数的差别只在于高次项 σ^q 。从而我们可以认为在某一频率附近，降阶后的模型有着足够的近似度。

由 Laplace 变换的定义不难得出原电路传输函数的泰勒展开形式为：

$$\begin{aligned} H(s) &= \int_0^{\infty} h(t)e^{-st} dt = \int_0^{\infty} h(t) \left[1 - st + \frac{1}{2}s^2t^2 - \dots \right] dt \\ &= \underbrace{\int_0^{\infty} h(t)dt}_{m_0} - s \underbrace{\int_0^{\infty} th(t)dt}_{m_1} + s^2 \underbrace{\frac{1}{2} \int_0^{\infty} t^2 h(t)dt}_{m_2} - \dots \\ &= m_0 + m_1 s + m_2 s^2 + \dots \end{aligned} \quad (18)$$

其中

$$m_i(t) = \frac{(-1)^i}{i!} \int_0^{\infty} t^i h(t) dt, \quad i = 1, 2, \dots \quad (19)$$

称为 $H(s)$ 的矩 (moment)。这是因为这种积分的形式与概率论和物理学中

“矩”的定义类似。

AWE 使用矩匹配 (moment-matching) 来确定降阶模型, 根据前面的定义, $\tilde{H}(s)$ 的前 q 个矩 m_0, m_1, \dots, m_{q-1} 必须保持和实际电路的矩相等。对于 $\tilde{H}(s)$, 它的矩可以用以上得到的矩的定义式得出, 把(*)式用到应用到 $\tilde{h}(t)$ 上我们就得到:

$$\tilde{m}_i = \frac{k_1}{p_1^{i+1}} + \frac{k_2}{p_2^{i+1}} + \dots + \frac{k_q}{p_q^{i+1}}, \quad i=1, 2, \dots \quad (20)$$

所以, 对于 q 阶的近似会有 q 个未知极点 ($p_1 \dots p_q$) 和 q 个未知余式 ($k_1 \dots k_q$), 一共是 $2q$ 个未知数。解 $2q$ 个未知数需要 $2q$ 个独立方程, 它们是:

$$\begin{aligned} -(k_1 + k_2 + \dots + k_q) &= m_0 \\ -\left(\frac{k_1}{p_1} + \frac{k_2}{p_2} + \dots + \frac{k_q}{p_q}\right) &= m_1 \\ -\left(\frac{k_1}{p_1^2} + \frac{k_2}{p_2^2} + \dots + \frac{k_q}{p_q^2}\right) &= m_2 \\ &\vdots \\ -\left(\frac{k_1}{p_1^{2q-1}} + \frac{k_2}{p_2^{2q-1}} + \dots + \frac{k_q}{p_q^{2q-1}}\right) &= m_{2q-1} \end{aligned} \quad (21)$$

注意到等式右边是实际电路对于某个特定激励响应的 $2q$ 个矩。响应的变量可以是电压或者电流, 而且只需要求解我们关心的响应。

必须注意到, 得到的这 $2q$ 个方程并不是简单的线性方程, 因此需要特殊的方法去求解。首先通过解下面这个方程组得到特征多项式的系数 a_0, a_1, \dots, a_{q-1} :

$$\begin{bmatrix} m_0 & m_1 & \dots & m_{q-1} \\ m_1 & m_2 & \dots & m_q \\ & \vdots & & \vdots \\ m_{q-1} & m_q & \dots & m_{2q-2} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{q-1} \end{bmatrix} = \begin{bmatrix} -m_q \\ -m_{q+1} \\ \vdots \\ -m_{2q-1} \end{bmatrix} \quad (22)$$

$$a_0 + a_1\tau + a_2\tau^2 + \cdots + a_{q-1}\tau^{q-1} + \tau^q = 0 \quad (23)$$

这个特征多项式的根的倒数就是未知的 q 个极点。然后把计算得到的极点的值代入方程组(*)，此时这已经是一个线性方程组了，就能够求得所有的余式 k_i 。当然，在求解的过程中仍然有很多特殊的情况，比如当有重极点的时候，方程组就会变得奇异，但这都可以通过一些技巧解决。

接下来的一个问题就是对于一个实际的电路，如何快速计算出这 $2q$ 个矩的值。而且这里不仅要考虑计算速度的问题，还要保证算法的稳定性和精度。

应用 AWE 算法计算互连线电流波形的一个例子是 Pillage 提出的 Rapid Interconnect Circuit Evaluation (RICE) [2]。RICE 的巧妙之处在于它没有直接通过矩阵运算去得到矩，而是利用电路自身的拓扑结构，通过反复迭代的直流分析，最后计算出电路的矩。能够应用这一技巧的主要原因是观察到 RC 互连线模型在拓扑结构上大致可以看成由电阻 R 组成的支撑树 (spanning-tree) 和由电容 C 组成的连枝 (branch) 构成。所以可以用对树结构遍历的方法快速求解矩，而对于电路中不符合这一要求的那些支路，也能通过其他的方法处理，但会牺牲掉一些性能。

3.1.2 迭代求矩

1. 把输入替换成与稳态值相等的直流激励源，所有电容替换成电流为 0 的电流源，所有电感替换成电压为 0 的电压源。
2. 对电路做直流分析，求得所有电压源的电流，作为其对应电感在本次迭代中的矩。同样也求得所有电流源的电压，作为其对应电容在本次迭代中的矩。
3. 把输入替换成 0 值激励源，并把电容电压源和电感电流源的值设为上一次迭代的矩分别与电容、电感自身的值相乘后的结果，然后返回执行第 2 步。
4. 以此类推，迭代 q 次后将得到电路的 q 阶矩。

这样的描述并不是很直观，能够这样做的推导如下：

还是利用图 2 中的电路，经过第 1 步中的替换，电路将变成如图 15 所示：

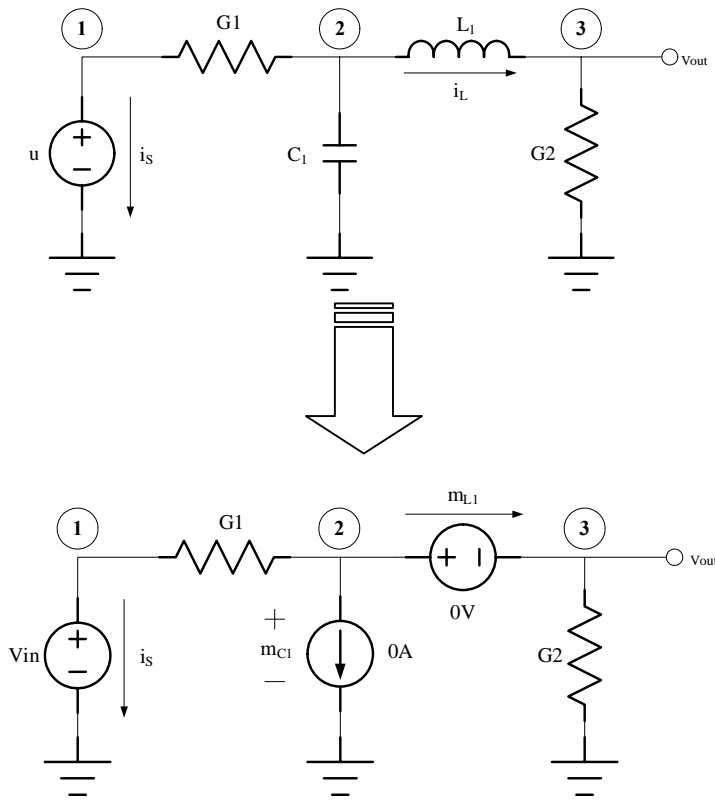


图 15 替换电感电容为特定的独立源

因为是做直流分析，且所有的电容电流源、电感电压源都是 0 值，所以该电路的 MNA 方程就退化为

$$0 = -\mathbf{G}\mathbf{x} + \mathbf{b}u \quad (24)$$

解得 $[\mathbf{x}]_0 = \mathbf{G}^{-1}\mathbf{b}u$ （以下的推导中将用 $[\]_i$ 表示第 i 次迭代）。电容和电感的 0 阶矩分别是对应电流源的电压和电压源的电流：

$$\begin{aligned} [m_{C1}]_0 &= v_2 = [0 \ 1 \ 0 \ 0 \ 0][\mathbf{x}]_0 \\ [m_{L1}]_0 &= i_L = [0 \ 0 \ 0 \ 1 \ 0][\mathbf{x}]_0 \end{aligned} \quad (25)$$

迭代的第 2 步，把电容、电感对应电源的值更新为上一次的矩与自身的乘积（图 16）：

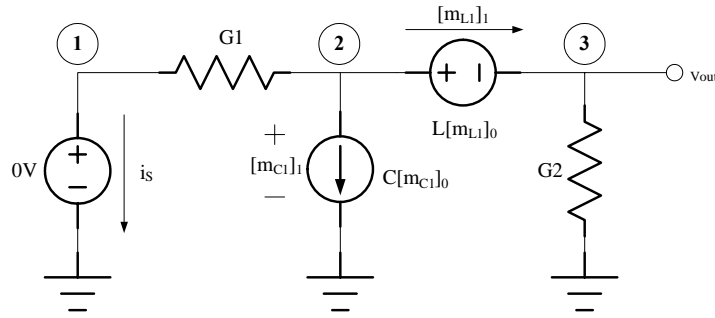


图 16 更新独立源的值

这个新电路的 MNA 方程变成：

$$\mathbf{C}[\mathbf{x}]_0 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & C_1[m_{C1}]_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & L_1[m_{L1}]_0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = - \underbrace{\begin{bmatrix} G_1 & -G_1 & 0 & 0 & 1 \\ -G_1 & G_1 & 0 & 1 & 0 \\ 0 & 0 & G_2 & -1 & 0 \\ \hline 0 & 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}}_{\mathbf{G}} \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ i_L \\ i_s \end{bmatrix}}_{[\mathbf{x}]_1} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = -\mathbf{G}[\mathbf{x}]_1$$

因此解得

$$[\mathbf{x}]_1 = -\mathbf{G}^{-1}\mathbf{C}[\mathbf{x}]_0$$

同样地，再把上述过程继续下去：

$$[\mathbf{x}]_2 = -\mathbf{G}^{-1}\mathbf{C}[\mathbf{x}]_1$$

$$[\mathbf{x}]_3 = -\mathbf{G}^{-1}\mathbf{C}[\mathbf{x}]_2$$

.....

经过 k 次迭代后就有：

$$\begin{aligned} [\mathbf{x}]_k &= -\mathbf{G}^{-1}\mathbf{C}[\mathbf{x}]_{k-1} \\ &= (-\mathbf{G}^{-1}\mathbf{C})^k \mathbf{G}^{-1}\mathbf{b}u \end{aligned} \quad (26)$$

因此原电路传输函数的 i 阶矩就应该是：

$$m_k = \frac{y}{u} = \frac{\mathbf{L}^T[\mathbf{x}]_k}{u} = \mathbf{L}^T (-\mathbf{G}^{-1}\mathbf{C})^k \mathbf{G}^{-1}\mathbf{b} \quad (27)$$

另一方面，由矩的定义——传输函数 $H(s)$ 的泰勒展开系数，有：

$$\begin{aligned}
H(s) &= \mathbf{L}^T (\mathbf{G} + s\mathbf{C})^{-1} \mathbf{b} \\
&= \mathbf{L}^T (\mathbf{I} + s\mathbf{G}^{-1}\mathbf{C})^{-1} \mathbf{G}^{-1} \mathbf{b} \\
&= \mathbf{L}^T [\mathbf{I} - s\mathbf{G}^{-1}\mathbf{C} + s^2(-\mathbf{G}^{-1}\mathbf{C})^2 + \dots] \mathbf{G}^{-1} \mathbf{b} \\
&= \sum_{k=0}^{\infty} [\mathbf{L}^T (-\mathbf{G}^{-1}\mathbf{C})^k \mathbf{G}^{-1} \mathbf{b}] s^k
\end{aligned} \tag{28}$$

所以通过迭代计算的到的矩正好就是我们想要的。事实上从上述推导过程也能够看出，每次迭代中直流分析的求解过程其实就是完成了矩阵求逆并相乘的计算。

3.1.3 直流分析

很显然，在求矩的每次迭代过程中，最主要的工作就是解一个直流电路。电路中只包含电阻、电压源和电流源。一般来说，直流分析有很多方法，关键是解一个线性方程。考虑互连线网络具有树状的结构，因此可以采用一种特殊的方法来求解，以提高 AWE 的效率。

比如对于图 17 (a) 中的一个 RLC 电路，如果把电阻、电感以及独立电压源所在的支路组成支撑树，电容所在的支路构成连枝，则它的拓扑图如图 17 (b) 所示。由于互连线的寄生参数大多数情况下，电阻和电感是沿着导线的，而电容多是对地电容和两根导线之间的耦合电容，所以上述假设去掉一些特殊的支路后几乎总能成立。

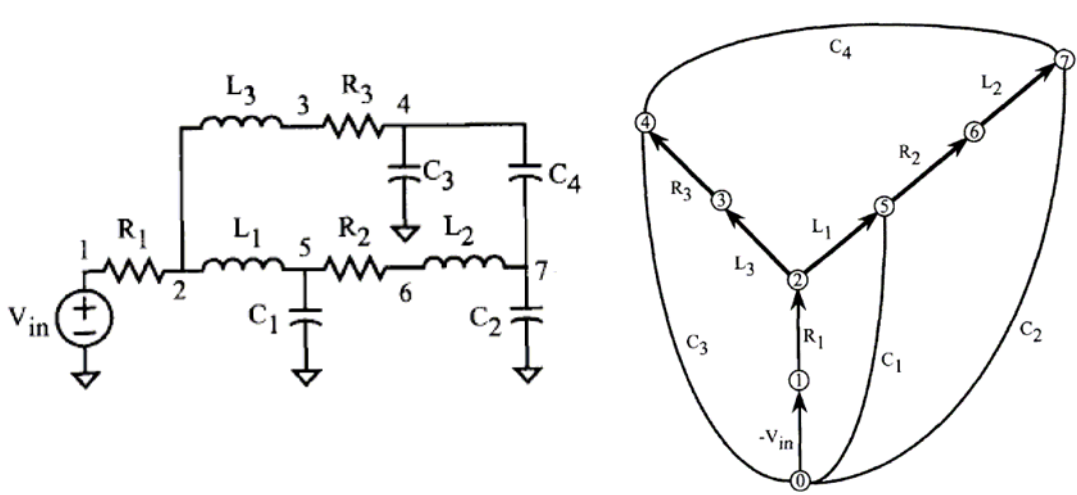


图 17 一个 RLC 电路与它的拓扑图（粗线支路为支撑树）

把 RLC 电路看成这样的树状结构的好处就是能够利用迭代的特点快速求解。在迭代过程中，电容被替换成了电流源，电感被替换成了电压源，所以对于所有的电容连枝，支路的电流都是已知的。而对于支撑树来说，电感支路的电压自然是已知的，而电阻支路如果能够求得电流的话那么电压也可以知道。

基于这种思想，只需要对支撑树进行两次深度遍历就可以完成一次直流分析：

第 1 步. 反向深度遍历

首先从一个叶节点开始，访问某一个节点的时候，把它所有从子节点来的树枝的电流与同它相连的连枝的电流求和，并由 KCL 定律可以马上知道求和结果就等于该节点向上树枝的电流（图 18）。

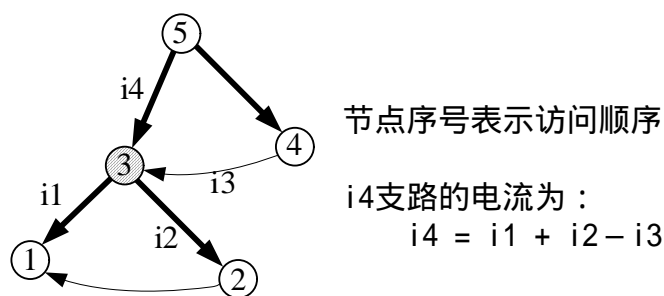


图 18 反向深度遍历

反向遍历保证了遍历的时候只有当从一个节点所有子节点来的支路的电流都已知时才计算该节点到父节点支路的电流。这一遍历过程一直进行到根节点为止，这样所有支路的电流都是已知的了。其中电感支路的电流就是我们要求的矩，而电阻支路的电流将在前向深度遍历中用来计算电阻两端的电压。

第 2 步 前向深度遍历

从根节点开始，计算每一个节点的子节点的电压。子节点的电压等于父节点的电压减去父节点到子节点的支路上的电压降。对于电压源支路和电感支路，因为在迭代之前的电感已经被替换成了电压源，所以它们的支路电压都是已知的。对于电阻支路，因为在反向深度遍历中已经求得了支路电流，所以只需要乘以电阻的阻值就能得到支路电压。节点访问的顺序为反向深度遍历的逆序，这保证了只有当一个节点的父节点电压已知时才计算该节点的电压。

遍历一直进行到访问过所有的节点为止，这样每一个节点的电压就是已知的了，把电容两端的电压相减就能求出电容支路上的电压，也即电容对应的电流源的

电压。

研究后可以发现，上述遍历过程只取决于电路本身的拓扑结构，是在电路存在的时候就已经确定的。虽然每次迭代中各电流源、电压源的值发生变化，但不改变计算的公式。因此，如果每次迭代都重新遍历就显得有些多余。一种更有效率的方法是可以仅对支撑树遍历一次，记录下每个支路电流、电压的计算公式，然后迭代的时候把具体的数值代入即可。

3.1.4 Elmore 延迟

如果把上述的降阶算法应用于 RC 电路，就能够发现互连线延迟计算中常用的 Elmore 延迟实际上就是一阶的 AWE 的结果。比如对于如图 19 所示的 N 阶 RC 阶梯：

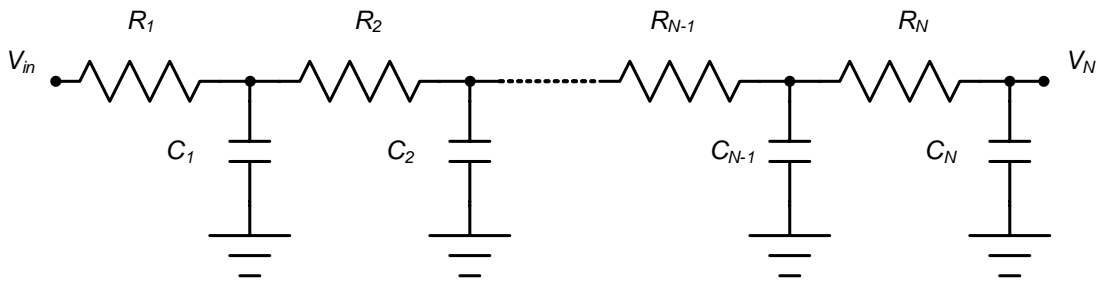


图 19 一个 N 阶 RC 阶梯

根据 Elmore 延迟， V_N 端的时间常数为 [27]：

$$\tau_N = \sum_{i=1}^N R_i \sum_{j=i}^N C_j = R_1(C_1 + C_2 + \dots + C_N) + R_2(C_2 + \dots + C_N) + \dots + R_N C_N \quad (29)$$

