Communications to SIMAI Congress,
ISSN 1827-9015, Vol. 1 (2006)

DOI: 10.1685/CSC06144

Optimization of electronic circuits

Jan ter Maten
Philips Research Laboratories
Hich Tech Campus 48, 5656 AE Eindhoven, The Netherlands

Jan.ter.Maten@philips.com

Achie Lin, Ahmed El Guennouni
Magma Design Automation
TUE Campus, Den Dolech 2, Dommel Building Z-Wing 8, 5612 AZ Eindhoven, The Netherlands
{Achie,Ahmed}@Magma-DA.COM

1 Introduction

Several types of parameters $\mathbf{p}$ influence the behaviour of electronic circuits and have to be taken into account when optimizing appropriate performance functions $f(\mathbf{p})$: design parameters $\mathbf{x}$, manufacturing process parameters $\mathbf{s}$, and operating parameters θ.
The performance functions $f(\mathbf{p})$ and the constraints $c(\mathbf{p})$ can be costly to evaluate and are subject to noise. For both, the dependency on $\mathbf{p}$ can be highly nonlinear.
In this paper we will describe our in-house developed method for optimization and our experiences with it. Also some new directions for further research will be described.

2 Constrained optimization by augmented Lagrangian

In this section we restrict our parameters $\mathbf{p}$ to the design variables $\mathbf{x}$, which can be geometrical quantities like transistor width W and length L. The designer can adjust them during the process of optimizing the performance of an actual design. The performance function $f(\mathbf{x})$ can be gain, transit frequency, phase margin, power dissipation, delay, slew rate, bandwidth, etc, that can be obtained from different types of analyses: DC, AC, Transient.
The search for the optimal values of the optimization variables (OVs) $\mathbf{x}$ can be formulated as a nonlinear constrained optimization problem in n variables with m constraints,

$$\begin{aligned} &\text{minimize} && f(\mathbf{x}), && \mathbf{x} = (x_1, x_2, \ldots, x_n), \\ &\text{subject to} && c_i(\mathbf{x}) \le 0, && i = 1, \ldots, m, \\ & && a_j \le x_j \le b_j, && j = 1, \ldots, n, \end{aligned} \tag{2.1}$$

where x_i denotes the i-th OV. The values of the objective function $f(\mathbf{x})$ and the constraining functions $c_i(\mathbf{x})$ are obtained from circuit simulation. The performance and stability of the optimization algorithm is affected by the *scaling* of the OVs, of $f(\mathbf{x})$ and of the $c_i(\mathbf{x})$ [5]. By introducing slack variables $s_i \ge 0$, the augmented Lagrangian penalty function can be written as [1, 10],

$$\Phi_{\mathrm{ALAG,s}}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \mathbf{s}) = \underbrace{f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i [c_i(\mathbf{x}) + s_i]}_{\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda})} + \sum_{i=1}^{m} \mu_i [c_i(\mathbf{x}) + s_i]^2, \tag{2.2}$$

in which $\mathcal{L}$ is the standard Lagrangian. The parameters λ_i and μ_i are Lagrange multipliers and penalty factors, respectively. Minimization over the slack variables s_i yields the simplified merit function that is used in Gridmom,

$$\Phi_{\rm ALAG}(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^{m}\left\{\mu_i \max\left[c_i(\mathbf{x}) - \frac{\lambda_i}{2\mu_i}, 0\right]^2 - \frac{\lambda_i^2}{4\mu_i}\right\}. \tag{2.3}$$

The optimal point $(\mathbf{x}^\star,\boldsymbol{\lambda}^\star)$ becomes a stationary point of $\mathcal{L}$ and satisfies the Karush-Kuhn-Tucker (KKT) conditions [1],

$$\begin{gathered}\nabla_x\mathcal{L}(\mathbf{x}^\star,\boldsymbol{\lambda}^\star) = 0,\\ c_i(\mathbf{x}^\star) \le 0, \quad \lambda_i^\star \le 0, \quad \lambda_i^\star c_i(\mathbf{x}^\star) = 0, \quad i = 1,\ldots,m.\end{gathered} \tag{2.4}$$

A basic observation is that there are $|\mu_i| < \infty$ such that $\Phi_{\rm ALAG}(.,.,\boldsymbol{\mu})$ has a local minimum in $(\mathbf{x}^\star,\boldsymbol{\lambda}^\star)$. Also the location becomes independent of μ_i (when large enough). For updating $\boldsymbol{\lambda}$ we observe:

$$\nabla\Phi_{\rm ALAG}(\mathbf{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = \nabla\mathcal{L}(\mathbf{x},\boldsymbol{\lambda}',\boldsymbol{\mu}) = 0 \rightarrow \lambda_i' = \lambda_i - 2\max[\mu_i c_i(\mathbf{x}),\lambda_i]$$

[cf also Primal-Dual methods]. The basic method used to solve the optimization problem

Algorithm 2.1 Method of Multipliers [5]

