

引用格式:高彬,叶佐昌,王燕. 基于反向传播的电路优化与模型参数提取方法[J]. 微电子学与计算机, 2019, 36(1): 79-84.
[GAO B, YE Z C, WANG Y. Model Parameter Extraction with Back Propagation Algorithms [J]. Microelectronics & Computer, 2019, 36(1): 79-84.]

基于反向传播的电路优化与模型参数提取方法

高 彬, 叶佐昌, 王 燕

(清华大学 微电子学研究所, 北京 100084)

摘 要: 传统的电路优化和模型参数提取方法假定仿真器为黑盒子, 通过修改参数或者优化变量得到仿真结果, 并用迭代的方式进行优化或模型参数提取. 由于仿真器并未提供梯度信息, 因此只能采用不依赖于梯度的优化方法, 限制了优化和参数提取的性能. 本文介绍了一种开源的电路仿真工具, 通过在给出仿真结果的同时计算出指定函数对指定参数的梯度, 为基于梯度的优化算法在电路参数提取、电路优化等领域的应用提供了便利. 使用该工具实现的梯度下降算法在电感模型和变压器模型参数提取问题上的实验表现显著优于模拟退火等非梯度方法和同一算法不使用其梯度的实现的版本. 实验结果证明, 本文提出的方法在电路优化和模型参数提取问题上的性能有很好的改进作用.

关键词: 仿真器; 电路器件仿真; 电路优化; 电路参数提取; 梯度算法; 优化算法;

中图分类号: TN401; TP301.6

文献标识码: A

文章编号: 1000—7180(2019)01—0079—06

Model Parameter Extraction with Back Propagation Algorithms

GAO Bin, YE Zuo-chang, WANG Yan

(Institute of Microelectronics, Tsinghua University, Beijing 100084, China)

Abstract: This work aimed at pushing forward the application on electronic fields of optimization algorithms which hire gradients. The authors developed an open source electronic simulating tool that gives out gradients of designed functions with respect to customized parameters of interest simultaneously when computing the desired state vectors of a circuit net list. These gradients power most gradient based optimization algorithms efficiently and relieve them from fetching the gradients expensively by secant methods. These optimization algorithms perform much better with the help of it, and also beat some non-gradient algorithms namely simulate anneal.

Key words: simulator; circuit-simulator; circuit-optimization; circuit-parameter-extraction; gradient-based-algorithms; optimization-algorithms

1 引言

在现代集成电路中, 有源器件和无源元件的建模是工艺开发和电路设计中不可缺少的步骤^[1]. 其中模型的参数提取成为模型建模中的关键. 过去的 20 年来, 人们针对不同的器件提出了各种模型的参数提取方法^[2-6].

在传统的模型参数提取中, 通常使用电路仿真工具, 结合优化算法来进行. 优化算法调整器件模型

参数, 送到仿真工具, 通过仿真得到人们关心的元器件的端口特性参数, 例如对于有源器件模型的 I-V 特性和 C-V 特性曲线, 以及无源元件的 S 参数或者 Y 参数曲线等. 然后计算这些端口特性曲线与实验测量的端口特性曲线之间的误差, 并通过调整模型参数来减小该误差.

在此类方法中, 电路仿真工具, 如 SPICE 等, 作为黑盒子, 只接受器件或者元件参数作为输入, 而输出是端口的特性曲线. 因此, 优化算法只能局限于不

需要依赖梯度的优化算法,例如粒子群优化^[7],差分进化算法^[8]等.由于缺乏梯度方向的指导,这类方法通常需要大量调用仿真器得到不同参数下的端口参数,导致效率较低.另外一类方法,可以通过差分的方式来得到误差相对器件参数的梯度,这种方法对于每个器件参数,都需要做一遍完整的电路仿真,计算效率仍然较低^[9].

本文借鉴近年来深度学习中反向传播^[10]的思想,研究构建电路仿真的计算图模型,并采用全反向传播的方式来进行误差的反向传播.并且结合现代的随机梯度下降方法的变种^[11-13],进行误差的优化.

理论上,每计算一次梯度,该方法只需进行一次电路仿真,再加上一次反向梯度传播.每次迭代的计算量大致相当于2次电路仿真所需的时间,远低于采用群优化的方法每次迭代所需时间.同时,在得到梯度之后,采用随机梯度下降方法及其变种,与群优化算法和差分进化算法相比,可以大大提高优化收敛所需的迭代次数.