下面我们将用 1 阶的 AWE 计算 V_N 端的延迟。

由于只计算一个极点和余式，所以降阶模型的传输函数可以写成：

$$\tilde{H}(s) = \frac{k_1}{s - p_1} \quad (30)$$

用树状结构画出电路的拓扑图（图 20），

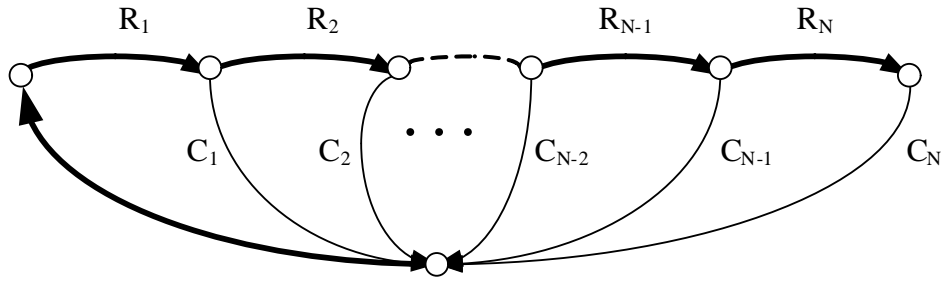


图 20 N 阶 RC 阶梯的拓扑图

首先把所有的电容替换成 0 A 的电流源，这相当于电容支路全部开路，容易计算得出此时各个节点上的电压都等于输入电压 V_{in} ，简单起见，我们设输入电压为 1V。那么所有电容的 0 阶矩就等于电容上的电压：1V。而 C_N 上的电压就是输出的矩：

$$m_0 = 1$$

接着进行下一次迭代，更新电容电流源的值为上一次的矩与电容值的乘积，因为上一次的矩都等于 1，所以电流源的电流就等于对应的电容值。

进行一次反向深度遍历，容易得出各个电阻支路上的电流分别是：

$$\begin{aligned} i_{RN} &= i_{CN} = C_N \\ i_{RN-1} &= i_{CN-1} + i_{RN} = C_{N-1} + C_N \\ &\dots\dots \\ i_{R1} &= i_{C1} + i_{R2} = C_1 + C_2 + \dots + C_N \end{aligned} \quad (31)$$

再进行一次前向深度遍历，得到：

$$\begin{aligned} [V_1]_1 &= 0 - V_{R1} = -i_{R1} R_1 = -R_1(C_1 + C_2 + \dots + C_N) \\ [V_2]_1 &= [V_1]_1 - V_{R2} = -R_1(C_1 + C_2 + \dots + C_N) - R_2(C_2 + \dots + C_N) \\ &\dots\dots \\ [V_N]_1 &= [V_{N-1}]_1 - V_{RN} = -R_1(C_1 + C_2 + \dots + C_N) - R_2(C_2 + \dots + C_N) - \dots - R_N C_N \end{aligned}$$

这样就求得了 1 阶矩

$$m_1 = [V_N]_1 = -\sum_1^N R_i \sum_{j=i}^N C_j \quad (32)$$

接下来需要求解 k_1 和 p_1 ，这需要两个方程：

$$\begin{cases} -k_1 = m_0 \\ -\frac{k_1}{p_1} = m_1 \end{cases} \quad (33)$$

把 (x) 式和 (x) 式代入，有

$$\begin{cases} k_1 = -m_0 = -1 \\ p_1 = m_1 = -\sum_1^N R_i \sum_{j=i}^N C_j \end{cases} \quad (34)$$

由此求得的时间常数也是

$$\tau_N = \frac{p_1}{k_1} = \sum_1^N R_i \sum_{j=i}^N C_j \quad (35)$$

这就说明了 Elmore 延迟其实是 AWE 的一种特例，即一阶的情形。

3.1.5 AWE 的性能

对于如图 21 的一个简单的 RC ladder：

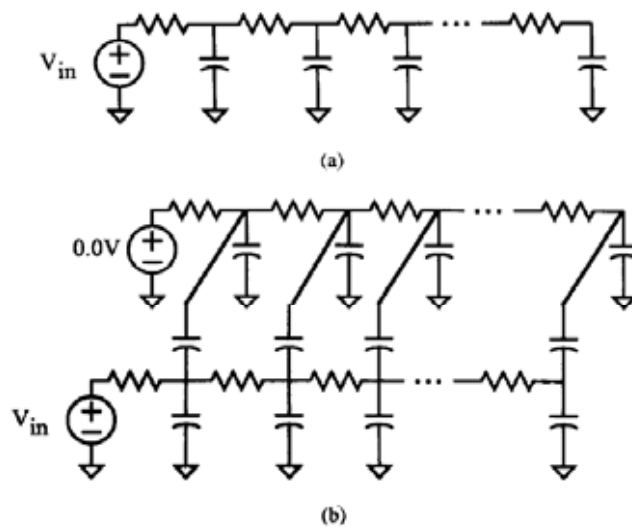


图 21 两种基本的 RC 互连线电路模型

RICE 可以达到比 PSPICE 快得多的速度[2]：

表 1 RICE 和 PSPICE 运行时间对比

类型	节点数	支路数	RICE (秒)				PSPICE (秒)
			分析	建立	计算	总计	
RC 树	2000	4000	0.49	0.11	0.07	0.67	98.15
	8000	16000	2.33	0.48	0.28	0.28	1170.67
耦合 RC 树	1600	4000	0.58	0.11	0.07	0.76	97.60
	6400	16000	3.38	0.44	0.30	3.02	908.7

同时还能保证相当高的精度：

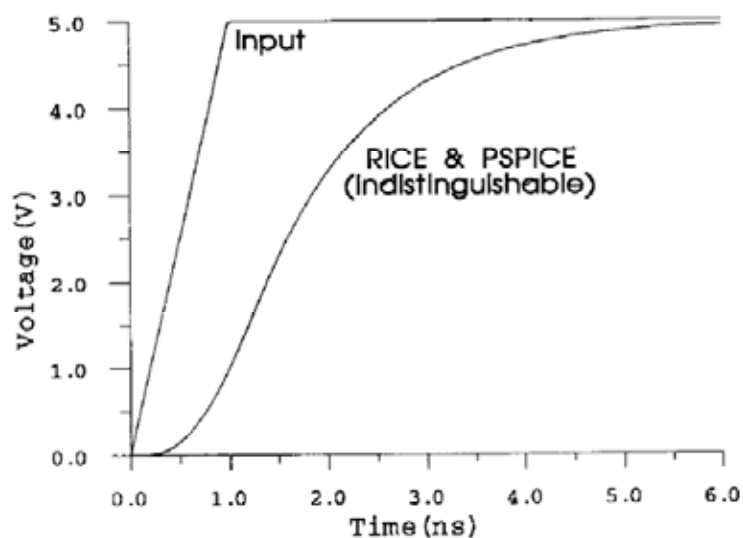


图 22 RICE 和 PSPICE 的仿真波形对比

可见，模型降阶后既提高了仿真速度，还依然可以满足一定的精度，这对于大规模的电路仿真尤其重要。

3.2 Pade Via Lanczos (PVL)

这种方法是 1995 年由 Bell 实验室的数学家 Roland W. Freund 提出[3]。它是一

种通过 Lanczos 过程计算大规模线性网络的 Laplace 域传输函数的 Pade 近似的一种算法。它的最大特点就是拥有非常高的稳定性，但同时保持了和通过矩匹配计算 Pade 近似的 AWE 算法一样的效率。而且这个算法还提供了一个方法可以计算所得到的极点与真正极点之间的误差有多大，从而保证了降阶计算的质量。

3.2.1 AWE 的缺点

虽然 AWE 在电路的模型降阶上取得了巨大的成功，但从实验的结果来看，它仍然存在不少限制，既有数值计算方面的，也有实际应用方面的。

首先 AWE 的一个最主要的不足就是降阶模型的阶数不能太高，也就只能产生少数几个精确的极点和零点。因此，对于 RC 互连线这种简单的电路来说这已经足够胜任了，但是如果对于更加复杂一些的电路，比如有电感或者互感的，AWE 往往不能得到精确的结果。

AWE 是依靠匹配降阶模型和实际电路的矩 (moment) 来确定降阶模型的极点和余式的，这是一种基于 Pade 近似的方法。在理想情况下，如果匹配足够多的矩，就可以任意提高算法的精度， $\tilde{H}(s)$ 就会无限接近 $H(s)$ ，但是事实上并不是这样。一般来说，在 10 阶以下，增加阶数还能够提高精度，但是超过 10 阶之后，精度就无法继续提高了。

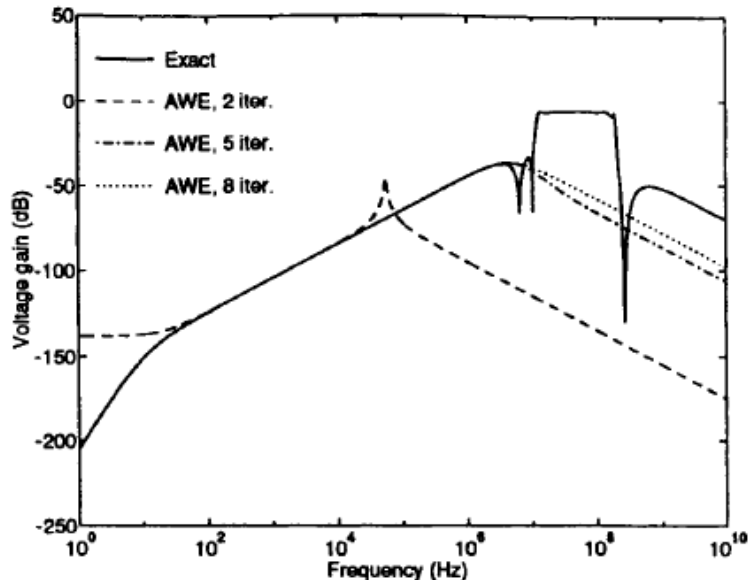


图 23 AWE 不同阶数的降阶结果对比

从图 23 中可以看出，5 阶的 AWE 相对于 2 阶在匹配的频率范围上有了很大的提高，但是当阶数增加到 8 阶后，并没有像想象的一样能够在 10^7Hz 以上的频率范围收敛于实际值，而对于 8 阶以上则已经和 8 阶几乎重合，无法进一步提高精度了。

AWE 在求解极点的过程中，必须要求解一个这样的方程：

$$\mathbf{M}_q \mathbf{a}_q = \begin{bmatrix} m_0 & m_1 & \cdots & m_{q-1} \\ m_1 & m_2 & \cdots & m_q \\ \vdots & \vdots & \ddots & \vdots \\ m_{q-1} & m_q & \cdots & m_{2q-2} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{q-1} \end{bmatrix} = \begin{bmatrix} -m_q \\ -m_{q+1} \\ \vdots \\ -m_{2q-1} \end{bmatrix} \quad (36)$$

在实际计算的时候就会发现，当阶数 q 增大后，方程的系数矩阵 \mathbf{M}_q 的条件数 $\text{cond}(\mathbf{M}_q)$ 急剧变大。 \mathbf{M}_q 的条件数表示舍入误差对方程数值求解结果的影响程度，条件数每增加 10 倍，求解结果的有效数字就丢失一位。所以对于用 16 为存储的浮点数来说，当 $\text{cond}(\mathbf{M}_q)$ 超过 10^{16} 后，计算的结果就已经毫无意义了。对于上面的例子，表 2 列出了 $\text{cond}(\mathbf{M}_q)$ 随 q 增加的趋势[3]：

表 2 图 23 例子中 \mathbf{M}_q 的条件数随阶数的变化

q	$\text{cond}(\mathbf{M}_q)$	$\text{cond}(\mathbf{M}_q^{(1)})$	$\text{cond}(\mathbf{M}_q^{(2)})$	$\text{cond}(\mathbf{M}_q^{(3)})$
2	9.59e+05	2.62e+00	2.62e+00	1.07e+00
3	2.91e+15	5.14e+03	5.12e+03	2.71e+01
4	3.18e+26	1.13e+09	1.13e+09	1.30e+07
5	2.16e+35	2.34e+12	2.33e+12	3.01e+10
6	3.68e+46	2.72e+17	2.38e+17	5.99e+15
7	1.97e+53	2.96e+17	6.19e+17	7.05e+15
8	3.34e+58	1.25e+18	1.78e+18	2.67e+16
9	1.66e+65	1.18e+18	1.52e+18	7.86e+15
10	1.30e+72	1.57e+18	7.50e+17	5.38e+16
20	5.41e+130	1.27e+18	1.92e+18	9.01e+16
30	2.87e+191	7.26e+18	3.51e+18	5.23e+17

AWE 中还提出了一种提出用频率变换的方法减缓条件数恶化的速度，不过即使是这样，也不能带来多大的改善。这些补救的方法在实际运用过程中的可操作性并不强，很多都是启发式的算法，不能自动根据实际问题应用，而且也增加了计算量。

导致这一问题的原因并不在于 Pade 近似的理论有什么差错，而是 AWE

在实际计算过程中带来的副作用。AWE 对于实际电路矩的计算是通过显示计算来进行的，即使是在 RICE 中用遍历和迭代的方法也是一样，实际上都是在计算：

$$\begin{aligned} m_k &= \mathbf{I}^T (-\mathbf{G}^{-1}\mathbf{C})^k (\mathbf{G}^{-1}\mathbf{b}) , k=0,1,\dots \\ &= \mathbf{I}^T \mathbf{A}^k \mathbf{r} \end{aligned} \quad (37)$$

这种对矩阵 \mathbf{A} 迭代相乘的结果就是随着 k 的增大， $\mathbf{A}^k \mathbf{r}$ 很快就收敛于 \mathbf{A} 最大特征值所对应的特征向量。由此计算出来的矩 m_k 也就仅包含了 \mathbf{A} 的一个特征值的信息，这并不符合我们的期望，理想情况下随着阶数的增加，矩应该越来越多的携带 \mathbf{A} 特征值的信息，只有这样才能提高降级模型对原电路的近似程度。

3.2.2 PVL 算法

Lanczos 算法是 1950 年由数学家 Cornelius Lanczos 提出的用于计算矩阵特征值的方法[11]，该算法通过迭代过程把一个方阵 \mathbf{A} 化简为一系列三对角矩阵。

算法的主要过程如下：

Lanczos 过程

输入： A, r, l 输出： V_q, W_q, D_q, T_q

0) 初始化，令 $\rho_1 = \|r\|_2, \eta_1 = \|l\|_2, v_1 = r/\rho_1, w_1 = l/\eta_1$ ，并设 $v_0 = w_0 = 0, \delta_0 = 1$ 。

For $n = 1, 2, \dots, q$ do:

- 1) 计算 $\delta_n = w_n^T v_n$
- 2) 计算 $\alpha_n, \beta_n, \gamma_n$

$$\alpha_n = \frac{w_n^T A v_n}{\delta_n}, \beta_n = \frac{\delta_n}{\delta_{n-1}} \eta_n, \gamma_n = \frac{\delta_n}{\delta_n - 1} \rho_n \quad (38)$$

- 3) 计算 v, w

$$v = A v_n - v_n \alpha_n - v_{n-1} \beta_n, w = A^T w_n - w_n \alpha_n - w_{n-1} \gamma_n \quad (39)$$

- 4) 令 $\rho_{n+1} = \|v\|_2, \eta_{n+1} = \|w\|_2$ ，并计算

$$v_{n+1} = \frac{v}{\rho_{n-1}}, w_{n+1} = \frac{w}{\eta_{n+1}} \quad (40)$$

5) 最后得到：

$$T_q = \begin{bmatrix} \alpha_1 & \beta_2 & 0 & \cdots & 0 \\ \rho_2 & \alpha_2 & \beta_3 & \ddots & \vdots \\ 0 & \rho_3 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \beta_q \\ 0 & \cdots & 0 & \rho_q & \alpha_q \end{bmatrix}, \tilde{T}_q = \begin{bmatrix} \alpha_1 & \gamma_2 & 0 & \cdots & 0 \\ \eta_2 & \alpha_2 & \gamma_3 & \ddots & \vdots \\ 0 & \eta_3 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \gamma_q \\ 0 & \cdots & 0 & \eta_q & \alpha_q \end{bmatrix} \quad (41)$$

$$V_q = [v_1 \ v_2 \ \cdots \ v_q], W_q = [w_1 \ w_2 \ \cdots \ w_q] \quad (42)$$

以及

$$D_q = W_q^T V_q = \text{diag}(\delta_1, \delta_2, \cdots, \delta_q) \quad (43)$$

Lanczos 算法得到的矩阵满足以下性质：

1. 向量 $\{v_n\}_{n=1}^{q+1}$ 和 $\{w_n\}_{n=1}^{q+1}$ 互相正交

$$w_j^T v_k = \begin{cases} \delta_j, & \text{if } j = k \\ 0, & \text{if } j \neq k \end{cases} \quad j, k = 1, 2, \cdots, q+1 \quad (44)$$

2. 三对角矩阵和 A 之间满足

$$\begin{aligned} AV_q &= V_q T_q + [0 \ \cdots \ 0 \ v_{q+1}] \rho_{q+1} \\ A^T W_q &= W_q \tilde{T}_q + [0 \ \cdots \ 0 \ w_{q+1}] \eta_{q+1} \end{aligned} \quad (45)$$

3. 从投影的角度来看， T_q 是 A 最佳的 q 阶近似，而这种近似仅用到了两个 Krylov 子空间的信息：

$$\begin{aligned} K_q(v_1, A) &= \text{span}\{v_1, Av_1, \cdots, A^{q-1}v_1\} \\ K_q(w_1, A^T) &= \text{span}\{w_1, A^T w_1, \cdots, (A^T)^{q-1}w_1\} \end{aligned} \quad (46)$$

