

# 器件模型参数优化提取的混合算法\*

郝 跃 贾新章

(西安电子科技大学微电子研究所, 西安)

1987年11月10日收到

本文论述了器件模型参数优化提取中常采用的 L-M 算法存在的一些问题。提出了用 (LM-BFGS) 混合算法提取器件模型参数, 这对优化收敛有极大的加速作用。本文例举了大量实验函数的计算结果并结合 MOSFET 的参数提取, 证明了该算法可以大大减少迭代的次数和计算非线性目标函数的次数。

**主题词:** 器件模型, 优化, 算法

## 一、引言

集成电路的优化设计和器件模型参数的优化提取中采用的目标函数是由理论响应与实际的相对误差构成的最小二乘函数。已经证明采用 L-M 方法<sup>[1]</sup>对最小二乘目标函数的优化无论对初值的要求和收敛速度方面都比其它非线性规划方法要好。在一般情况下, L-M 方法具有超线性收敛速度。但是, 当目标函数的具体形式较复杂或初值赋得不当时, 目标函数的值可能较大, 由 L-M 方法所产生的下降方向会较大地偏离目标函数的下降方向, 甚至完全偏离下降方向。这样的结果使 L-M 方法的超线性收敛的优点丧失, 甚至出现不收敛的情况。

在集成电路器件参数优化提取过程中, 初值的选取往往带有一定的经验估计, 因此容易引起最初迭代时目标函数较大。这时用 L-M 方法的收敛速度较慢, 甚至发散。为了解决这个问题, 本文提出了用 (LM-BFGS) 混合算法。根据各种实验函数的例证并结合 SPICE 中的 MOS3 模型<sup>[2]</sup>参数的优化提取, 证明了以上问题能得到很好解决。

## 二、参数优化的混合算法

设集成电路器件模型参数所组成的向量为  $\mathbf{z} = (z_1, z_2, \dots, z_N)^T$ 。参数优化提取的目标函数为

$$E(\mathbf{z}) = \frac{1}{2} \sum_{i=1}^M r_i^2(\mathbf{z}) = \frac{1}{2} \mathbf{r}^T(\mathbf{z}) \mathbf{r}(\mathbf{z}) \quad (1)$$

\* 本工作得到国家自然科学基金资助

式中  $r_i$  为理论与实际响应函数相对误差(残差)的第  $i$  个抽样点, 并且  $M > N$ . 参数优化问题可以描述为寻找最佳点  $\mathbf{z}^*$ , 使得

$$\min_{\mathbf{z}} E(\mathbf{z}) = \min_{\mathbf{z}} \frac{1}{2} \sum_{i=1}^M r_i^2(\mathbf{z}) = \frac{1}{2} \mathbf{r}^T(\mathbf{z}^*) \mathbf{r}(\mathbf{z}^*) \quad (2)$$

一般情况下, 对问题 2 这样的非线性最小二乘形式, 在第  $k$  步优化迭代中的新搜索方向  $s^k$  可以根据

$$B^k s^k = -g^k \quad (3)$$

得到, 新的改进点  $\mathbf{z}^{k+1}$  为

$$\mathbf{z}^{k+1} = \mathbf{z}^k + \alpha^k s^k \quad (4)$$

式中  $\mathbf{g}^k = \nabla E(\mathbf{z}^k) = A(\mathbf{z}^k) \mathbf{r}(\mathbf{z}^k)$ ,  $A$  为  $\mathbf{r}(\mathbf{z})$  的 Jacobian 矩阵, 即  $a_{ij} = \partial r_i(\mathbf{z}) / \partial z_j$ .  $\alpha^k$  为一维搜索得到的最佳步长。为了使  $s^k$  为  $E(\mathbf{z})$  的下降方向,  $B^k$  为对称正定矩阵。对具有二阶收敛性的算法  $B^k$  近似为  $E(\mathbf{z})$  的 Hessian 矩阵  $G(\mathbf{z})$ . 对式(1)求二阶偏导:

$$B^k = G(\mathbf{z}^k) = \nabla^2 E(\mathbf{z}^k) = \left( AA^T + \sum_{i=1}^M r_i \nabla^2 r_i \right)_k \quad (5)$$

不同的非线性梯度方法就是对  $B$  用不同的近似函数。其中 L-M 方法表述  $B$  为

$$\begin{aligned} B^k &= LM(A^k, \lambda_k) \\ LM(A, \lambda) &= (AA^T + \lambda A_{TT}) \end{aligned} \quad (6)$$

式中  $A_{TT} = \text{diag}(AA^T)$ ,  $\lambda$  为阻尼系数。 $\lambda A_{TT}$  的引入是为了防止在极值点附近  $B^k$  出现病态。从式(5)和(6)知, 当残差  $r_i$  较小时,  $LM(A, \lambda)$  是对 Hessian 矩阵的极好近似, 这时 L-M 算法能保证超线性收敛且调用目标函数的次数也较少。但是当所赋的初值偏离  $\mathbf{z}^*$  太远使得  $r_i$  很大时, 式(5)右端的第二项很可能接近或超过第一项的影响。这时  $B = LM(A, \lambda)$  会偏离对称正定性, 从而使  $s^k$  偏离下降方向, 使收敛速度变慢, 甚至发散。