这两方面优势的结合,使得采用误差反向传播方式进行模型参数提取,相对于传统的方法有着极大的优势.实验结果证实了该方法的有效性.

2 背景

2.1 电路仿真原理

电路器件仿真问题是求解一个电路器件的电学特性的问题.所以电学仿真就是将一个电路器件用理想器件描述出来后,求解这组理想器件组成的网络的电学特征的过程.

将一个电路表示成一张有向图 $G(V, E)$, 其中电流满足

$$I_{u,v} = -I_{v,u} \quad (\forall (u, v) \in E) \quad (1)$$

电压满足

$$U_{u,v} = -U_{v,u}, \quad \forall (u, v) \in E \quad (2)$$

从而用来建模电路器件的每个理想器件的电学特性都可以表示成此电路网络中的物理量的微分方程:

$$f(I, \dot{I}, U, \dot{U}) = 0 \quad (3)$$

一个电路网络中的结点回路电势和结点电流满足基尔霍夫电流定律:

$$\sum_{(u,v) \in E} I_{u,v} = 0, \quad \forall u \in V \quad (4)$$

选定一个电势为0的结点,再通过基尔霍夫定律可以将整个电路网络中的所有器件的电学特性联系起来得到一个微分方程组:

$$\frac{d}{dt}[q(x(t))] + j(x(t)) = s(t) \quad (5)$$

式中, $x(t) \in \mathbb{R}^N$ 是在 t 时刻电路结点的状态向量; N 是网络中改进结点的数目; j 是电流; q 是电荷密度; s 是源激励.

这是一组偏微分方程,其中的未知量就是对应电路的改进结点上的电学变量.

解出了这个偏微分方程组,就得到了电路在某一时刻的全部状态信息,待求的所有关键输出变量都是这些状态的函数.从而用这些状态信息可以计算出所有关键输出变量的值,即电路仿真结果.

将这些方程整理成如下标准形式:

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

在某一静态工作点附近进行线性化,求解其小信号模型,可以把这个时域上的方程组通过拉普拉斯变换转换到频域上,改写成这样的等价问题:

$$\begin{aligned} sEX &= AX + Bu \\ Y &= CX + Du \end{aligned} \quad (7)$$

对这个方程组施以牛顿-拉普森迭代法通常可以得到它的一个数值解.而在电路参数提取和电路优化这样的实际问题中,也不难选定一个有物理含义的有意义的初值,从这样的初值一般会收敛到一个有意义的正确解.牛顿-拉普森迭代是一种二阶收敛速率的算法.

2.2 电路参数提取和优化

在电路器件建模的过程中,是电路参数提取是一个重要的问题.电路网结点在某一时刻的状态向量 x , 源激励 s , 电荷密度 q 和电流 j 都由电路参数 p 来确定.

$$\frac{d}{dt}[q(x(t, p), p)] + j(x(t, p), p) = s(t, p) \quad (8)$$

式中, $p \in \mathbb{R}^P$ 是电路参数向量, P 是电路参数的数目.

集成电路器件的等效模型至少包括两部分内容:集成电路器件等效模型的拓扑结构和提取该模型的电学物理量参数 p .

集成电路器件等效模型的拓扑结构来自对器件本身构造的物理特性的抽象,具有坚实的物理意义和现实基础.附加在其上的电路参数则是这个模型准确描述该器件的关键.

电路参数提取要解决的是:在等效电路模型中找到一组参数,使得该模型的输入输出特性与真实

测量结果尽可能相近. 定义损失函数 loss :

$$\text{loss}(p) = \sum_k ||Y_k(p) - Y_k^*|| \tag{9}$$

这是一个优化问题:

$$p^* = \underset{p}{\operatorname{argmin}} \sum_k ||Y_k(p) - Y_k^*|| \tag{10}$$

式中, Y 是系统输出; Y^* 是器件电学特性的实际观测值; k 是激励的频率; p^* 是等效电路中最优的电学参数.