4. T_q 实际上是 A 投影到 $K_q(v_1, A)$ 上且正交投影到 $K_q(w_1, A^T)$ 的结果。而这种投影的近似度即使是在 q 远小于 A 的阶数的情况下也是很好的。

T_q 的特征值是对 A 的特征值的近似。这一点可以从式(45)看出来。事实上，如果迭代的次数足够多， v_{q+1} 和 w_{q+1} 就会变得很小，以至于可以忽略。这样式(45)就简化成：

$$AV_q = V_q T_q \quad (47)$$

所以若 λ 和 x 分别是 T_q 的特征值和对应的特征向量，则有

$$\begin{aligned} T_q x &= \lambda x \\ \Rightarrow A(V_q x) &= AV_q x = VT_q x = V_q \lambda x = \lambda(V_q x) \end{aligned} \quad (48)$$

即 λ 也是 A 的特征值，对应的特征向量为 $V_q x$

5. Lanczos算法还有一个最重要的性质，就是 T_q 在匹配最多矩的意义下，也是 A 的最好的近似。正是这一个性质使得Lanczos算法能够在电路模型降阶中发挥独特的作用，前面提到的AWE就是通过匹配矩来确定降阶模型的。

这一性质可以从式(45)中看出。在式(45)的等号两端同时乘上 A ，可以得到

$$\begin{aligned} A^j r &= \rho_1 A^j v_1 \\ &= \rho_1 A^j v_1 \\ &= \rho_1 A^j V_q e_1 \\ &= \rho_1 V_q T_q^j e_1, \quad j = 0, 1, \dots, q-1 \end{aligned} \quad (49)$$

其中右乘 $e_1 = [1 \ 0 \ \dots \ 0]^T \in \mathbb{R}^q$ 表示左边矩阵的第一列。

同样的我们还可以导出

$$\begin{aligned} l^T A^j &= \eta_1 e_1^T (\tilde{T}_q^T)^j W_q^T \\ &= \eta_1 \delta_1 e_1^T T_q^j D_q^{-1} W_q^T, \quad j = 0, 1, \dots, q-1 \end{aligned} \quad (50)$$

因此由矩的表达式

$$m_k = l^T A^k r = (l^T A^{k'}) (A^{k''} r) \quad (51)$$

必然存在这样的 k' 和 k'' 使得满足

$$\begin{aligned} m_k &= (\eta_1 \delta_1 e_1^T T_q^{k'} D_q^{-1} W_q^T) (\rho_1 V_q T_q^{k''} e_1) \\ &= \underbrace{(\eta_1 \delta_1 \rho_1)}_{l^T r} e_1^T T_q^{k'} D_q^{-1} \underbrace{(W_q^T V_q)}_{D_q} T_q^{k''} e_1 \\ &= (l^T r) \cdot (e_1^T T_q^k e_1) \end{aligned} \quad (52)$$

所以有

$$\begin{aligned}
H_q(s) &= l^T r \cdot e_1^T (I - sT_q)^{-1} e_1 \\
&= l^T r \sum_{k=0}^{\infty} e_1^T T_q^k e_1 s^k \\
&= \sum_{k=0}^{2q-1} m_k s^k + O(s^{2q})
\end{aligned} \tag{53}$$

这证明了 $H_q(s) = H(s) + O(s^{2q})$ ，即 $H_q(s)$ 是 $H(s)$ 的近似，且匹配了前 $2q$ 个矩。

PVL 算法

输入： C, G, b, l

输出： p, k

- 1) 令 $A = -G^{-1}C, r = G^{-1}b$
- 2) 执行 q 步 Lanczos 过程， $T_q = \text{Lanczos}(A, l, r)$
- 3) 计算 T_q 的特征值分解

$$T_q = S_q \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_q) S_q^{-1} \tag{54}$$

并令 $\mu = S_q^T e_1, \nu = S_q^{-1} e_1$

- 4) 计算 H_q 的极点和余式，

$$p_j = 1/\lambda_j, \quad k_j = \frac{l^T r \cdot \mu_j \nu_j}{\lambda_j}, \quad j = 1, 2, \dots, q, \quad \lambda_j \neq 0 \tag{55}$$

如果 $\lambda_j = 0$ ，则令

$$k_\infty = \sum_{\substack{j=0 \\ \lambda_j=0}}^q l^T r \cdot \mu_j \nu_j \tag{56}$$

3.2.3 PVL 的优点

PVL 和 AWE 一样，最后也是满足了降阶模型和实际电路的若干个矩相等，从

而保证了降阶模型的传输函数能够达到一定的近似度。但是 PVL 的高明之处就在于避开了显式计算电路的矩，这在 AWE 中被证明是导致精度不能提高的主要原因。PVL 的三对角化过程实际上是用到了两个 Krylov 子空间 $K_q(v_1, A)$ 和 $K_q(w_1, A^T)$ 的信息。这两个空间的基 $\{r_1, Ar_1, \dots, A^{q-1}r_1\}$ 如果乘上 l^T 正好是 AWE 中的矩。但是从递归的过程可以看出，PVL 并不直接计算 $A^k r_1$ ，而是一边计算一边做投影，相当于正交化的过程。所以说 PVL 是隐式的矩匹配，而且在三对角化得到 T_q 后实际上已经得到了降阶模型的传输函数 $H_q(s)$ ，并不需要像 AWE 一样计算极点和余式。如果要求极点也是比较容易的，

AWE 的另一个不足之处是没有一个理论的证明能够预测降阶模型近似的程度究竟有多少。PVL 可以给出极点近似的误差范围。

首先我们可以从(55)中看出， q 阶 Pade 近似 H_q 的极点就是 Lanczos 矩阵 T_q 特征值的逆。另一方面，实际电路传输函数 H 的极点也是 A 的特征值的逆。因此对于 H_q 的极点的近似程度的衡量就可以通过检验 T_q 的特征值和 A 的特征值的接近程度来进行。

令 $\lambda_j \in \mathbb{C}$ ， $s_j \in \mathbb{C}^q$ 是 T_q 的一对特征值和特征向量。因此就有，

$$T_q s_j = \lambda_j s_j \quad (57)$$

然后我们令

$$z_j = V_q s_j \quad (58)$$

其中 V_q 是 Lanczos 过程中定义的。利用(57)，(58)以及式(45)，就可以得到

$$Az_j - \lambda_j z_j = v_{q+1} \rho_{q+1} s_{qj} \quad (59)$$

其中 s_{qj} 是向量 s_j 的最后一个元素。对(59)取模且由于 $\|v_{q+1}\|_2 = 1$ ，则有

$$\frac{\|Az_j - \lambda_j z_j\|_2}{\|A\|_2 \cdot \|s_j\|_2} = \frac{\rho_{q+1} |s_{qj}|}{\|A\|_2 \cdot \|s_j\|_2} \quad (60)$$

更进一步，我们还可以估计 $\|A\|_2$ 的大小，

$$\|A\|_2 \geq \max_{n=1,2,\dots,q} \{\|Av_n\|_2, \|A^T w_n\|_2\} = n(A) \quad (61)$$

结合式(60)和式(61)就有

$$\frac{\|Az_j - \lambda_j z_j\|_2}{\|A\|_2 \cdot \|s_j\|_2} = \frac{\rho_{q+1} |s_{qj}|}{n(A) \cdot \|s_j\|_2} = Q_j \quad (62)$$

这样 Q_j 就表示了 (λ_j, z_j) 与 A 的特征值对的近似程度有多好，也就检验了 PVL 得到的极点 $p_j = 1/\lambda_j$ 的品质。因此就可以用 Q_j 来过滤出那些极点是收敛于 H 的真正的极点的。

3.2.4 PVL 的性能

用 PVL 算法处理一个混合信号 ASIC 标准单元的完整的电源网格[12]，如[13]描述的包括了一些衬底接触和衬底耦合/去耦。这个模型包含了 1074 个电源总线段、36 个单元模型、一个粗略的 $10 \times 10 \times 1$ 的衬底网格。产生的 MNA 矩阵的大小有 1766×1766 。我们想知道单元的开关电流对 VDD 和 GND 的影响。图 24 给出了由 PVL 算法的 50 次迭代得到的相应传输函数的幅度和相位，与 AC 扫描的结果的比较。可以看到符合程度相当高，一直到 5GHz，PVL 的结果和 AC 扫描基本完全重合。

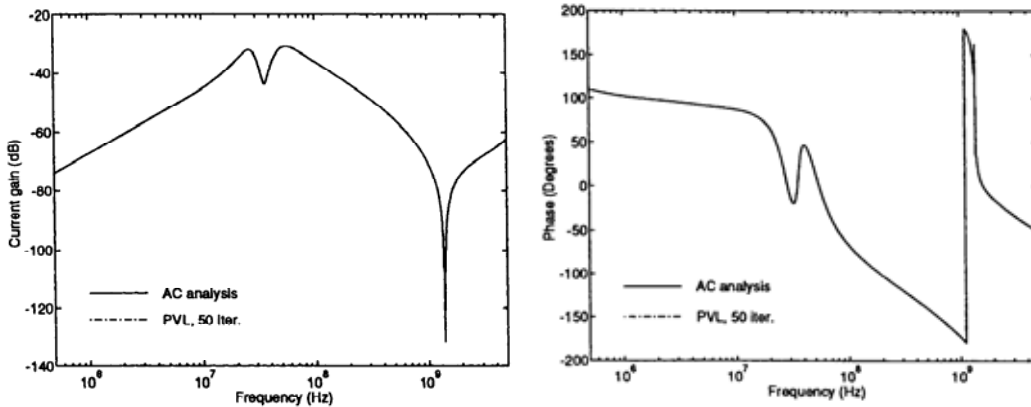


图 24 混合信号电路仿真结果，左图：幅度，右图：相位

3.3 无源互连线宏模型降阶算法 (PRIMA)

这个算法是 1998 年由 Odabasioglu 等人发表的[7]，它主要针对 block Arnoldi 算法[14]不能保持模型无源性的问题加以改进。对于一个电路系统，通常都可以把它分为线性元件和非线性元件两个集合，而所有的线性元件可以组成一个 N 口网络（图 25），它的特性完全由端口上的 Y 参数或者其他参数决定（图 26）。

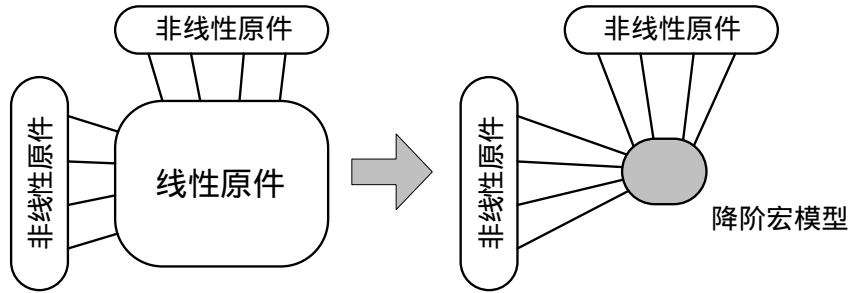


图 25 线性元 件等效为一个降阶的宏模型

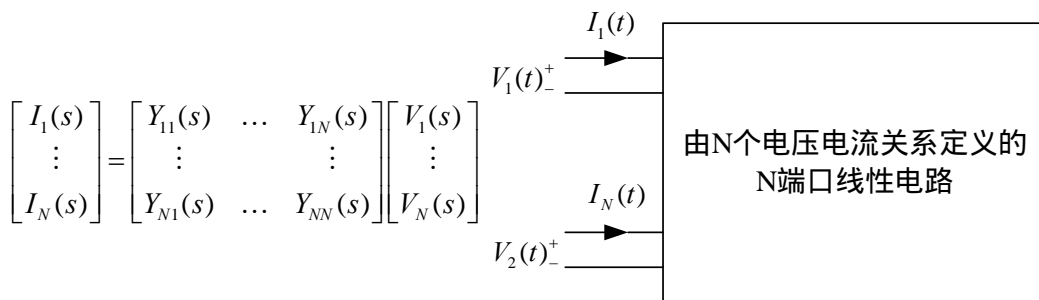


图 26 多口网络的表示式

3.3.1 电路的无源性

一般来说电路中的线性原件的数量远大于非线性原件，所以模型降阶主要是作用在线性原件组成的网路上。把降阶后得到的宏模型再和非线性原件一同放入仿真器。在求解降阶模型的过程中，不但对模型的精度有所要求，而且还需要考虑两个重要的电路特性。

1. 稳定性 (Stability)

一个系统，如果对任意的有界输入，其零状态响应也是有界的，则称该系统是有界输入有界输出 (BIBO) 稳定的系统，简称稳定系统[28]：

对所有的激励信号

$$|e(t)| \leq M_e \tag{63}$$

其响应 $r(t)$ 满足

$$|r(t)| \leq M_r \tag{64}$$

则称该系统是稳定的。 M_e, M_r 都是有界正值。

一般难以验证以上的稳定性定义，但是还有一个比较简单的判据：

若传输函数 $H(s)$ 的全部极点位于 s 平面的左半平面（不包括虚轴），即所有的极点都有负实部，那么系统是稳定的。

稳定性在 AWE 和 PVL 中都很容易检验，而且在 PVL 中还可以证明，如果降阶后带来了正实部的极点，那么有两种情况：1. 实际电路本身的确是不稳定的，这时 Q_j 就会很小。2. 由 PVL 算法带来的不稳定极点。这种极点的余式都很小，所以几乎不影响降阶模型的传输函数 H_q ，因此可以直接把它们去除。

2. 无源性 (Passivity)

无源系统表示一个系统无法自己产生能量，而可以从其他的激励源吸收能量。可以证明，无源系统必然是稳定系统，但是反之不成立[15]。对于电路仿真模型来说，保持无源系统的无源性是需要满足的一个重要特性，因为稳定但非无源的宏模型当和其他稳定、甚至无源的负载连接的时候仍然可以产生不稳定的系统。模型降阶的主要对象是互连线网络，通常都是由 RLC 组成的，因此是无源系统。当把降阶得到的宏模型放入仿真器和其他的模型一起仿真时，并不能够保证整个系统是稳定的，也就不能得到可靠的结果。

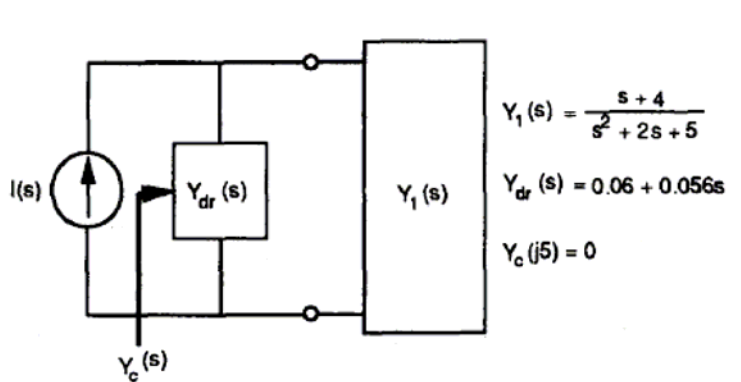


图 27 一个由稳定模型组成的不稳定系统

比如在图 27 中，模型 $Y_1(s)$ 是稳定的，和它并联的另一个系统 $Y_{dr}(s)$ 是由无源元件组成的，因此也是稳定的。两者组成一个整体后的输入点导纳为 $Y_C(s) = Y_1(s) + Y_{dr}(s)$ ，很显然它有一个虚轴上的极点，是不稳定的系统。

为了体现无源性的重要意义，我们考虑一个有损传输线，由 40 个集总 RLC 段的模型表示[7]。模型的降阶算法使用 PVL 和 PRIMA (5 极点)。虽然 Pade 近似得到的所有极点都是稳定的（负实部），但整个系统却是不稳定的。而 PRIMA 的近似度和真实响应十分吻合（图 28）。

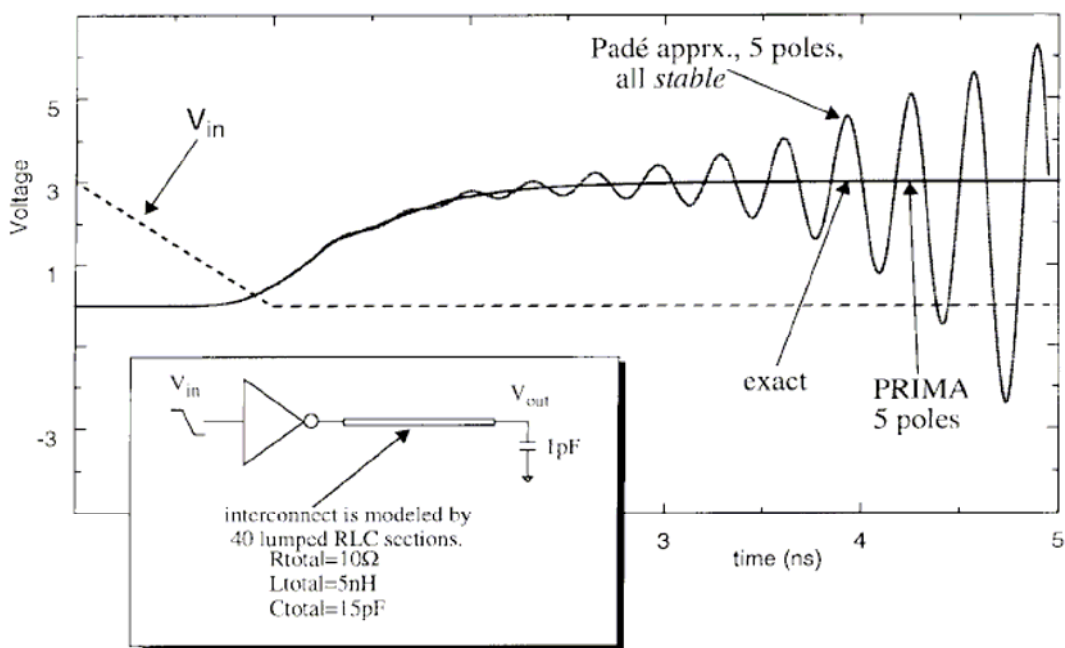


图 28 稳定但非无源的宏模型

3.3.2 PRIMA 算法

PRIMA 也是用 Pade 近似去寻找降阶模型的，而且为了保证数值稳定性，没有去显式计算矩，而是用 Arnoldi 过程求得 Krylov 子空间 $K_q(A, r) = span\{r, Ar, \dots, A^{q-1}r\}$ 的一组正交基。Arnoldi 过程实际上是一种改进的 Gram-Schmidt 正交化，算法流程如下：