考虑到在非线性规划中, 目前最好的变尺度方法为 BFGS 法<sup>[3]</sup>。该方法具有超线性收敛速度, 且  $B^k$  矩阵在任意  $r_i$  情况下都能保证与 Hessian 矩阵有较好的近似。只是对于最小二乘目标函数的优化其收敛速度一般情况下比 L-M 法慢, 特别是计算目标函数的次数比 L-M 法要多。根据 L-M 和 BFGS 法的性质与特点, 我们提出用 (L-M-BFGS) 混合算法对器件参数进行优化提取。在  $r_i$  较大时, 采用 BFGS 法;  $r_i$  较小时采用 L-M 算法。即

$$B^k = \begin{cases} \text{BFGS}(B^{k-1}, \delta^{k-1}, \gamma^{k-1}) & \text{如果 } \Delta(B_1^k, \delta^{k-1}, \gamma^{k-1}) < \Delta(B_2^k, \delta^{k-1}, \gamma^{k-1}) \\ \text{LM}(A^k, \lambda_k) & \text{其它情况} \end{cases} \quad (7)$$

式中:

$$\text{BFGS}(B, \delta, \gamma) = B + \frac{\gamma \gamma^T}{\delta^T \gamma} - \frac{B \delta \delta^T B}{\delta^T B \delta}$$

$$\delta^{k-1} = \mathbf{z}^k - \mathbf{z}^{k-1}$$

$$\gamma^{k-1} = g^k - g^{k-1} = A^k A^{kT} \delta^{k-1} + (A^k - A^{k-1}) r^k$$

$\Delta(\cdot)$  为 Al-Baali 和 Fletcher 提出的一个测度算式<sup>[4]</sup>, 该测度算式表征的近似的 Hessian 矩阵与 Hessian 矩阵之间的误差,

$$\Delta(B^k, \delta^{k-1}, \gamma^{k-1}) = (a^2 - 2b + 1)^{1/2} \quad (8)$$

式中  $a = \frac{\gamma^T H \gamma}{\delta^T \gamma}$ ,  $b = \frac{\delta^T \delta}{\delta^T B \delta}$ ,  $H$  为所采用  $B$  矩阵的逆 ( $H = B^{-1}$ ). 式(7)中  $B_k^k = \text{BFGS}(B^{k-1}, \delta^{k-1}, \gamma^{k-1})$ ,  $B_k^k = LM(A^k, \lambda_k)$ . 所以式(7)的判据说明了: 如果 LM 算法  $B^k$  矩阵与 Hessian 矩阵的误差比 BFGS 法的大, 则采用 BFGS 方法, 否则采用 L-M 算法. 采用混合算法的结果可以保证  $r$  较大时  $B^k$  同样具有对称正定性, 使其超线性收敛在不同初值情况下均能满足.

### 三、算法比较及 MOSFET 参数提取

为了验证 LM-BFGS 方法的可行性, 用不同的实验函数对算法进行测试并与 L-M 和 BFGS 方法比较.

#### 1. 最简单的实验函数

$$r_1(z_1, z_2) = z_1^2 + 10z_1z_2 + 4z_2^2 + 0.7401006 = 0$$

$$r_2(z_1, z_2) = z_1^2 - 3z_1z_2 + 2z_2^2 - 1.0201228 = 0$$

其最优解为  $z^* = (0.37655, -0.43795)^T$ . 表 1 给出了在不同初值情况下三种方法的比较(初值选  $z^0 = z_1^0$ , 收敛精度为  $E(z^*) \leq 10^{-12}$ ).

表 1 几种算法比较

初值 \ 收敛次数	方法	L-M	BFGS	混合法
-1		4	8	7
-10		7	11	7
-10 <sup>2</sup>		10	12	7
-10 <sup>3</sup>		F	14	7
-10 <sup>4</sup>		F	17	7
-10 <sup>5</sup>		F	48	8

(F 表示发散)

从计算结果看出, 当初值接近于  $z^*$  时, L-M 法迭代次数比其它方法少. 随着初值偏离  $z^*$ , L-M 法的迭代效果与 BFGS 法接近. 但当  $z^0$  偏离  $z^*$  较大时 L-M 法发散. 从表中也可以看出混合的 LM-BFGS 法在初值变化范围较大的情况下仍具有稳定的收敛速度, 这正是优化设计所需要的.