3 方法

3.1 梯度推导

在每轮迭代中, 保存并累加电路改进结点的电学变量的迭代增量 u 对电路关键参数变量 p 的梯度, 得到改进结点的状态向量 x 对电路关键参数 p 的雅可比矩阵:

$$\frac{dx}{dp} = \sum_{n=1}^{n_{\text{loop}}} \frac{\partial u_n}{\partial p} \tag{11}$$

式中, n_{loop} 表示迭代的步数.

迭代收敛之后, 只要构造关键输出变量 o 对优化目标值 o^* 的一个解析的损失函数, 比如均方误差 (MSE), 便可以推导出损失函数 loss 对关键输出变量 o 的梯度:

$$\nabla_o \text{loss}(o; o^*) \tag{12}$$

式中, ∇ 表示求梯度运算.

接着计算出关键输出变量的损失函数 loss 对电路改进结点的状态向量 x 的梯度:

$$\nabla_x \text{loss} = \nabla_o \text{loss}^T J_o(x) \tag{13}$$

式中, $J_o(x)$ 表示求雅可比矩阵.

从而用这个累积的梯度算出了待求的关键输出变量的损失函数 loss 对关键电路参数变量 p 的梯度:

$$\nabla_p \text{loss} = \nabla_o \text{loss}^T J_o(x) \sum_{i=1}^{n_{\text{iter}}} \frac{du_i}{dp} \tag{14}$$

式中, n_{iter} 是牛顿 - 拉普森迭代的步数.

3.2 电路优化与参数提取方法

在无源器件的建模问题中, Y - 参数是一个常用的是器件特性指标, 反映了该器件的一些根本特征.

如图 1 所示, Y - 参数 $Y_k(p)$ 是一个矩阵, 其中元素 $y_{i,j}$ 表示在第 j 端口施加 k 频率的单位幅值的激励时, 第 i 端口的电流响应.



图 1 Y- 参数

无源器件建模问题, 可以选择 Y - 参数作为观测

量, 通过对 Y - 参数进行拟合, 计算出该器件的等效模型.

有了仿真器提供的梯度, 这个问题变得更加容易求解. 以下本文将举电感和变压器的参数提取问题为例, 用梯度下降法与粒子群优化作对比, 说明一个提供梯度推导的仿真器给参数提取问题的求解带来的好处.

4 实验结果

4.1 电感模型参数提取问题

集成电路中的电感与理想电感之间有一些差别. 这是由片上电感中的电流趋肤效应以及对硅片衬底的电容效应等一系列寄生效应引起的. 在许多应用场合, 仿真中这些效应必须加以充分考虑.

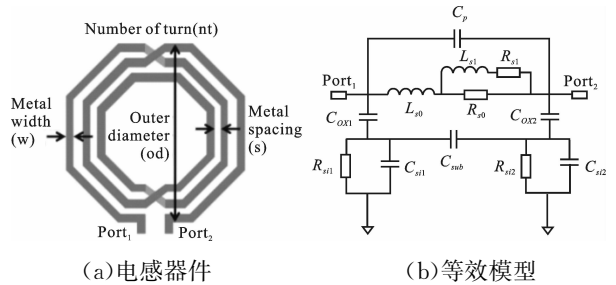


图 2 电感模型

如图 2(a) 所示的这种电感器件, 人们构造了图 2(b) 这样的等效模型来刻画它的电学特性.

通过对这个器件进行测量, 可以画出它的输入输出特性曲线. 这条曲线概括了这个器件的实际作用, 也是本问题的优化目标.

图 2(b) 中的各个理想元件的参数就是人们需要求解的对象.

这个实际问题中, 这些参数大约的数量级范围是可以相关的物理知识粗略地预估的. 不难给出一个良好的初值, 从这样的一些初值出发, 许多优化算法常常可以收敛出一个很好的解. 多数电路参数提取类的问题都是如此. 同时, 初值的选取亦不是本文所要讨论的重点, 下文假定已经知道了所有参数大致的数量级. 初值将在邻近的一个数量级内随机选取. 将比较一种不需要依赖梯度的优化算法 (粒子群优化算法) 与一种需要依赖梯度的优化算法 (Adam 梯度下降算法) 在我们的仿真器中使用仿真器提供的梯度和不使用仿真器提供的梯度时的性能表现.

梯度下降法是一种通过不断将误差对优化参数的梯度的一定倍数加到优化参数上来减小误差, 从而求解优化问题的方法.