Arnoldi 过程

输入： A, r, q

输出： $V_q, v_{q+1}, H_q, h_{j+1,j}$

0) 设 $v_1 = r / \|r\|_2$

For $j = 1, 2, \dots, q$

1) 令 $w = Av_j$ ，正交化：

For $i = 1, 2, \dots, j-1$

$$h_{i,j} = w^T v_i$$

$$w = w - h_{i,j} v_i$$

2) 令 $h_{j+1,j} = \|w\|_2$ ， $v_{j+1} = w / h_{j+1,j}$ ，返回执行 1)

最后得到

$$\begin{aligned} V_q &= [v_1 \quad v_2 \quad \dots \quad v_q] \\ H_q &= (h_{i,j}), \quad j = 1, 2, \dots, q \end{aligned} \quad (65)$$

V_q 中的 q 个向量是相互正交的：

$$V_q^T V_q = I_q \quad (66)$$

且它们就是 $K_q(A, r)$ 的一组基。

H_q 是一个 $q \times q$ 的上 Hessenberg 矩阵，其中的元素是 Gram-Schmidt 正交化的系数。且它们之间满足以下性质：

$$AV_q = V_q H_q + h_{j+1,j} v_{q+1} e_q^T \quad (67)$$

如果不考虑最后列的误差，就有 $AV_q = V_q H_q$ ，两边左乘 V_q^T 就有

$$V_q^T AV_q = H_q \quad (68)$$

Arnoldi 降阶算法

把 Arnoldi 算法应用在降阶领域并不是 PRIMA 首先提出的[5]，一种传统的方法就是在得到了 $K_q(A,r)$ 的基之后，我们对 MNA 方程(3)进行变量替换：

$$x_n = V_q z_q \quad (69)$$

注意这里 x_n 是一个 n 维向量， z_q 是一个 q 维向量， $V_q \in \mathbb{R}^{n \times q}$ 。如果 $q \ll n$ ，则现在系统的状态变量的个数比原来少了很多。把式(69)代入 MNA 方程，并左乘 $V^T G^{-1}$ 有：

$$\begin{aligned} -V_q^T (G^{-1}C)V_q \dot{z}_q &= (V_q^T V_q)z_q - V_q^T (G^{-1}b)u \\ y &= l^T V_q z_q \end{aligned} \quad (70)$$

把式(66)和式(68)代入得到了新的系统方程，这个系统只有 q 阶了：

$$\begin{aligned} H_q \dot{z}_q &= z_q - V_q^T r u \\ y &= l^T V_q z_q \end{aligned} \quad (71)$$

降阶系统的传输函数就是

$$\tilde{H}(s) = l^T V_q (I_q - sH_q)^{-1} V_q^T r \quad (72)$$

PRIMA 算法

PRIMA 的方法是直接对 MNA 矩阵中的 C, G, b, l 进行降阶，因此保持了无源性。设 X 就是 Arnoldi 过程得到的 Krylov 子空间的一组正交基，在变量替换 $x_n = X\tilde{x}_q$ 后，对 MNA 方程左乘 X^T ，方程变为：

$$\begin{aligned} (X^T C X) \dot{\tilde{x}}_q &= -(X^T G X) \tilde{x}_q + (X^T b)u \\ y &= (l^T X) \tilde{x}_q \end{aligned} \quad (73)$$

所以得到的降阶模型的 MNA 矩阵为：

$$\begin{aligned} \tilde{C} &= X^T C X & \tilde{G} &= X^T G X \\ \tilde{b} &= X^T b & \tilde{l} &= X^T l \end{aligned} \quad (74)$$

这里之所以用 X 表示Arnoldi算法中的 V_q 是因为PRIMA算法的本质是要寻找一个合同变换矩阵，它的列向量是Krylov子空间 $K_q(A, r)$ 的一组正交基，而Arnoldi过程只不过是产生这组基的一种方法而已，并不局限于Arnoldi，任何可以找到这组基的算法都可以用于PRIMA。可见PRIMA对变换矩阵的要求并不十分苛刻，实际上只需要满足以下两个条件即可：

$$\begin{aligned} \text{colsp}(X) &= K_q(A, r) \\ X^T X &= I_q \end{aligned} \quad (75)$$

3.3.3 保持矩

由于是直接对MNA矩阵降阶，所以降阶模型的矩的计算方法和原来一样：

$$\tilde{m}_i = \tilde{l}^T (-\tilde{G}^{-1}\tilde{C})^i (\tilde{G}^{-1}\tilde{b}) = \tilde{l}^T \tilde{A}^i \tilde{r} \quad (76)$$

对式(76)展开并作适量代数运算后可以证明

$$\tilde{m}_i = l^T A^i r = m_i, \quad i = 1, 2, \dots, q \quad (77)$$

说明 q 阶的降阶模型（指的是降阶后状态变量的个数为 q 个）能够保持实际电路的前 q 个矩。PVL 的 q 阶模型可以保持原电路的前 $2q$ 个矩，所以 PRIMA 在阶数相等的条件下精度还是要比 PVL 差一点，没有匹配最多的矩。

3.3.4 保持无源性

PRIMA 的优越性在于它能够保持降阶模型的无源性。如果用无源网络端口上的导纳矩阵来表示，降阶系统无源的充分必要条件是：

- 1) 对于所有的复数 s ，都有 $\tilde{Y}(s^*) = \tilde{Y}^*(s)$ ，“*”表示共轭。
- 2) $\tilde{Y}(s)$ 对于所有的满足 $\text{Re}(s) > 0$ 的复数 s 和所有的复向量 z 都有

$$\text{Re}(z^H \tilde{Y}(s)z) \geq 0 \quad (78)$$

由于降阶后的矩阵 $\tilde{G}, \tilde{C}, \tilde{b}, \tilde{l}$ 都是实的，因为变换矩阵 X 是实的，所以条件 1) 总是满足的。而第二个条件通过一些代数运算也可以证明，详细的证明过程请参看文献[7]。这里需要注意的一点是在证明无源性的过程中，没有对变换矩阵 X 作出什么要求，因此只要降阶系统是通过式(74)的变换得出的，总是可以保证无源性成立。这个性质为 X 的选取提供了便利。

3.4 保持结构的互连线降阶模型 (SPRIM)

PRIMA 算法的优点是能够在匹配矩的同时保持电路的无源性 (passivity)，但是它没有保持电路结构的特征，这使得 PRIMA 的模型难以综合成实际电路。而且另一方面，相对于其他的矩匹配降阶技术，PRIMA 也只匹配了它们一半的矩。SPRIM 算法[8]对 PRIMA 做了改进，既拥有 PRIMA 的优点：保持无源性，还能够保持电路结构，而且在相同的计算量下，矩匹配的数量是 PRIMA 的两倍，因此精度也有所提高。

3.4.1 变换矩阵的选取

PRIMA 是用合同变换的方法直接对 MNA 方程的矩阵降阶的，这个变换矩阵 X 需要满足的条件是它的列向量必须是 Krylov 子空间的基：

$$\text{colsp}(X) = K_q(A, r) \quad (79)$$

已经证明，由此得出的降阶模型

$$\begin{aligned} \tilde{C} &= X^T C X & \tilde{G} &= X^T G X \\ \tilde{b} &= X^T b & \tilde{l} &= X^T l \end{aligned}$$

能够保持 q 个矩不变。

通过研究证明过程不难发现，式(79)的条件还可以放宽，变换矩阵 X 可以包含更多的列，只要 X 满足列向量空间包含 Krylov 子空间 $K_q(A, r)$ 即可：

$$K_q(A, r) \subseteq \text{colspan}(X) \quad (80)$$

而降阶的方法和 PRIMA 一样。

在第一章中已经提到，MNA 矩阵并不是任意形状的，而是具有一定的结构特征。简单起见，只考虑电流源输入的 Z 参数模型，其 MNA 矩阵为

$$\begin{aligned} (G + sC)X &= BI_s(s) \\ V_i &= B^T X \end{aligned} \quad (81)$$

其中各 MNA 矩阵都可以写成分块的形式

$$G = \begin{bmatrix} G_1 & G_2^T \\ -G_2 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix} \quad (82)$$

传输函数等于：

$$Z(s) = B^T (G + sC)^{-1} B \quad (83)$$

现在我们把原来 $n \times q$ 阶的变换矩阵 X 按照 G 的分块大小写成：

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \quad (84)$$

然后令新的变换矩阵等于

$$\tilde{X} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \quad (85)$$

这时 \tilde{X} 是一个 $n \times 2q$ 的矩阵，且容易证明

$$\text{colspan}(\tilde{X}) \supseteq \text{colspan}(X) \quad (86)$$

所以式(80)的条件可以满足。所以由这个新的变换矩阵得到的降阶模型一定至少可以匹配和 PRIMA 一样多的矩。而事实上还可以证明它可以匹配两倍于 PRIMA 的矩，更可贵的是降阶后模型的分块形式得以保留。

3.4.2 SPRIM 的性能

虽然降阶后的模型比 PRIMA 大了一倍，但降阶的计算量主要集中在求解 Krylov 子空间上，所以并没有比 PRIMA 增加多少开销。另一方面，SPRIM 可以写成二阶形式，即

$$\tilde{Z}(s) = \tilde{B}_1^T (s\tilde{C}_1 + \tilde{G}_1 + \frac{1}{s} \tilde{G}_2^T \tilde{C}_2^{-1} \tilde{G}_2)^{-1} \tilde{B}_1 \quad (87)$$

其中 $\tilde{G}_1, \tilde{G}_2, \tilde{C}_1, \tilde{C}_2, \tilde{B}_1$ 是降阶后所保持的分块。可见，利用二阶形式后，SPRIM 和 PRIMA 的状态空间是相等的。

3.5 其他的降阶方法

3.5.1 节点消去法

前文介绍的模型降阶方法都是属于 Krylov 子空间投影技术，本质上都是 AWE 不断改进的结果。这些方法的一个缺点是降阶后的模型是不可实现的。这就是说，虽然这些方法保持了端口上的电器特性（对传输函数近似），但是最终结果都是把

一个 RLC 电路变成了数学上等价的宏模型——矩阵、极点等。对这些降阶模型的实际应用需要特殊的仿真器，或者至少对标准的仿真器进行改造。

实际上还有一种完全不同的思路，称为可实现的 (realizable) 模型降阶。这种方法希望能够做到 SPICE-in, SPICE-out 的降阶过程，降阶后得到的还是可以直接仿真的 SPICE 网表，不但更加直观，而且无需对仿真器作任何改造，拿来直接使用。

TICER (Time Constant Equilibration Reduction) [10]就是基于这种思想提出的一个比较实用的方法。它的优点在于：

- 降阶得到可实现的 RC 网络
- 保持指定的内部节点不变 (即确保关键节点不被影响)
- 可以处理比较大的网络
- 保持直流和交流特性
- 控制精度

TICER 指出了节点的时间常数的重要性。由简单的电路分析知识，对于一个 RC 网络，任何一个节点根据和它相连的电阻、电容的大小，都可以计算出一个时间常数。简单来说，一个节点的时间常数就是该节点到其他节点以及到地的总电容除以和它连接的电导之和。

我们分析电路的时候，通常只考虑某一个频率范围内的电路行为。针对这个频率范围，可以把节点根据时间常数分类为快节点、慢节点和普通节点。过快和过慢的节点实际上在关心的频率范围内对电路的影响并不大，所以可以把这些节点从网络中去除，来减小电路的规模。这种消去电路中快节点和慢节点的过程称为时间常数均衡，原始电路节点的时间常数分布的很广，均衡后将集中在某一个频率范围附近。

TICER 在消去节点的时候利用了忽略小量的近似方法，在消去的同时，周围的节点之间添加电阻和电容。每次处理一个节点，并优先考虑度数低的节点，对于稀疏图 (边数比较少，对于大多数电路都是成立的)，TICER 的计算量和消去的节点数成正比。

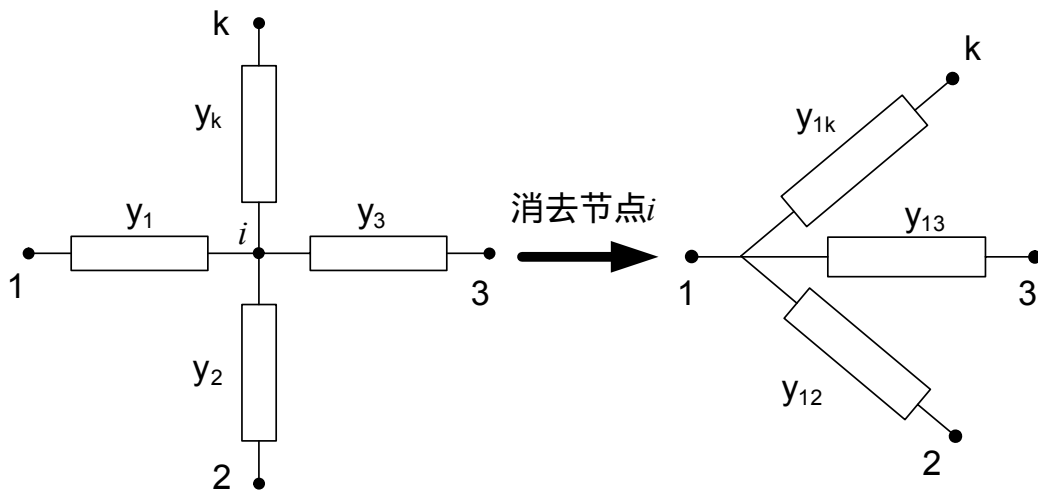


图 29 节点消去的过程

这种节点消去的降阶法特别适用于参数抽取工具。这些工具通常在版图发生几何变换的地方插入节点，比如转角、过孔等，因此抽取得到的参数有很多小的值，也有一些比较大的值。所以节点的时间常数变化会比较大，正好可以利用此方法消去一些节点减小网表的规模。节点消去法可以作为其他模型降阶方法的一种前导工序，寄生参数抽取后通过节点消去，再输入投影降阶算法，可以获得更好的效果。

TICER 只是针对 RC 电路，近几年来，对于包含电感的电路是研究的热点。RCLK 压缩[16]，广义 $Y-\Delta$ 变换[17]，支路融合[18]等技术的提出逐渐解决了电感和互感的问题。到目前为止，这一领域仍是值得研究的方向。

3.6 基于可观测标准型的模型降阶算法

模型降阶问题最初是来源于自动控制领域。控制理论解决线性系统问题往往是抽象成一个状态空间求解，在这方面有着成熟的理论。电路的改进节点方程 (MNA) 实际上也可以看成是状态方程。下面我们将提出一种基于状态变量分析的模型降阶算法。

3.6.1 系统状态方程

系统状态方程是一阶联立微分方程组，方程左边是各状态变量的一阶导数，而右边是状态变量和激励函数的线性组合。状态方程的建立有多种方法，直接法多用于计算机辅助网络分析设计，间接法可由系统的输入输出方程或者系统函数来建立状态变量方程。

系统状态方程一般总是可以表示成矢量矩阵的形式

$$\begin{aligned}\frac{d}{dt}x(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t)\end{aligned}\quad (88)$$

从方程的形式上可以看出，这和电路的改进节点方程十分相像，区别仅在于微分方程系数矩阵的位置一个在等号右边，一个在等号左边。需要指出的是这并没有本质区别，MNA 方程中的矩阵 C 如果可逆的话直接就可以得到相应的状态方程。对于一般不可逆的情况，总是可以通过适当的变量替换，消去一些冗余的节点电压或者支路电流，使左边的系数矩阵可逆，从而也能够转换成状态方程的形式。

3.6.2 系统的可观测性

系统的可观测性是指通过观测有限时间内的输出量，能否识别出系统初始状态的能力，也表示系统内部的状态的不同能否传播到系统输出的能力。如果原系统用状态方程表示，在给定输入后，能在有限时间 $(0, t_1)$ 内根据系统的输出唯一确定出系统的所有起始状态，则称系统完全可观测，若只能确定部分起始状态，则此系统不完全可观测。一个系统是否可观测的判据可以从定义求得：

用 Laplace 变换求解状态方程有

$$Y(s) = Y_{zp}(s) + Y_{zs}(s) = C(sI - A)^{-1}u(0^-) + [C(sI - A)^{-1}B + D]U(s) \quad (89)$$

变换到时域后输出为

$$y(t) = Ce^{At}x(0^-) + [Ce^{At}B + D\delta(t)] * u(t) \quad (90)$$

其中等号右边的第一项是零输入响应，第二项是零状态响应。

因为在讨论可观测问题时，输入是给定的，所以上式的第二项是确定的，为了讨论方便，可以假设 $u(t) = 0$ ，所以有

$$y(t) = Ce^{At}u(0^-) \quad (91)$$

由哈密尔顿-凯莱定理可得

$$y(t) = C(a_0 I + a_1 A + \cdots + a_{k-1} A^{k-1})u(0^-) \quad (92)$$

其中 c_0, c_1, \dots, c_{k-1} 是矩阵 A 的特征方程的系数。

改写成向量的形式为

$$y(t) = \begin{bmatrix} a_0 & a_1 & \cdots & a_{k-1} \end{bmatrix} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix} u(0^-) \quad (93)$$

上式表明，输出 $y(t)$ 是系统所有起始状态的线性组合，因而要求在有限时间内根据输出唯一确定 $u(0^-)$ ，就要求矩阵

$$N = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix} \quad (94)$$

满秩。这是系统完全可观测的充要条件，矩阵 N 就是判别矩阵。

仔细观察判别矩阵可以发现，这和投影降阶算法中的 Krylov 子空间的基颇为相似。我们在分析 AWE 算法缺点的时候就已经提到过，由于当 k 增大后，后面的向量很快就变得和前面的向量线性相关，导致了数值不稳定。也就是说，判别矩阵 N 是很难满足完全可观测的条件的。事实上的确是这样，对于大多数的电路来说，可观测性都非常低，可以理解为对于输出来说，系统有很大的冗余度，因此如果能去掉冗余的部分，系统自然就被“压缩”了。模型降阶需要做的正是这些。

3.6.3 可观测标准型

在状态方程的建立过程中，状态向量的选取并不唯一，同一个系统可以选择不同的状态向量，列出不同的状态方程。既然这些状态方程描述的是同一个系统，则这些状态向量之间有一定的关系。已知一个状态向量，可以通过线性变换产生另一个状态向量：

若 n 阶列向量 x 是系统的状态向量， P 是一个 $n \times n$ 的可逆方阵，则

$$z = Tx$$

也是系统的一个状态向量。

把 (*) 式代入状态方程有

$$\begin{aligned} P^{-1}\dot{z} &= AP^{-1}x + Bu \\ y &= CP^{-1}z + Du \end{aligned} \quad (95)$$