#### 2. 不同实验函数下的算法比较

取目前比较流行的优化算法实验函数<sup>[5]</sup>, 对上述的三种算法收敛的迭代次数 ( $NN$ ), 计算目标函数的次数 ( $NF$ ) 和计算梯度次数 ( $NG$ ) 进行比较, 其结果如表 2 所

示。从表中可清楚说明 LM-BFGS 方法的三项性能都是最好的。

表 2 三种优化算法的比较

实验函数	L-M			BFGS			混合法		
	NN	NF	NG	NN	NF	NG	NN	NF	NG
Rosenbrock	15	98	15	17	169	18	8	72	8
Powell(0)	12	93	12	20	167	21	12	89	12
Powell(1)	21	144	21	20	178	23	13	90	15
Powell(2)	21	117	21	20	231	21	15	108	17
Helix	20	75	20	15	62	16	12	53	12
Davidoni	11	91	11	6	56	7	4	46	4
Box	27	97	27	35	129	37	17	73	18

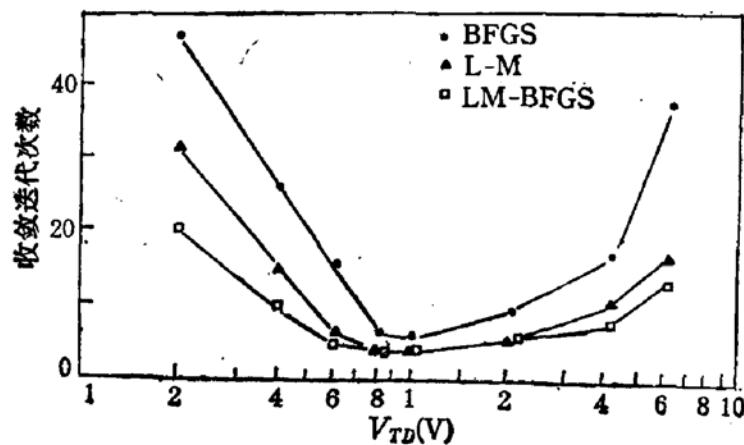


图 1 不同初值下三种算法收敛次数比较

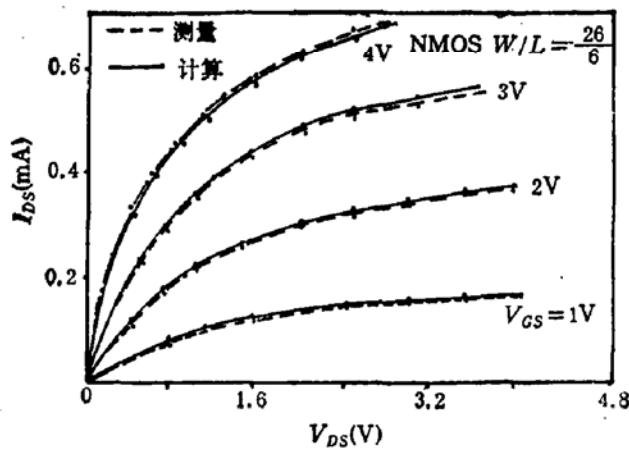


图 2 MOSFET 直流特性比较

### 3. MOSFET 模型参数的最优化提取

MOS3 模型包括了 15 个直流模型参数，设  $r_i = (1 - I_{DSi}^M / I_{DSi})$ ， $I_{DSi}$  为由模型参数计算得到的第  $i$  个源漏电流抽样点， $I_{DSi}^M$  为相应的实测抽样点，确定抽样点的状态变量为  $(V_{GSi}, V_{DSi}, V_{BSi})$ 。图 1 为只改变零偏阈值电压  $V_{TO}$  的初值而固定其它 14 个参数

初值时,几种算法的收敛次数受初值影响的情况。

从图 1 中看出,当  $z^0 \rightarrow z^*$  时,几种算法的收敛次数几乎相差不大。当初值偏离  $z^*$  较大时,LM-BFGS 混合算法比其它两种算法要好得多。图 2 为对器件进行参数最优化提取后计算的理论曲线与实测曲线的比较。其各点的最大误差为 8.27%, 这能较好地满足器件分析和机辅设计的要求。

通过实验函数和 MOSFET 器件模型参数优化结果证明,采用 LM-BFGS 混合算法的收敛性比较稳定。在初值偏离最优解较远的情况下,该算法仍然保证稳定的超线性收敛。这对器件参数的优化提取及其它领域的最小二乘目标函数的优化设计都是极为有益的。

本文工作期间得到了朱吉良同志的帮助,特表感谢。

### 参 考 文 献

- [1] D. W. Marquardt, *J. Soc. Indst. Appl. Maths.*, 11, 431 (1963).
- [2] A. Vladimirescu and S. Liu, ERL Memo, UCB/ERL M80/7, Electronics Res. Lab. Univ. of California Berkeley (1980).
- [3] D. F. Shanno, *Maths. Compt.*, 24, 647 (1970).
- [4] M. Al-Baali and R. Fletcher, *J. Oper. Res. Soc.*, 36, 405 (1985).
- [5] M. G. David, *ACM Trans. Math. Softw.*, 7, 348(1981).

## A Hybrid Algorithm for Parameter Optimal Extraction of Device Model

Hao Yue and Jia Xinzhang

(Microelectronics Research Institute, Northwest Telecommunications Engineering Institute, Xi'an)

### Abstract

Some problems of L-M algorithm used to extract optimal parameters of device model are described, and a hybrid algorithm for the parameter optimal extraction is suggested. This paper gives some examples calculating the popular testing functions and extracting device parameters for MOSFET model, and demonstrates that the algorithm can reduce a large number of the iterative numbers for convergence and the numbers for calculating objective function.

**KEY WORDS:** Device model, Optimization, Algorithm