在此实验中取学习率 η 初值为 2.0, 学习率的下降倍率 $\gamma = 0.99$:

$$p \leftarrow p - \eta \cdot \nabla \text{loss}(p) \tag{15}$$

$$\eta \leftarrow \gamma \cdot \eta \tag{16}$$

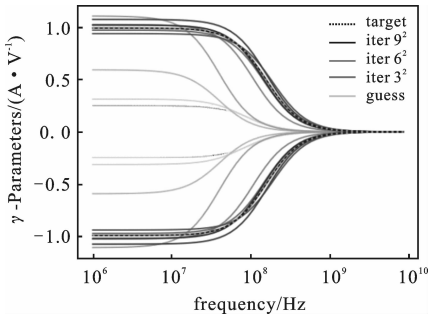
粒子群优化是从一些随机初始解出发, 不断在它的周围随机寻找一个更优的解的优化方法.

在对照实验中取每轮的搜索范围为上一轮迭代值的 10%, 学习率的下降倍率也取 0.99.

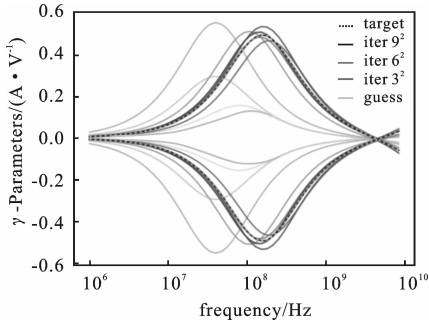
如果梯度未知, 梯度下降算法退化为差分割线法. 这是梯度下降时不知道梯度时的做法. 通过对解的每一维做一点小的扰动, 观察损失函数的变化来求出割线斜率, 从而估计出梯度来进行迭代:

$$\frac{\partial \text{loss}}{\partial p} \approx \frac{\Delta \text{loss}}{\Delta p} \tag{17}$$

如图 3 所示, 梯度下降算法能够正确处理这个问题. 增益曲线在 100 次迭代之内收敛到了目标值.



(a) 实部



(b) 虚部

图 3 电感模型 Y-参数的一条梯度下降收敛轨迹

使用同一组初值, 将此方法与带模拟退火的粒子群优化算法 (SA-PSO) 进行对比. 实验结果如图 4 所示. 不难发现通过使用我们的仿真工具, 梯度下降算法在收敛速度上较粒子群优化算法有明显优势.

然而在算法的实际执行时间方面, 如图 5 所示, 没有仿真器提供梯度推导的情况下, 采用割线法近似地寻找梯度的值, 需要多次仿真, 消耗了大量的计

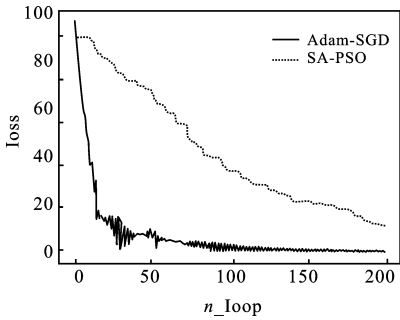


图 4 梯度下降算法与粒子群优化算法收敛速度比较

算, 因此表现并不能明显优于模拟退火算法. 使用了仿真器提供的梯度后, 省略了割线法求解梯度的过程, 大大节约了计算量, 速度提升了数倍.

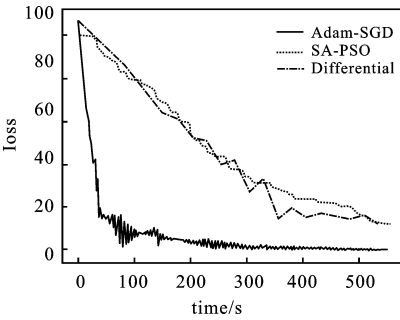


图 5 3 种算法 Y 参数收敛时间曲线

表 1 给出的是用于绘制上图的部分实验数据.