为了化成标准的状态方程形式，对上式左乘 P 后得

$$\begin{aligned} \dot{z} &= (PAP^{-1})z + (PB)u \\ y &= (CP^{-1})z + Du \end{aligned} \quad (96)$$

然后我们令

$$\begin{aligned} \hat{A} &= PAP^{-1}, \hat{B} = PB \\ \hat{C} &= CP^{-1}, \hat{D} = D \end{aligned} \quad (97)$$

就得到了一个在新的状态向量下的状态方程：

$$\begin{aligned} \frac{d}{dt}z(t) &= \hat{A}z(t) + \hat{B}u(t) \\ y(t) &= \hat{C}z(t) + \hat{D}u(t) \end{aligned} \quad (98)$$

这里输入和输出都不发生变化，只有状态变量换成了新的。

线性系统理论可以证明，通过选择适当的变换矩阵 P ，总是可以把状态方程化成一种特殊的形式：

$$\begin{aligned} \hat{A} = \mathbf{PAP}^{-1} &= \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & & \ddots & \\ & & & & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \cdots & -\alpha_{n-1} \end{bmatrix}, \\ \hat{B} = PB &= \begin{bmatrix} CB \\ CAB \\ \vdots \\ CA^{n-1}B \end{bmatrix}, \quad \hat{C} = CP^{-1} = [1 \ 0 \ \cdots \ 0] \end{aligned} \quad (99)$$

其中 $a_0, \alpha_1, \dots, \alpha_{n-1}$ 是矩阵 A 的特征方程的系数。这种形式称为可观测标准型 [29]。

3.6.4 MOROC 算法

基于以上分析，以及电路的可观测性较弱的假设，我们提出基于可观测标准型

的模型降阶算法 (Model Order Reduction using Observability Canonical form)。得到的降阶模型形式上和可观标准型一样，但阶数要小很多。

这里我们仍然采用改进节点法习惯的表示方法，需要指出的是，正如前文所述，这同状态空间表示法没有本质的区别，两种表示中的矩只是 Laplace 变换的展开点不同而已。

步骤：

1) 用 Arnoldi 算法计算 Krylov 子空间

$$K_q(\mathbf{A}^T, \mathbf{l}) = \text{span}[\mathbf{l}, \mathbf{A}^T \mathbf{l}, \dots, (\mathbf{A}^T)^{q-1} \mathbf{l}] \quad (100)$$

2) 找到系数 $\alpha = [\alpha_0 \ \alpha_1 \ \dots \ \alpha_{q-1}]$ ，满足：

$$(\mathbf{A}^q)^T \mathbf{l} \approx -\alpha_0 \mathbf{l} - \alpha_1 \mathbf{A}^T \mathbf{l} - \dots - \alpha_{q-1} (\mathbf{A}^T)^{q-1} \mathbf{l} \quad (101)$$

3) 降阶后的系统定义为：

$$\mathbf{A}_q = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & & \ddots & \\ & & & & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{q-1} \end{bmatrix}, \mathbf{r}_q = \begin{bmatrix} \mathbf{l}^T \mathbf{r} \\ \mathbf{l}^T \mathbf{A} \mathbf{r} \\ \vdots \\ \mathbf{l}^T \mathbf{A}^{q-1} \mathbf{r} \end{bmatrix}, \mathbf{l}_q = [1 \ 0 \ \dots \ 0]^T = \mathbf{e}_1 \quad (102)$$

3.6.5 矩匹配

由降阶模型的特殊形式不难证明，MOROC 总是可以匹配原系统的 q 个矩。可以发现，矩阵 \mathbf{A}_q 每自乘一次，由 1 组成的对角线就向右平移一列，而且降阶后 \mathbf{r}_q 的每一个元素实际上都是原系统的矩，因此在 MOROC 算法中，对于 $k = 0, 1, \dots, q-1$ 都有

$$\begin{aligned}
\tilde{m}_k &= \mathbf{l}_q^T \mathbf{A}_q^k \mathbf{r}_q \\
&= \mathbf{e}_1^T \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \\ & \ddots & & \ddots & \vdots \\ & & & 0 & 1 \\ & & * & & \\ & & & & \end{bmatrix} \begin{bmatrix} m_0 \\ m_1 \\ \vdots \\ m_{q-1} \end{bmatrix} \\
&= m_k
\end{aligned} \tag{103}$$

3.6.6 α 和 \mathbf{r}_q 的计算

为了得到 $(\mathbf{A}^q)^T \mathbf{l}$ 由 Krylov 子空间的一组基 $(\mathbf{A}^i)^T \mathbf{l}$ ($i = 0, 1, \dots, q-1$) 线性表出的系数，首先要把 Krylov 子空间正交化，采用 Arnoldi 算法[5]可以避免正交化过程中 \mathbf{A}^i 当 i 增大而逐渐恶化。记 $\mathbf{W}_q = [\mathbf{l} \quad \mathbf{A}^T \mathbf{l} \quad \cdots \quad (\mathbf{A}^{q-1})^T \mathbf{l}]$ ，得到的一组标准正交基为

$\tilde{\mathbf{W}}_q = [w_1 \quad w_2 \quad \cdots \quad w_q]$ ，Arnoldi 算法保证了它们之间有以下关系：

$$\mathbf{W}_q = \tilde{\mathbf{W}}_q \tilde{\mathbf{H}}_q \tag{104}$$

其中 $\tilde{\mathbf{H}}_q$ 是一个上三角矩阵。

Arnoldi 算法并不能直接得到 $\tilde{\mathbf{H}}_q$ ，而是一个 Hessenberg 阵 \mathbf{H}_q ，满足：

$$\mathbf{A} \tilde{\mathbf{W}}_q = \tilde{\mathbf{W}}_q \mathbf{H}_q + h_{q+1,q} w_{q+1} e_q^T \tag{105}$$

为了找到 \mathbf{H}_q 和 $\tilde{\mathbf{H}}_q$ 之间的关系，把迭代过程展开，对于 $j = 1, 2, \dots, q$ ：

$$\begin{aligned}
\mathbf{w}_{j+1} &= \frac{1}{h_{j+1,j}} \left(\mathbf{A} \mathbf{w}_j - \sum_{k=1}^j h_{kj} \mathbf{w}_k \right) \\
&= \frac{1}{h_{j+1,j}} \left\{ \mathbf{A} \left[\frac{1}{\tilde{h}_{jj}} \left(\mathbf{A}^{j-1} \mathbf{b} - \sum_{k=1}^{j-1} \tilde{h}_{kj} \mathbf{w}_k \right) \right] - \sum_{k=1}^j h_{kj} \mathbf{w}_k \right\} \\
&= \frac{1}{h_{j+1,j} \tilde{h}_{jj}} \left\{ \mathbf{A}^j \mathbf{b} - \sum_{k=1}^{j-1} \left[\tilde{h}_{kj} \left(\sum_{l=1}^{k+1} h_{lk} \mathbf{w}_l \right) \right] - \sum_{k=1}^j \tilde{h}_{jj} h_{kj} \mathbf{w}_k \right\} \\
&= \frac{1}{h_{j+1,j} \tilde{h}_{jj}} \left\{ \mathbf{A}^j \mathbf{b} - \sum_{k=1}^j \left(\sum_{l=k-1}^j \tilde{h}_{kj} h_{lk} \right) \mathbf{w}_k \right\}
\end{aligned} \tag{106}$$

由此可得

$$\mathbf{A}^j \mathbf{b} = \sum_{k=1}^{j+1} \left(\sum_{l=k-1}^j \tilde{h}_{kj} h_{lk} \right) \mathbf{w}_k = \sum_{k=1}^{j+1} \tilde{h}_{k,j+1} \mathbf{w}_k \tag{107}$$

这表明 $\tilde{\mathbf{H}}_q$ 的每一列可以通过前一列迭代产生：

$$\tilde{h}_{k,j+1} = \sum_{l=k-1}^j \tilde{h}_{lj} h_{kl} \tag{108}$$

其中 $\tilde{h}_{11} = \|b\|_2$, $k=1,2,\dots,j$; $j=1,2,\dots,q-1$

由于 \mathbf{H}_q 是 Hessenberg 阵 , 式(108)写成矩阵形式就是

$$\tilde{\mathbf{H}}_q^{(j+1)} = \mathbf{H}_q \tilde{\mathbf{H}}_q^{(j)} \tag{109}$$

其中 $\tilde{\mathbf{H}}_q^{(j)}$ 表示 $\tilde{\mathbf{H}}_q$ 的第 j 列。

在计算的时候 , 正交化多进行一步 , 以得到 $(\mathbf{A}^q)^T \mathbf{1} = \tilde{\mathbf{W}}_q \boldsymbol{\beta}_q$, 其中 $\boldsymbol{\beta}_q = \mathbf{H}_q \tilde{\mathbf{H}}_q^{(q)}$

。又因为 $(\mathbf{A}^q)^T \mathbf{1} = -\mathbf{W}_q \boldsymbol{\alpha}$, 所以有

$$\boldsymbol{\alpha} = -\tilde{\mathbf{H}}_q^{-1} \boldsymbol{\beta}_q \tag{110}$$

通常 $\tilde{\mathbf{H}}_q$ 比较接近于奇异 , 但由于它是一个上三角阵 , 所以实际上并不要求求逆。

对于 \mathbf{r}_q ，注意到它是原系统各阶矩组成的列向量，直接计算会增加计算量，且数值结果逐渐恶化。如果利用 Arnoldi 的结果则不需要直接计算 \mathbf{r}_q ，因为

$$\mathbf{r}_q = \mathbf{W}_q^T \mathbf{r} = \tilde{\mathbf{H}}_q^T \tilde{\mathbf{W}}_q^T \mathbf{r} \quad (111)$$

由降阶模型的形式还可以看出，当阶数 q 提高时，只要把 Arnoldi 算法多进行几步就可以得到新的 $\tilde{\mathbf{H}}_q$ ， \mathbf{r}_q 也只需要更新最后一行，仅 是要重新计算的。

3.6.7 检验稳定性

一个稳定的系统要求它的极点都有负实部。在 AWE 的分析中我们已经知道降阶系统的极点和 \mathbf{A}_q 的特征值的倒数相等，这就要求 \mathbf{A}_q 的特征值都在左半平面。

从 \mathbf{A}_q 的形式容易看出，它的特征多项式为：

$$f(\lambda) = \lambda^q + \alpha_{q-1} \lambda^{q-1} + \cdots + \alpha_1 \lambda + \alpha_0 \quad (112)$$

因此降阶模型的稳定性可以直接用 来检验，而不必求出每一个极点。Hurwitz 判据提供了一种检验的方法[28]。比如对于多项式

$$B(s) = b_n s^n + b_{n-1} s^{n-1} + \cdots + b_0$$

将其分为 $M(s)$ 和 $N(s)$ 两部分：

$$\begin{aligned} M(s) &= b_n s^n + b_{n-2} s^{n-2} + b_{n-4} s^{n-4} + \cdots \\ N(s) &= b_{n-1} s^{n-1} + b_{n-3} s^{n-3} + b_{n-5} s^{n-5} + \cdots \end{aligned} \quad (113)$$

并辗转相除得

$$\frac{M(s)}{N(s)} = q_1 s + \frac{1}{q_2 s + \frac{1}{\ddots \frac{1}{q_n s}}} \quad (114)$$

Hurwitz 判据告诉我们， $B(s)=0$ 的全部根位于 s 左半平面的充要条件是全部系数 q_i 为正值 ($i=1, 2, \dots, n$)。

4 实验与分析

在本章中，我们将在 Matlab 平台上，对以上的几种模型降阶算法分别进行实验，观察对比它们的特点和性能。

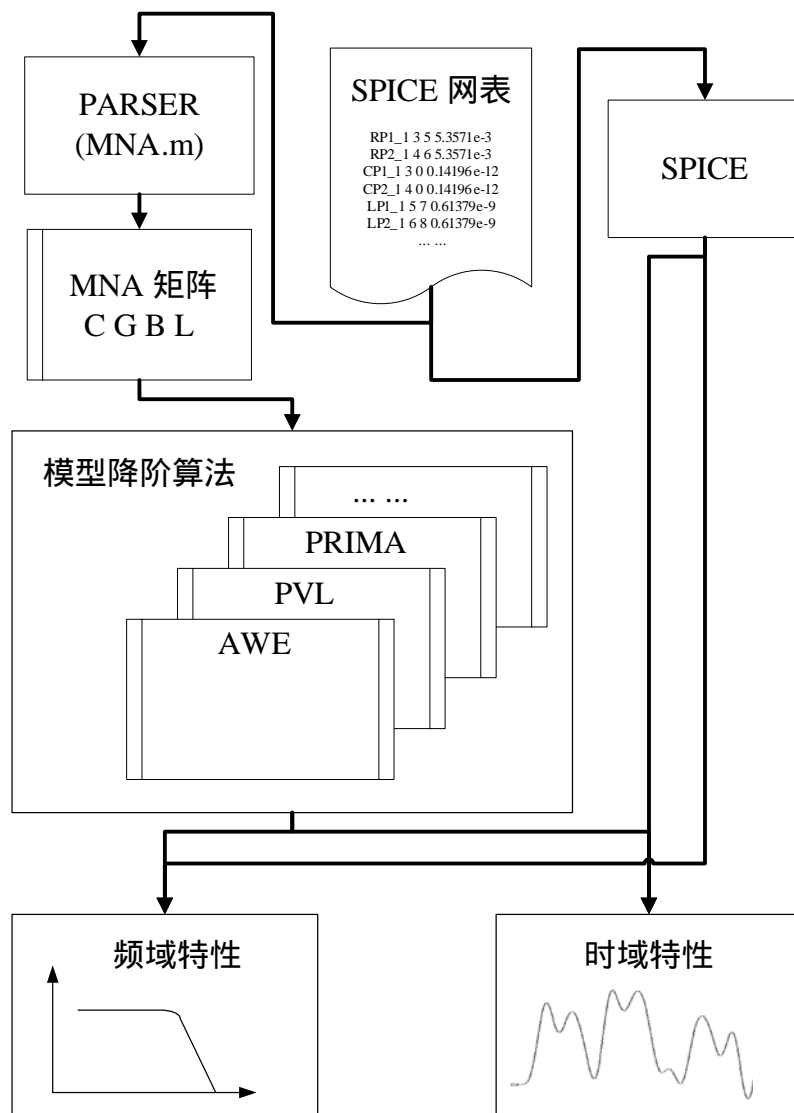


图 30 模型降阶实验流程

实验所用平台：

CPU：AMD Athlon 3000+，2.0GHz

内存：512Mb DDR400

操作系统：WindowsXP Service Pack 2

仿真平台：Matlab 7.1

参考软件：Hspice Ver. 2003.09

4.1 RLCG 互连线

这是一个互连线的 RLCG 模型，由两根电容耦合的导线组成，导线之间通过这些电容产生串扰。如图 31 所示，每根导线都由几段 π 模型建模。总共有 14 个电阻，12 个电容，6 个电感组成。这里电感只考虑了导线上的自感。

其中每一段的电阻电容的参数如下：

$$R_{\pi} = 5.3571 \text{ m}\Omega, \quad L_{\pi} = 0.6138 \text{ nH}, \quad C_{\pi} = 0.1420 \text{ pF}, \\ G = 10^{-75} \text{ }\Omega^{-1}, \quad C_x = 0.0155 \text{ pF}, \quad C_L = 1 \text{ pF}$$

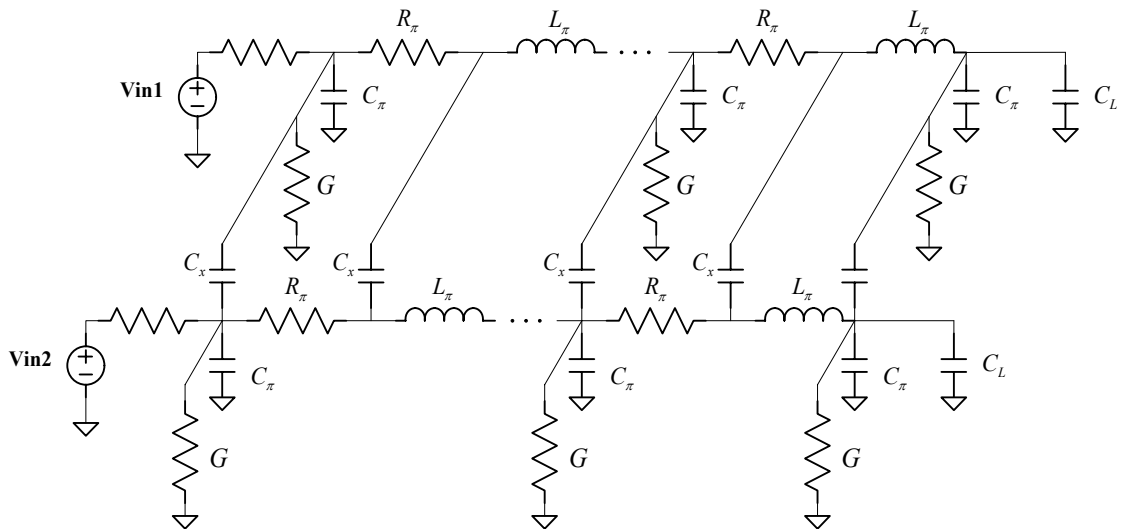


图 31 两根电容耦合的 RLCG 互连线

仿真时为了观察两根互连线之间的影响，我们令一根互连线保持低电平，另一根互连线的输入端接一个理想阶跃信号，从 0V 上跳到 5V。

4.1.1 幅度频率响应

首先我们用不同的方法：AWE、PVL、PRIMA，并以不同的阶数对该电路进行频域分析。这个电路的幅频曲线在低频部分（从 1Hz~100MHz）都比较平坦，由电感引起的一些尖峰发生在 1GHz~100GHz 的频率段中。由于各算法在低频部分的近似度都比较好，所以以下我们只观察分析高频段的幅频曲线。

第一根导线输出段的响应

AWE

应用 AWE 的 2 阶、6 阶、10 阶分别对图 31 中的模型降阶后的幅度频率响应如图 32：

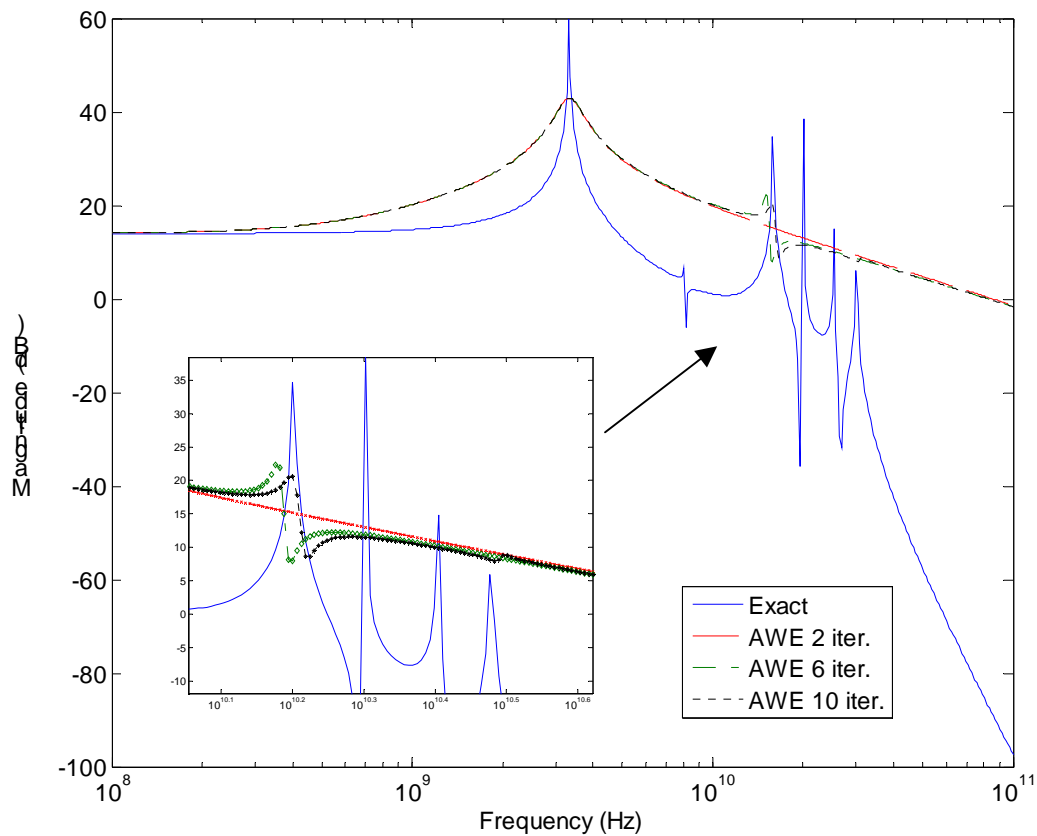


图 32 AWE 对两根电容耦合的 RLCG 互连线的仿真结果

AWE 在低频部分的拟合度保持的不错，但是对于高频部分的尖峰基本上无法逼近。当阶数从 2 阶提高到 6 阶后，似乎有开始收敛于第三个尖峰的趋势，不过即使阶数再继续提高到 10 阶，从小图中也可以看到，只是从位置上比 6 阶的曲线更加靠近了，但是提高十分有限。同时第一个尖峰也没有能够进一步提高近似度。

表 3 AWE 算法中 M 矩阵条件数随阶数 q 的变化

阶数	M 矩阵的条件数	阶数	M 矩阵的条件数
2	4.1632e+020	6	4.1574e+108
3	3.3642e+042	7	4.0811e+130
4	3.0990e+064	8	1.4004e+152
5	2.4055e+087	9	1.4152e+195

表 3 是 AWE 算法求解过程所用的 M 矩阵的条件数随阶数 q 变化的趋势。可以很明显看到，矩阵的条件数恶化的非常快，即使 Matlab 内部采用双精度浮点数来存储数据，5 阶以上精度就已经几乎没有提高的空间了。这一点和我们分析 AWE 缺点的讨论符合，正是 AWE 显式计算导致降阶模型的阶数不可能太高的原因。

PVL

应用 PVL 算法 3 阶、8 阶、12 阶和 13 阶降阶后的幅度频率响应如图 33 所示。

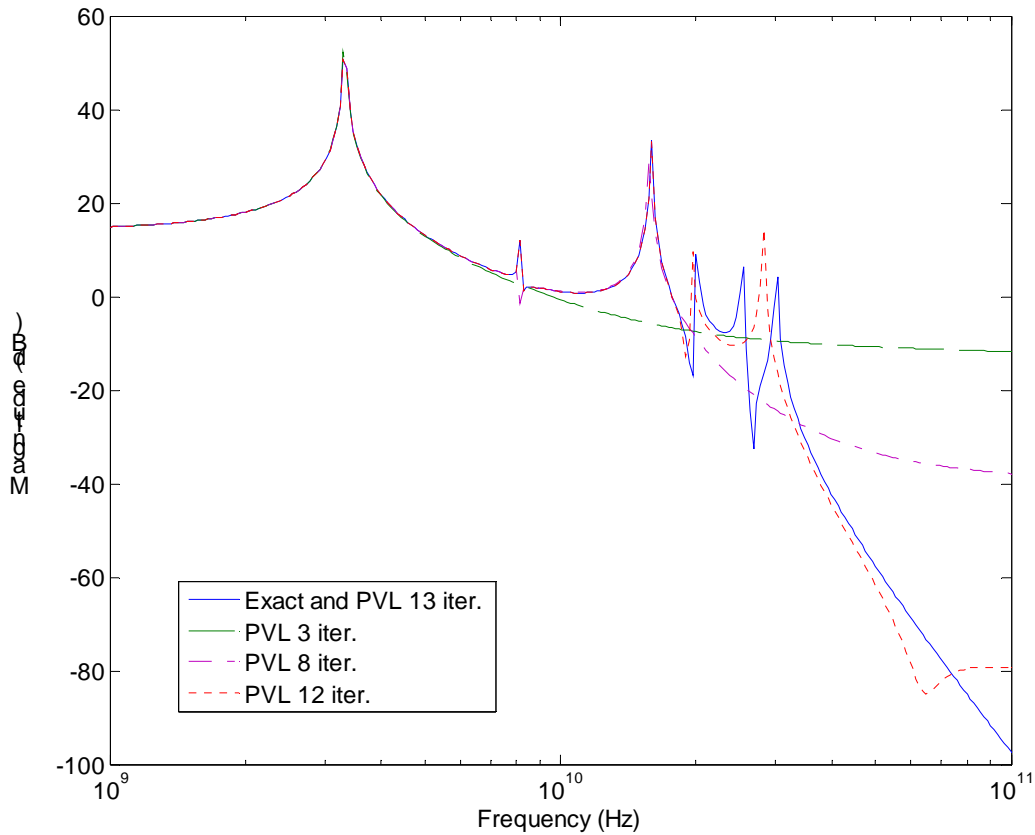


图 33 PVL 对两根电容耦合的 RLCG 互连线的仿真结果

PVL3 阶的时候就已经可以完全逼近 2.5GHz 处的第一个尖峰了。这说明 PVL 的精度和收敛速度是比较快的，主要得益于 PVL 算法中的 Lanczos 三对角化过程，没有像 AWE 那样显式求矩，所以避免了矩阵数值恶化的发生。而且 PVL 随着阶数提高，能够近似的频率范围也越来越大，分别在 8 阶和 12 阶的时候逼近了第三和第四个尖峰，而且当 q 增加到 13 阶之后，已经和未降阶模型的曲线完全一致了，从图中来看两条曲线就是重合的。

PRIMA

接着我们再用 PRIMA 对同样的模型降阶，给出了 4 阶、8 阶和 15 阶的结果，如图 34 所示。

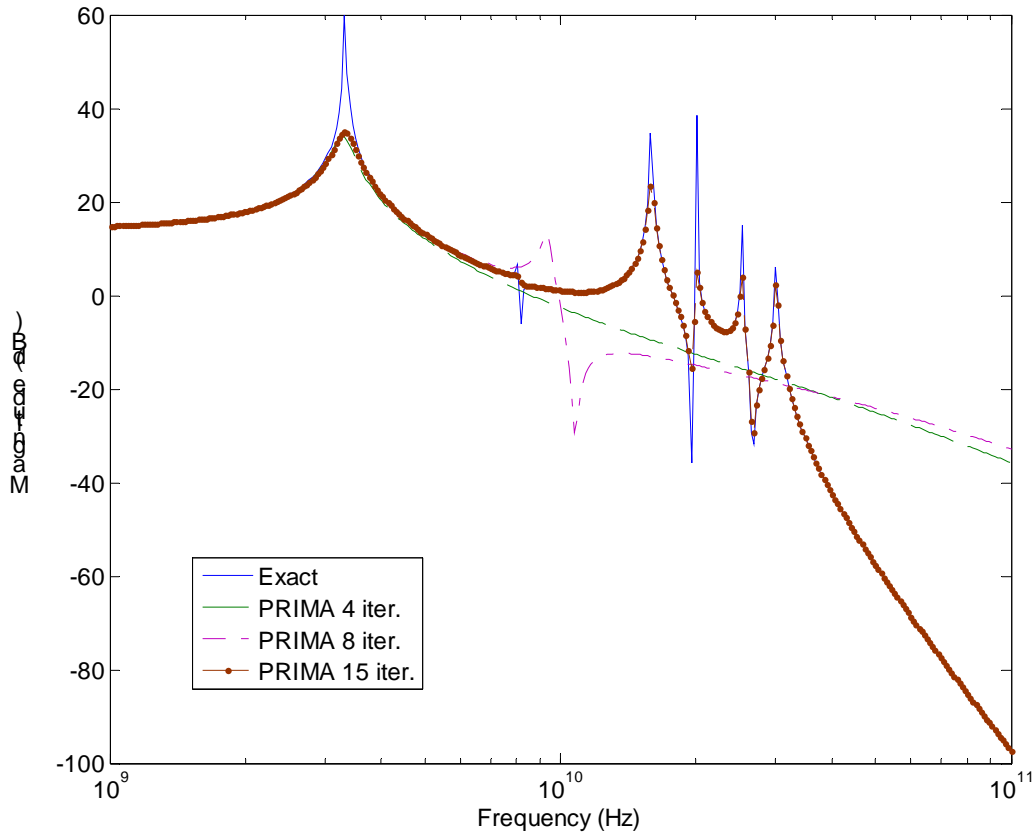


图 34 PRIMA 对两根电容耦合的 RLCG 互连线的仿真结果

PRIMA 的逼近速度比 PVL 要慢一些，到 8 阶的时候第二个尖峰还没有匹配到，而 PVL 已经逼近到第三个尖峰了。另外还可以发现的一点是 PRIMA 所近似的尖峰没有 PVL 那么尖锐，到 15 阶的时候 PRIMA 已经可以基本收敛到曲线的准确值了，但是在尖峰的毛刺部分还是没有做到。PRIMA 收敛速度上不如 PVL 的原因是当阶数为 q 时，PRIMA 只匹配 q 个矩，而 PVL 要匹配 $2q$ 个矩，比 PRIMA 整整多了一倍，因此精度上也提高了。

MOROC

MOROC 对同样的模型降阶的幅频曲线如图 35 所示。

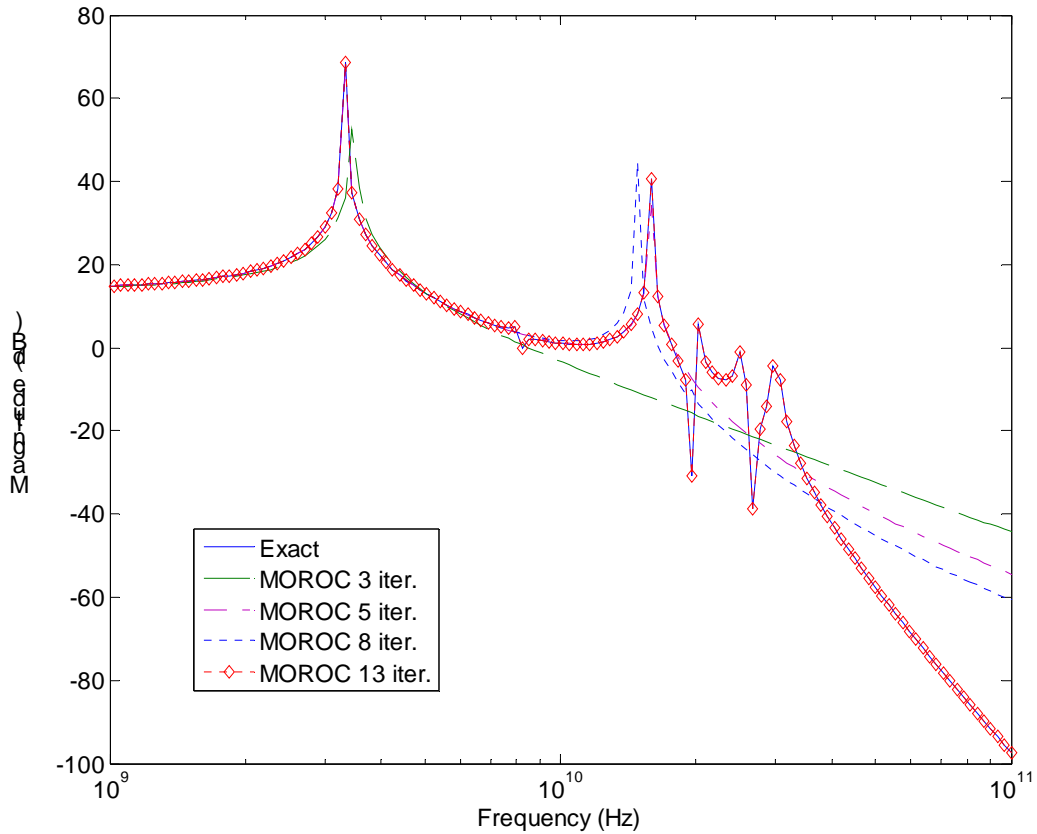


图 35 MOROC 对两根电容耦合的 RLCG 互连线的仿真结果

可以看到，MOROC3 阶基本上可以近似第一个尖峰，但位置有一点点偏。到 5 阶已经能够匹配到 10GHz 以上（第三个尖峰）了。8 阶在高频部分 5 阶压得更低，但是尖峰略有偏移。13 阶的 MOROC 已经可以完全收敛到未降阶的曲线了，可以说其逼近的程度和 PVL 类似。

MOROC 的优点是可以在不计算极点的情况下快速对降阶模型的稳定性进行检验，表 4 是用 Hurwitz 判据检验 10 阶和 13 阶的结果。

表 4 Hurwitz 判据中的 q_i

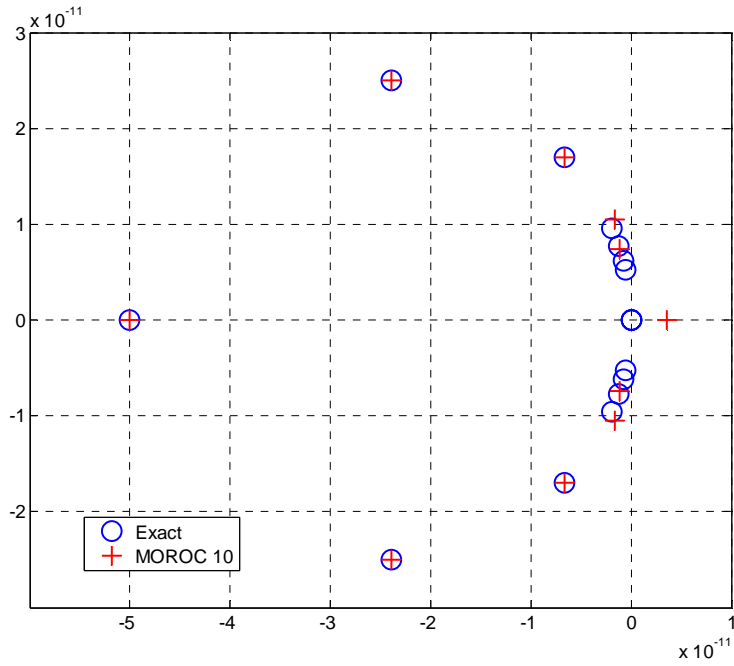
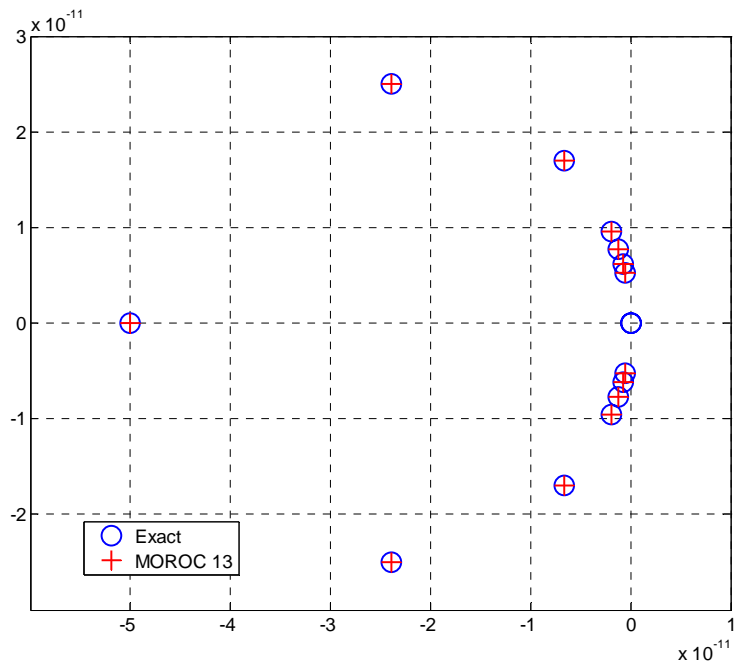
(a) MOROC 10 阶

i	q_i	i	q_i
1	8.8404E+09	6	2.1797E+12
2	2.7479E+10	7	6.0786E+08
3	4.3636E+10	8	-2.4678E+12
4	8.7784E+10	9	2.2057E+11
5	5.1128E+10	10	6.3336E+10

(b) MOROC 13 阶

i	q_i	i	q_i
1	8.3285E+09	8	4.6239E+11
2	2.5732E+10	9	4.3911E+10
3	4.1645E+10	10	2.5202E+12
4	7.1487E+10	11	9.2306E+09
5	6.9871E+10	12	4.0797E+13
6	1.5464E+11	13	6.9265E+08
7	8.6377E+10		

从表中可以看到 10 阶降阶模型有 1 个 q 值小于 0，因此可以推断有 1 个极点有正实部，该模型是不稳定的。而 13 阶的模型所有 13 个 q 值都大于 0，因此模型是稳定的。进一步计算得到降阶后矩阵 A_q 的特征值（倒数即是极点）的分布也验证了上述的推断。