表 1 电感模型不同算法的误差下降及时间开销

迭代步数	粒子群优化	割线法	TF-Spice 梯度下降
1	99.2 (26s)	94.4(55s)	100.7(6s)
9	99.1(66s)	39.9(247s)	50.9(27s)
36	79.2(134s)	5.2(870s)	10.2(95s)
81	47.7(258s)	5.0(1 940s)	1.5(211s)

* 注: 实验环境为 1.6GHz intel core i5 / 4GB 1600MHz ddr3 RAM / Mac OS X Sierra

4.2 变压器模型参数提取问题

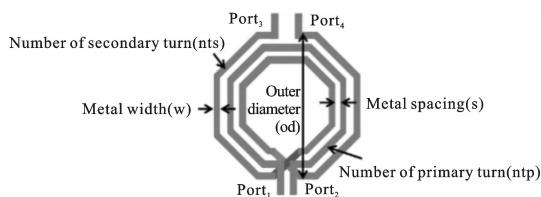
根据上文的推导, 在越高维的问题上, 我们的仿真工具提供的梯度越能改善算法的性能. 如下实验说明了这个问题的.

如图 6(a) 所示, 变压器是一个四端口的网络, 它的 Y 参数是一个 4×4 的矩阵. 可以使用图 7(b) 的网络来描述它的电学特性.

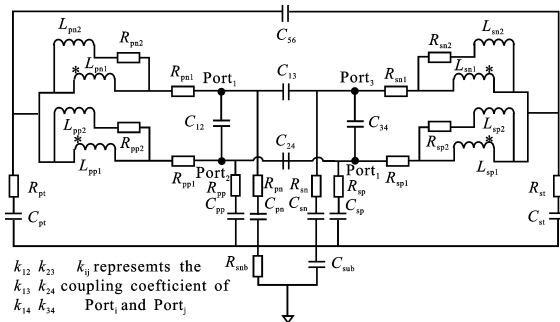
此处使用与上一小节相同的三种算法, 在经验认为合理的范围内随机选定了相同的几组初值.

对这个问题上做与上一小节相同的实验, 得到了如图 7 所示, 收敛轨迹.

如图 8 所示, 在迭代次数意义下的算法收敛速

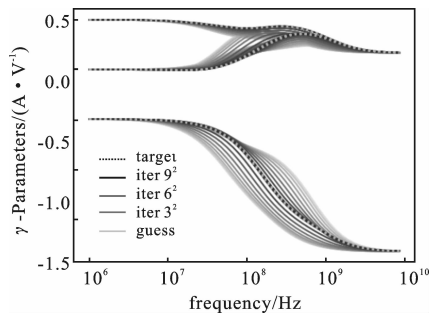


(a) 四端口变压器的网络

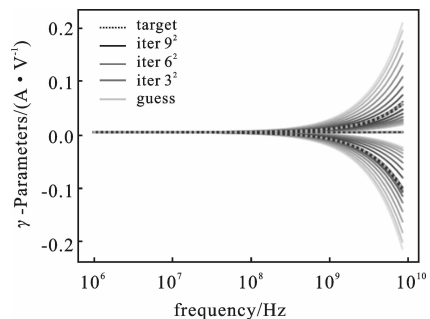


(b)电学特性

图 6 四端口变压器的一种模型



(a) 实部



(b) 虚部

图 7 变压器模型 Y-参数的一条梯度下降收敛轨迹

度方面,梯度下降依然优于带模拟退火的粒子群优化.在运行时间意义下,得到的实验结果非常支持理论的结果,因为本问题的维数较高,使用割线法寻找梯度,运行时间性能没有胜过模拟退火.

如图 9 所示,使用了我们的仿真工具提供的梯度,削减掉了数值求解梯度的时间开销,打破了效率的瓶颈,运行速度大幅提升.

表 2 给出了更详细的一些数据.