图 36 MOROC 降阶模型 (10 阶) A_q 的特征值分布图 37 MOROC 降阶模型 (13 阶) A_q 的特征值分布

对于 MOROC，虽然不能够像 PRIMA 一样保证无源性，降阶模型无源性的检验还

是比较容易的。如果降阶模型表示的是电路的 Y 参数模型，那么模型无源的充分条件就是在所有频率范围内 $H(j\omega) + H^*(j\omega) > 0$ 。图 38 就是对 MOROC 13 阶模型的检验，横轴是频率，纵轴是 $H(j\omega) + H^*(j\omega)$ 取常用对数后的值。可以看出，在整个频率范围内，该值都是大于 0 的，因此该降阶模型是无源的（因此也是稳定的）。

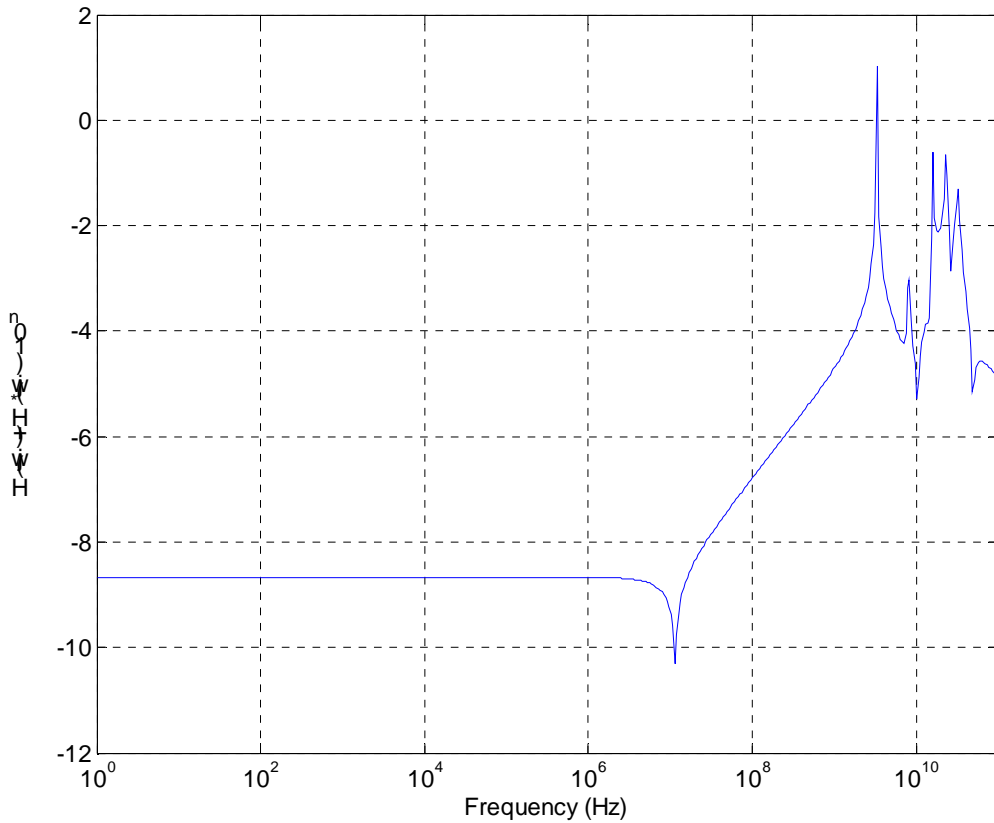


图 38 MOROC 降阶模型的无源性检验

第二根导线输出端的响应

图 39 的曲线反映的是导线 1 对导线 2 串扰的频率特性。

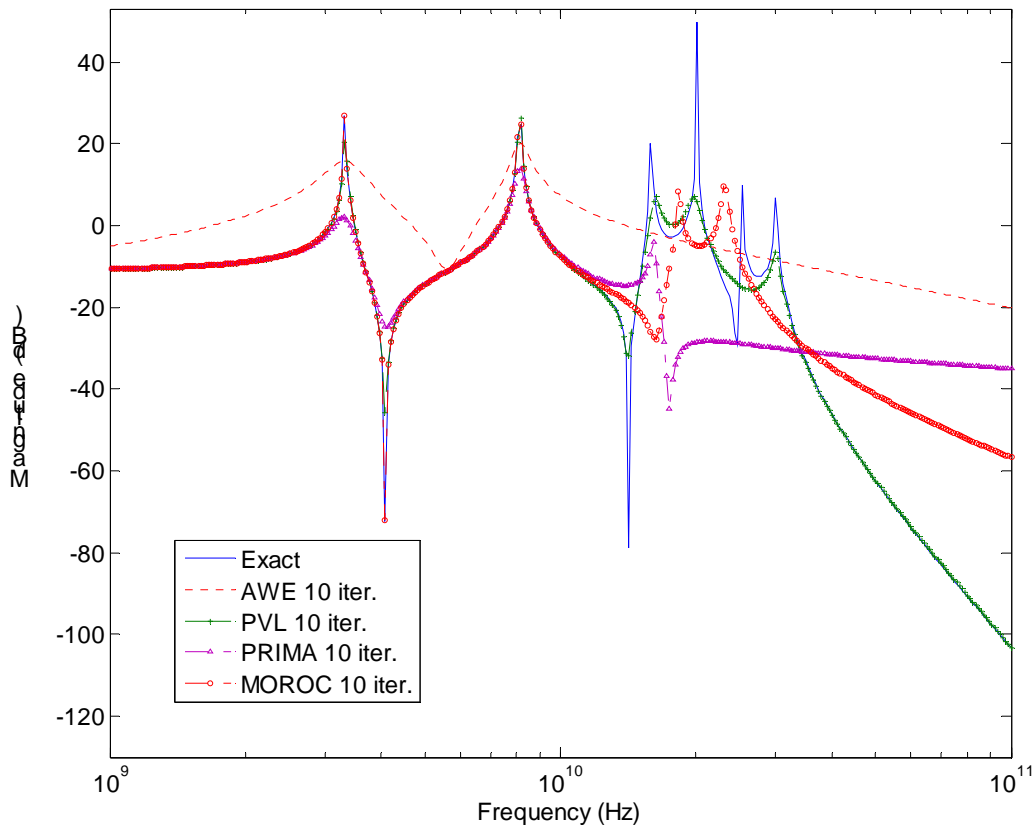


图 39 端口 2 的幅度频率响应

图中列出了各降阶算法 10 阶的结果，不难看出这和前面的分析一致。仅从近似程度来看，有 $PVL > MOROC > PRIMA > AWE$ 。但是各个算法又有自己的特点。AWE 由于是第一个被应用到电路中的模型降阶算法，性能自然逊色一些，但是它所提出的矩匹配的思想一直被后来的算法所使用，而且由于 1 微米以上的工艺并不需要考虑过多的信号完整性问题，也没有电感、互感等的影响，所以 AWE 仍然有用武之地。AWE 的一阶形式——Elmore 延迟也是工程师在草稿纸上估算信号延迟的一个常用的方法。PRIMA 的特点在于能够保证降阶模型的无源性。因为降阶得到的宏模型通常不是单独仿真，而是需要作为一个模块同其他的模型（包括线性和非线性）联合仿真。在这种要求下，仅仅保证降阶模型的稳定性是不够的，两

个稳定的模型组成的系统也有可能是不稳定的。因为我们降阶的对象往往是无源的 RLC 电路，因此希望得到的降阶模型也能够保持无源性，从而保证仿真中不会出现不稳定的现象。PRIMA 最初提出来的时候正是基于这个目标，采用 Arnoldi 算法得到 Krylov 子空间的一组正交基，然后直接对 MNA 矩阵作合同变换以得到降阶模型。数学上能够证明这样得到的模型必然是满足保持无源性的。不过 PRIMA 只能匹配其他算法一半的矩，因此精度会差一些。我们提出的 MOROC 算法是从状态方程的角度入手，针对电路系统的可观测性非常弱的特点，把系统降阶成可观测标准型的形式。用的方法依然是 Krylov 子空间投影的方法，但是得到的模型结构简单，易于检查稳定性，且实验表明其精度介于 PVL 和 PRIMA 之间。

4.1.2 串扰的时域波形

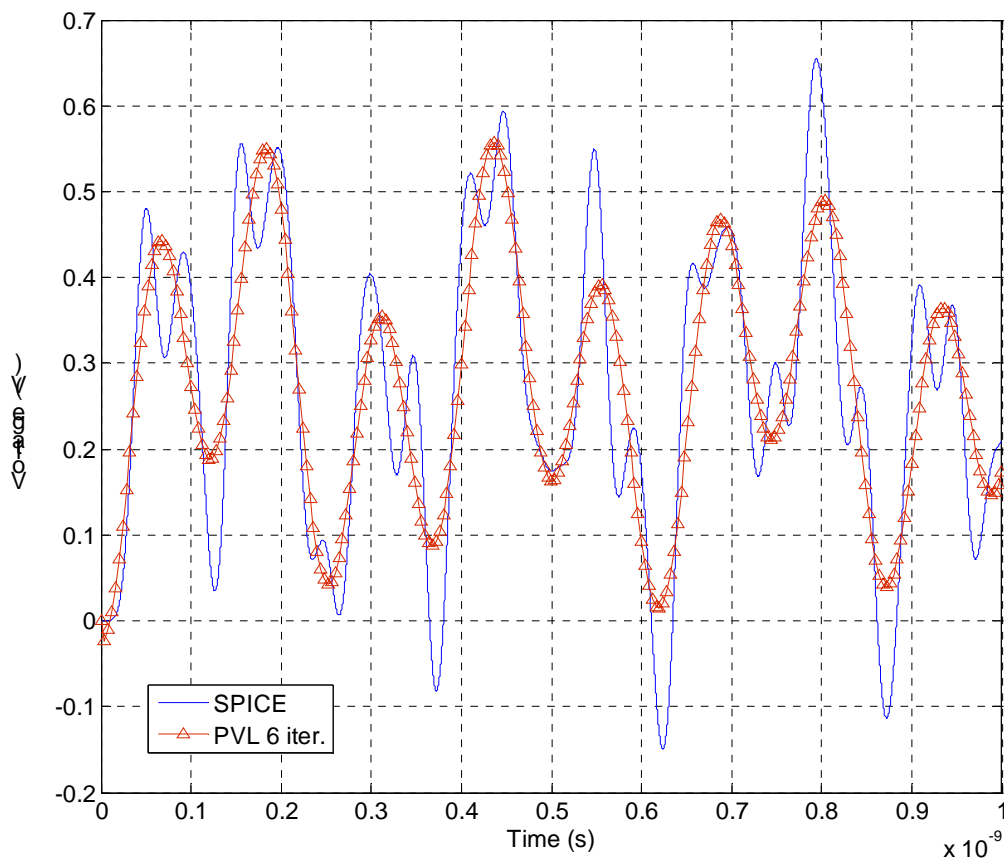


图 40 串扰的时域波形，PVL 6 阶

图 40 是用 6 阶的 PVL 降阶后由 Matlab 得到的时域波形，同时用 SPICE 的波形作为参考，这里认为 SPICE 的波形是最接近于准确值的。由于一个是 Matlab 的结果，一个是 SPICE 的结果，考虑到代码的执行效率差别较大，所以没有作 CPU 时间上的比较，但是模型从 23 阶降到 6 阶，对于时域仿真来说是可以提高不少速度的。这个波形是在一根导线的输入段加从 0V 到 5V 跳变的激励，第二根导线输出端看到的串扰波形。由于有大量寄生参数的存在，且电感效应比较明显，所以产生了高频的振荡。从波形图中可以看到，6 阶的波形对最明显的振荡波形近似的已经非常好了。在 1ns 内这样的振荡一共有 8 个峰，所以可以估算出这是周期大约 1.25ns 的振荡。这正好是我们上一小节中幅频响应的第二个尖峰，大约在 8GHz 处。

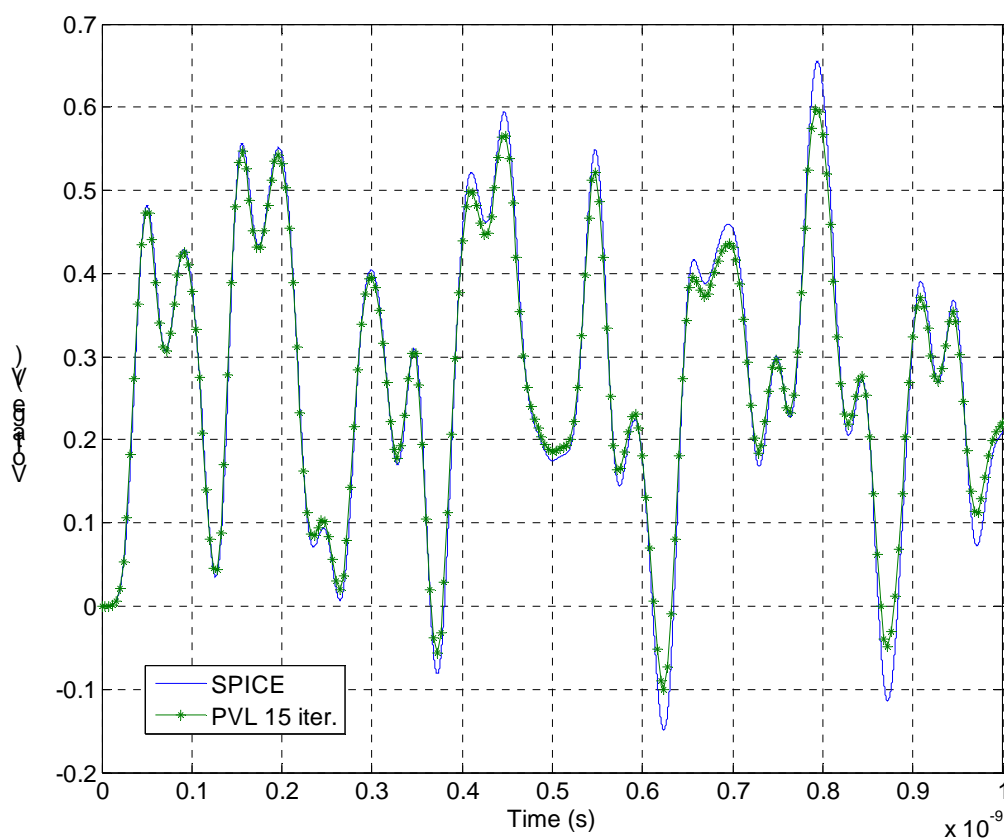


图 41 串扰的时域波形，PVL 15 阶

用 15 阶的模型得出的波形的近似度要更高一些，如图 41 所示。波形的细节

部分都能够吻合的比较好了。

4.2 5 根耦合导线 1

这个例子是 5 根耦合的导线（图 42），每根导线的参数为： $R = 30 \text{ m}\Omega$ ， $C = 0.3 \text{ pF}$ ， $L = 0.3 \text{ nH}$ ，相邻导线之间通过 $C_{couple} = 0.15 \text{ pF}$ 的电容耦合。现在每根导线用 4 段 RLC 模型建模，包括导线之间的耦合电容，最后得到的 MNA 矩阵是 52 阶。现在第一根导线的一端加上一个电平跳变，其他导线都接地。

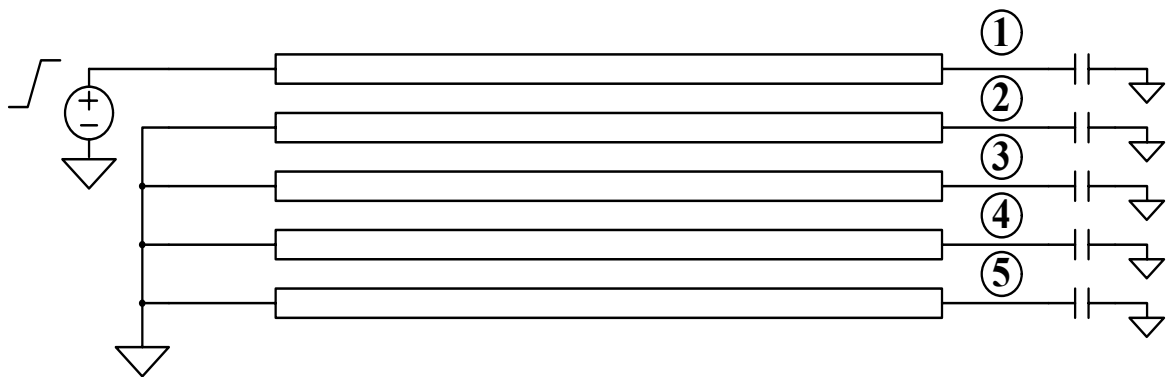


图 42 5 根耦合互连线

为了观察其他 4 根导线受到导线 1 的串扰的影响，通过模型降阶的方法降到 20 阶后用时域仿真得到的波形如图 43 所示。

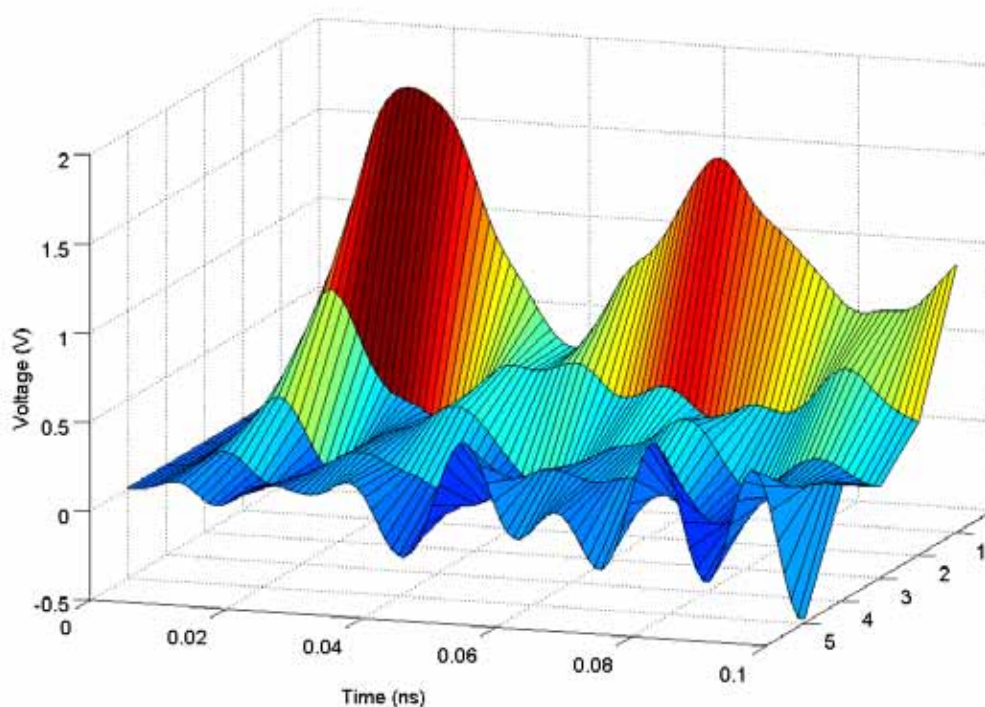


图 43 各端口的串扰波形，PVL 20 阶

最后面是入侵者 (aggressor) 的波形，前面的 4 个波形分别是受害者 (victim) 2、3、4、5 的波形。1 号线加上的是阶跃信号，稳态电平 1V，由于寄生参数的影响，激起了强烈的振荡，串扰从 1 到 5 的传播十分明显，且到 5 号线时的振幅仍然非常大，达到了 0.5V。这里没有画出 SPICE 的波形，因为 20 阶时的波形已经和 SPICE 的结果几乎一致了。