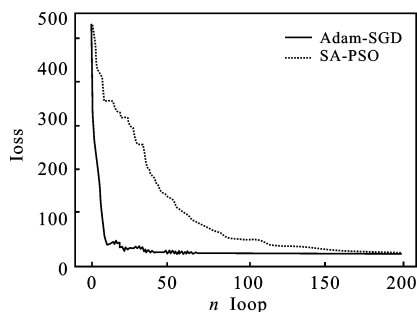


图 8 梯度下降算法与粒子群优化算法收敛速度比较

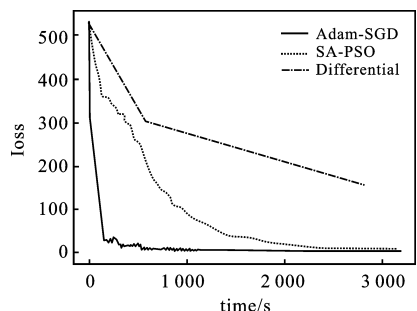


图 9 3 种算法 Y 参数收敛时间曲线

表 2 变压器模型不同算法的误差下降及时间开销

迭代步数	粒子群优化	割线法	TF-Spice 梯度下降
1	523.8(1ks)	331.5(34ks)	301.8(1ks)
9	353.7(14s)	24.0(487ks)	37.94(14ks)
36	198.5(23ks)	9.33(816ks)	4.03(23ks)
81	49.6(45ks)	3.15(1566ks)	3.59(46ks)

* 注:实验环境为 1.6 GHz intel core i5 / 4GB 1600 MHz ddr3 RAM / Mac OS X Sierra

5 结束语

我们开发了一种新的电路仿真工具,主张在对关键指标进行仿真时提供对关键参数的梯度,以方便后续的电路优化工作.

在此仿真工具的帮助下,基于梯度的方法在电路参数提取、电路优化领域的应用变得前所未有的简单容易和高效.这种仿真工具架起了一座基于梯度的算法和电路仿真、电路参数提取问题的桥梁.

我们可以基于此仿真工具将电路优化领域的许多关键问题做得更好.

参考文献:

- [1] Yao J, Ye Z, Wang Y. An automatic parameter extraction and scalable modeling method for transformers in RF circuit[C]// Proc. IEEE RFIC Symp. , Seattle, WA, USA, 2013: 57-60.
- [2] Gao W, Jiao C, Liu T, et al. Scalable compact circuit

- model for differential spiral transformers in CMOS RFICs[J]. IEEE Trans. Electron Devices, 2006, 53(9): 2187-2194.
- [3] Leite B, Kerherve E, Begueret J B, et al. An analytical broadband model for Millimeter-Wave transformers in silicon technologies[J]. IEEE Trans. Electron Devices, 2012, 59(3): 582-589.
- [4] Guo J, Tan T. A broadband and scalable model for on-chip inductors incorporating substrate and conductor loss effects[J]. IEEE Trans. Electron Devices, 2006, 53(3): 413-421.
- [5] Cao Y. Frequency-independent equivalent-circuit model for onchip spiral inductors[J]. IEEE J. Solid-State Circuits, 2003, 38(3): 419-426.
- [6] Yue C, Wong S. Physical modeling of spiral inductors on silicon[J]. IEEE Trans. Electron Devices, 2000, 47(3): 560-568.
- [7] Mandal S, Sural S, Patra A. ANN and PSO based synthesis of on-chip spiral inductors for RF ICs[J]. IEEE Trans. ComputerAided Design Integr. Circuits Syst., 2008, 27(1): 188-192.
- [8] Liu B, Lu J, Wang Y, et al. An effective parameter extraction method based on memetic differential evolution algorithm[J]. Microelectron. J., 2008, 39(12): 1761-1769.
- [9] Price K, Storn R M, Lampinen J A. Differential evolution. a practical approach to global optimization [M]. Berlin, Germany: Springer, 2005.
- [10] Rumelhart D E, Hinton G E, Williams R J. Learn-ing representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533-536.
- [11] Amari Shun Ichi. Natural gradient works efficiently in learning [J]. Neural computation, 1998, 10(2): 251-276.
- [12] Pascanu Razvan, Bengio Yoshua. Revisiting natural gradient for deep networks[EB/OL]. <https://arxiv.org/abs/1301.3584v4>, arXiv preprint arXiv: 1301.3584, 2013.
- [13] Diederik P Kingma, Jimmy Ba. Adam: a method for stochastic optimization[EB/OL]. <https://arxiv.org/abs/1412.6980v8>. arXiv:1412.6980 [cs. LG].

作者简介:

高彬男, (1988-), 硕士研究生. 研究方向为电路优化与模型参数提取方法.

E-mail: gaob13@mails. tsinghua. edu. cn.

叶佐昌男, (1980-), 博士, 副教授. 研究方向为集成电路设计验证方法学.

王燕女, (1967-), 博士, 教授. 研究方向为半导体物理与器件、深亚微米器件及纳米电子器件建模与 EDA 工具开发、微波/毫米波集成电路设计.