4.3 5 根耦合导线 2

这个实验同样也是一组耦合导线，结构以及耦合方式与前一例相同，电阻、电容的参数不同。现在导线 1 和导线 3 的输入端同时加入阶跃信号，并假设 1 是入侵者，3 是受害者，观察在导线 3 输出端看到的信号完整性受影响的情况。

情形一：导线 1 跳变方向和导线 3 相同

导线 1：0→1，导线 3：0→1

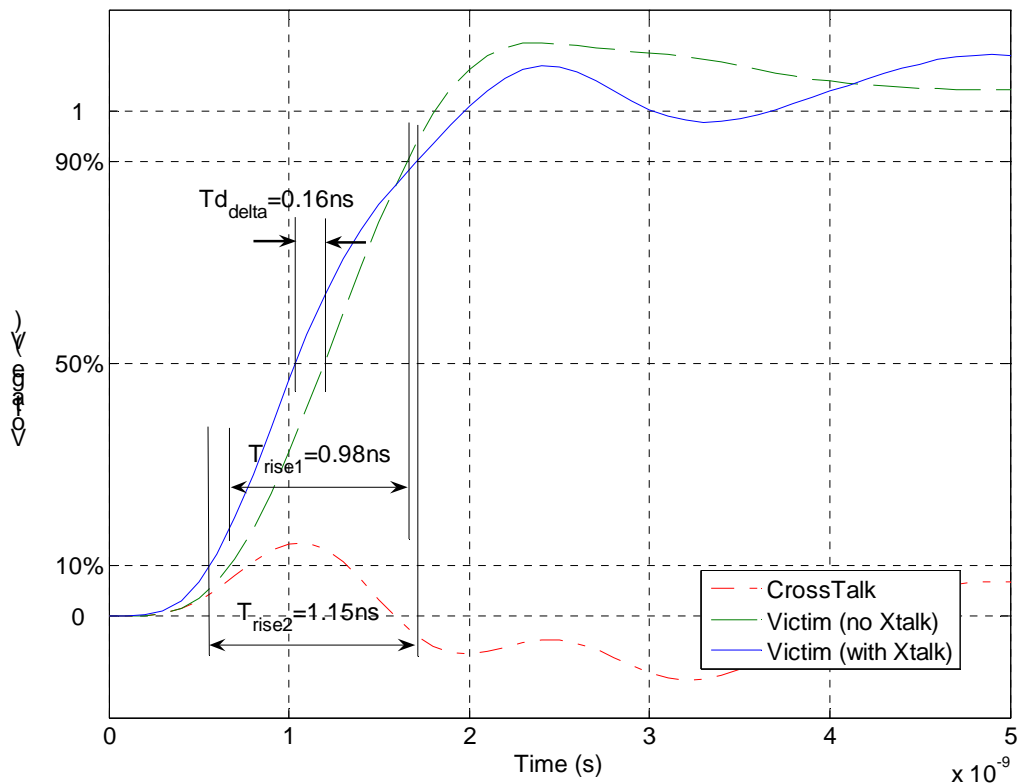


图 44 导线 3 上的瞬态波形（有串扰和无串扰，时延减少）

图 44 给出了用模型降阶算法得到的导线 3 上的波形。图中的 3 条波形分别表示了 1 对 3 的串扰、3 在无串扰时的波形、3 在有串扰时的波形。根据习惯，把 VDD 的 50% 定义为计算时延的阈值，把电压从 VDD 的 10% 变化到 90% 的时间定义为上升时间。则由图中可以看到，在有串扰情况下，导线 3 的传输时延要比没有串扰时少了 0.16ns。而有无串扰时的上升时间分别是 1.15ns 和 0.98ns，两者相差 0.17ns。

情形一：导线 1 跳变方向和导线 3 相反

导线 1：1→0，导线 3：0→1

这时的波形如图 45 所示。

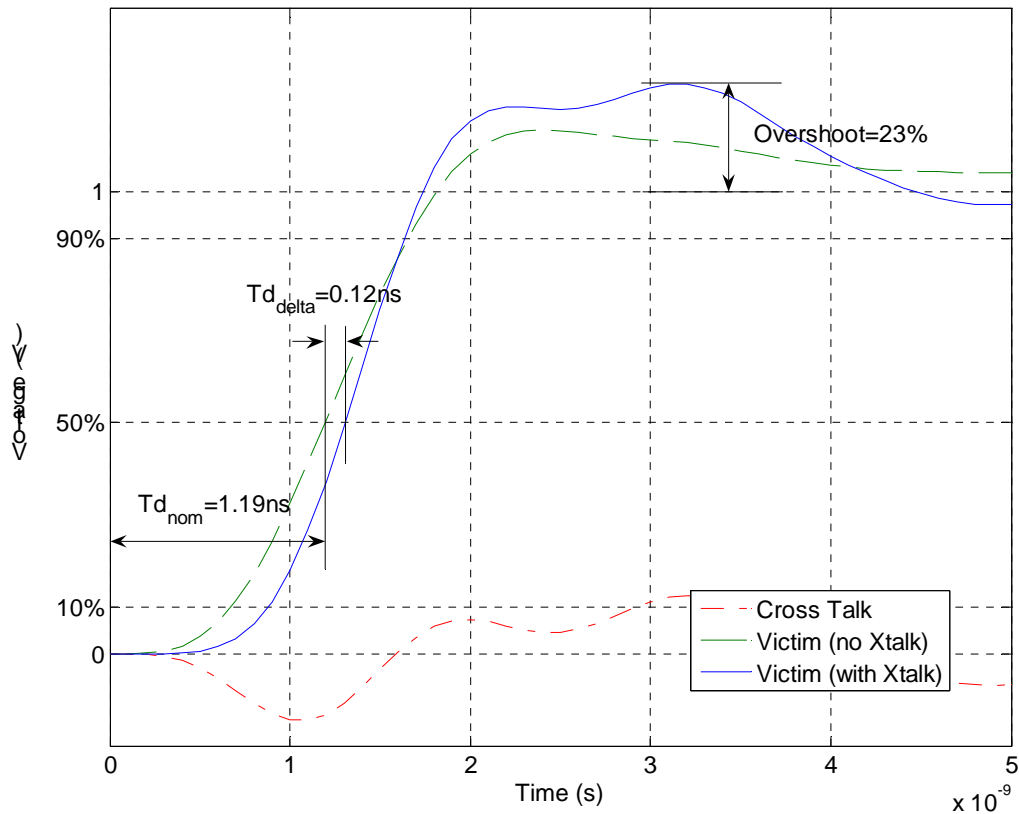


图 45 导线 3 上的瞬态波形（有串扰和无串扰，时延增加）

可以看到，导线 3 自身的时延是 1.19ns，但此时由于受到串扰的影响而增加了 0.12ns。同时阶跃响应的过冲（overshoot）也达到了 23%，相对原来增加了不少。这是仅考虑有一个入侵者的情况，如果附件的几根导线同时都发生跳变，这些串扰也将会叠加，对信号完整性造成更大的损害。

5 结束语

集成电路产业的发展离不开电子设计自动化技术的进步。随着设计水平的提高，越来越多的二级效应已经不能够忽视，信号完整性问题也逐渐引起了人们的重视。但是芯片规模和工作频率的日益增高给信号完整性的分析带来的巨大的困难，用传统的时域仿真方法很难在一个人们可以接受的时间内得到分析的结果。正是为了解决这一问题，模型降阶技术逐渐发展了起来。

模型降阶从概念上来说就是通过数学的方法快速得到原电路的一个等效模型，这个模型从输入输出端口的传输函数上来看非常接近原电路，但是规模却减小了很多，因此在数值仿真的时候能够提高速度、减小内存消耗。对于大规模集成电路的互连线来说，这一方法尤其重要。首先因为互连线通常在寄生参数抽取以后规模巨大，直接计算难以满足要求。其次互连线一般只包含线性无源原件，比如电阻、电容、电感，对于模型降阶算法来说这样问题就简化了许多，并且有不少降阶算法利用了线性电路的一些性质。

本文中对模型降阶的几个重要的算法做了研究，主要阐述了投影降阶的数学原理，分析比较了基于这一思想的各种算法的优缺点。同时提出了基于可观测标准型的模型降阶算法，推导证明了该算法能够匹配原系统矩，且具有模型形式简单和容易检验稳定性的特点。在本文的实验部分，我们得到了几种互连线的模型用各种算法降阶后的模型，并对这些模型分别做了频域和时域的比较。实验结果表明模型降阶算法能够以比原电路小得多的模型达到接近于 SPICE 的精度，这对于大规模集成电路的信号完整性分析具有重大的意义。

模型降级技术目前还在发展，包括考虑工艺参数异动、非线性、符号化的模型降级技术在未来的几年中将会是研究的热点。随着 65 纳米，甚至 45 纳米工艺的逐渐普及，对信号完整性的要求可以说是越来越强烈，因此模型降阶作为其中的关键技术也值得我们更多的关注和研究。

参考文献

- [1] L. T. Pillage and R. A. Rohrer, Asymptotic waveform evaluation for timing analysis, IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, 1990, 9(4), pp. 352-366.
- [2] Curtis L. Ratzlaff and Lawrence T. Pillage, "RICE: Rapid Interconnect Circuit Evaluation Using AWE", IEEE Trans. Computer-Aided Design, vol. 13, No. 6, June 1994, pp. 763-776
- [3] P. Feldman and R. W. Freund, Efficient Linear Circuit Analysis by Pade Approximation Via the Lanczos Process, IEEE Trans. Computer - Aided Design of Integrated Circuits and Systems, 1995, 14(5), pp. 639-649.
- [4] P. Feldman and R. W. Freund, Reduced-order modeling of large linear subcircuits via a block Lanczos algorithm, IEEE/ACM Proc. DAC, June 1995, pp. 474-479.
- [5] L. M. Silveira, M. Kamon, and J. White, Efficient reduced-order modeling of frequency-dependent coupling inductances associated with 3-D interconnect structures, IEEE Trans. Components, Packaging, and Manufacturing Technology, 1995, 19(2), pp. 283-288.
- [6] K. J. Kerns, I. L. Wemple, A. T. Yang, "Stable and Efficient Reduction of Substrate Model Networks using Congruence Transformation", ICCAD, Nov. 1995, pp. 207-214.
- [7] A. Odabasioglu, M. Celik, and L. T. Pillage, PRIMA: Passive reduced-order interconnect macro modeling algorithm, IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, 1998, 17(8), pp. 645-654.
- [8] Freund, R. W., "SPRIM: structure-preserving reduced-order interconnect macromodeling", Computer Aided Design, 2004. ICCAD-2004. IEEE/ACM International Conference on 7-11 Nov. 2004, pp. 80-87.
- [9] Xiao dong Yang, Chung kuan Chen, Walter H. Ku, Robert J. Carragher, Hurwitz stable order modeling for RLC interconnect Trees, IEEE/ACM pro. ICCAD, Nov. 2000, pp. 222-228.
- [10] Bernard sheehan, "TICER: Realizable reduction of extracted RC circuits," in IEEE/ACM pro. ICCAD, Nov. 1999, pp. 200-204.

- [11] C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," J. Res. Nat. Bur. Standards, vol. 45, 1950, pp. 255-282.
- [12] T. J. Schmerbeck, R. A. Richetta, and L. D. Smith, "A 27 MHz mixed analog/digital magnetic recording channel DSP using partial response signalling with maximum likelihood detection," Tech. Dig. IEEE Int. Solid-State Circuit Conf., Feb. 1991, pp. 136-137, 304, and pp. 96, 97, 256
- [13] B. R. Stanisic, R. A. Rutenbar, and L. R. Carley, "mixed-signal noise-decoupling via simultaneous power distribution design and cell customization in rail", Proc. IEEE Custom Integrated Circuits Conf., 1994, pp. 533-536.
- [14] D. L. Boley, "Krylov space methods on state-space control models," Circuits Syst. Signal Processing, vol. 13, no. 6, 1994, pp. 733-758.
- [15] Seok-Yoon Kim, Nanda Gogal and Lawrence T. Pillage, "Time-Domain Macromodels for VLSI Interconnect Analysis," IEEE TRAN. Computer-Aided Design of Integrated Circuits and Systems, Vol. 13, No 10, Oct. 1994, pp 1257-1270.
- [16] C. S. Amin, M. H. Chowdhury, and Y. I. Ismail, "Realizable RLCK circuit crunching," in Proc. DAC, Anaheim, CA, 2003, pp. 226-231.
- [17] Z. Qin and C. Cheng, "Realizable parasitic reduction using generalized $Y-\Delta$ transformation," in Proc. DAC, Anaheim, CA, 2003, pp. 220-225.
- [18] B. N. Sheehan, "Branch merge reduction of RLCM networks," in Proc. ICCAD, San Jose, CA, 2003, pp. 658-664.
- [19] Carlos P. Coelho, Joel Phillips, and L. Miguel Silveira, "A Convex Programming Approach for Generating Guaranteed Passive Approximations to Tabulated Frequency-Data," IEEE Trans. Computer-Aided Design, vol. 23, no. 2, Feb. 2004, pp. 293-301.
- [20] J. Morsey and A. C. Cangellaris, "PRIME: passive realization of interconnect models from measured data," in Proceedings of the IEEE 10th Topical Meeting on Electrical Performance of Electronic Packaging, Oct. 2001, pp. 47-50.
- [21] Elfadel I M, Ling D D. "A block rational Arnoldi algorithm for multipoint passive model order reduction of multiport RLC networks," in IEEE/ACM International

- Conference on Computer Aided Design, San Jose, California, 1997, pp. 66–71.
- [22] K. Henrik, A. Olsson, “Model Order Reduction with Rational Krylov Methods,” Doctoral thesis in numerical analysis, Stockholm, 2005.
- [23] Y. Su, J. Wang, X. Zeng, Z. Bai, C. Chiang and D. Zhou, SAPOR: Second-order Arnoldi method for passive order reduction of RCS Circuits, Proceedings of IEEE/ACM International Conference on Computer-Aided Design (ICCAD 2004), pp.74-79.
- [24] Z. Qi,, H. Yu, P. Liu, S. X.-D. Tan, L. He, “ Wideband passive multi-port model order reduction and realization of RLCM circuits ” , IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, (TCAD) vol 25, No. 8, Aug 2006, pp. 1496-1509.
- [25] Raghavan V,Bracken J E,Rohrer R A. AWESpice:a general tool for the accurate and efficient simulation of interconnect problems. In:Proceedings 29th ACM/IEEE Design Automation Conference.ACM, New York, 1992, pp 87-92.
- [26] 王豪行, 电路计算机辅助设计, 上海: 上海交通大学出版社, 1995, 2-5
- [27] J. M. Rabaey 等著, 周润德等译, 数字集成电路——电路、系统与amp;设计, 第二版[M], 北京: 电子工业出版社, 2004.
- [28] 胡光锐, 信号与系统, 上海: 上海交通大学出版社, 1995.
- [29] 尤昌德, 线性系统理论基础, 北京: 电子工业出版社, 1985, 154-155
- [30] Celtic NDC User Guide, v6.1, Cadence, 2006, 61-68.

附录

无

致谢

时光荏苒，岁月如梭，两年多的研究生学习生涯即将结束。在此我要真心感谢很多曾经给予我关怀和帮助的人。首先要感谢我的导师施国勇教授，没有他一年来的悉心指导，就没有我的这篇论文。施老师敏捷的思路，广博的学识和严谨的治学态度给我留下了深刻的印象。我能够深深感受到导师对教学工作的热情，对科学研究的热爱，以及那对学生认真负责的精神。施老师在授课时循循善诱，不厌其烦地给我们讲解知识难点，且他那带有西方思维的教学方式，常常能为我们这些应试教育制度下的学生带来一些新鲜的空气。

其次要感谢我们的大家庭——上海交通大学微电子学院的老师和同学们。感谢郭炜老师、汪宁老师、谢凯年老师、韩泽耀老师、赵峰老师、梁阿磊老师和程秀兰老师，正是你们的授课把我领进了微电子行业的大门，让我对这个学科充满了兴趣。尤其要感谢郭炜老师，在各方面都给过我帮助，您曾经对我的教导我一直不会忘记。还要感谢实验室的师兄师姐们，彭莉、王兵、罗伟毅、李宇飞、余宙、宋飞，我的进步离不开你们，当我在迷茫的时候，你们这些前辈给了我无私的帮助。特别感谢彭莉大姐，你和蔼可亲，把我们这群学生当成自己的弟弟妹妹一样，两年来在学习和生活上都给了我许多关怀和帮助，在实验室里我们都称你为大姐，今后你也一直是我们的大姐。

还要感谢我的同窗好友，彭瑞华、孙传名、冯鑫、张耀辉、洪杰、蒋剑飞、戴剑臻、陈志凯、郭向东、马潇、赵大猷、谢俊、林明亮、陈微微、严青、郑秋华等，大家是否都很怀念曾经在实验室的那段时光？虽然每天在闵行和徐汇之间早出晚归，来回奔波的日子很辛苦，但是大家结伴而行，苦中作乐，同舟共济，蓬勃向上，在学习中发现快乐，在生活中寻找激情，一起度过了这段一生中永远都不会忘记的日子。

最后我要衷心感谢我的父母，没有你们无微不至的关怀和支持，我就不可能成才。不管遇到什么困难和挫折，你们始终是我的坚强后盾，是我的避风港湾，你们的爱是天下最伟大最无私的！

邓立群

2006年11月26日于上海家中

攻读学位期间发表的学术论文

1. 邓立群, 施国勇, “一种基于可观测标准型的模型降阶算法”, 《计算机仿真》, 已录用