1: Start: $k=1$, $x_i = x_i^{(0)}$, $\lambda_i = \lambda_i^{(0)} = 0$, $\mu_i = \mu_i^{(0)}$
2: Set $\Phi(\mathbf{x}) = \Phi_{\rm RSM}(\mathbf{x},\boldsymbol{\lambda}^{(k-1)},\boldsymbol{\mu}^{(k-1)};\mathbf{x}^{(k-1)})$
3: Minimize: $\mathbf{x}^{(k)} = \text{argmin}_{\mathbf{x}}\Phi(\mathbf{x})$
4: If (update μ_i) $\mu_i^{(k)} = 10 * \mu_i^{(k-1)}$ (but $< \mu_{\max}$!)
5: Else update $\lambda_i^{(k)} = \lambda_i^{(k-1)} - 2\max[\mu_i^{(k-1)}c_i(\mathbf{x}^{(k-1)}),\lambda_i^{(k-1)}]$
6: Endif
7: Test for optimality
8: If not optimal set $k = k+1$ and goto 2.

is the Method of Multipliers (Algorithm 2.1).
In step 3 a trust region approach is applied on a response surface model around $\mathbf{x}^{(k-1)}$ to solve the minimization problem (for details see [5]). In step 4 we keep $\mu_i^{(k)} < \mu_{\max}$. This is based on the basic observation made above: in the end we may trust that only the $\lambda_i^{(k)}$ should do the work. Note also that preventing $\mu_i^{(k)}$ to become too large also keeps the problem well conditioned.
Convergence of a related algorithm is discussed in [10], assuming that $f, c_i \in C^2$. The problem of the reduced smoothness, caused by eliminating the slack variables in (2.3), is discussed in the appendix of [11].
In our case the whole approach is applied on a grid (i.e. all $\mathbf{x}^{(k)}$ are projected to the nearest grid point), that subsequently is refined, like in [4]. This prevents the algorithm from going too fast to a small scale (and then get stuck in a local minimum). However, in practice equally appreciated, is that it also allows to store and re-use expensive parts of $f(\mathbf{x})$ and of $c_i(\mathbf{x})$ during the building of the response surface model in the optimization process.
At initialization we apply a Uniform Design approach which is based on number theory.

Actually this approach is problem independent. Here improvements can be expected by introducing additional techniques, like pattern search methods [7], that add more information from the problem itself to obtain initial starts for the above method. Also Kriging techniques [2, 3, 6] can be used in this phase, which additionally offer a more global optimization aspect, when also applied later in the process. However, experience shows that only a small reduction in the number of overall function evaluations was obtained.

3 Toward Robust Design

The additional types of parameters, manufacturing process parameters $\mathbf{s}$ and operating parameters $\boldsymbol{\theta}$, introduce additional requests.

- The first ones, $\mathbf{s}$, reflect statistical process fluctuations, like T_{ox} (oxide thickness), and V_{th0} (threshold voltage), substrate doping, characterized by their distribution functions.
- The operating parameters $\boldsymbol{\theta}$ include supply voltage V_{sup} and temperature T. Here additional constraints are found that require $c(\mathbf{x}, \boldsymbol{\theta}) \leq 0$ should hold for all $\boldsymbol{\theta}$ within some interval.

Most statistical parameters appear as transistor model parameters and cannot be modified by the designer if a fully qualified production process is assumed. However, production variation must be taken into account. As IC technologies scale down to finer feature sizes, it becomes increasingly difficult to control the relative process variations. Hence large variations can be observed. Also non-normality must be taken into account. The APEX algorithm [8] addresses this point. A response surface model of f is built that is quadratic in the process variations. The method is heavily based on the explicit calculation of the stochastic moments of f and on AWE-techniques (with their drawbacks) to efficiently approximate the transfer function. Here more robust model order reduction techniques may be applied.
An interesting test function to consider the effect of operating parameters is

$$c(x, \theta) = (\theta - 2)^2 \frac{x \ \sin(x)}{2}, \quad \text{with} \ \ x \in [0, 6], \ \ \theta \in [0, 5]. \tag{3.1}$$

In the worst-case analysis in [9] first linearization with respect to θ is done, and a θ_{WC} is derived. Next, this value is used in the linearization with respect to x, and a x_{WC} is determined. This process is iteratively repeated in a Gauss-Seidel-like manner. For several problems this algorithm gives reasonable results and yield estimations (as is confirmed by our own experiments). However, the algorithm appears not to be robust on the above constraint function c, which gives way for further research.

REFERENCES

1. M.S. Bazaraa, H.D. Sherali, C.M. Shetty: *Nonlinear programming: theory and algorithms*, 2nd edition, John Wiley and Sons, New York (1993).
2. W.C.M. van Beers, J.P.C. Kleijnen: *Kriging interpolation in simulation: a survey*, In: R.G. Ingalls, M.D. Rossetti, J.S. Smith, B.A. Peters (Eds.): *Proc. of the 2004 Winter Simulation Conference*, Association for Computing Machinery, New York, 2004.

3. W.C.M. van Beers: *Kriging metamodeling for simulation*, PhD-Thesis, Univ. of Tilburg, The Netherlands, 2005.
4. C. Elster, A. Neumaier: *A grid algorithm for bound constrained optimization of noisy functions*, IMA J. Numer. Anal., Vol. 15, pp. 585–608, 1995.
5. T. Heijmen, C. Lin, J. ter Maten, M. Kole: *Augmented Lagrangian algorithm for optimizing analog circuit design*, in: A. Buigis, R. Ciegis, A.D. Fit: *Progress in industrial mathematics at ECMI 2002*, Mathematics in Industry 5, Springer, pp. 179–184, 2003.
6. D.R. Jones, M. Schonlau, W.J. Welch: *Efficient global optimization of expensive black-box functions*, J. of Global Optimization, Vol. 13, pp. 455–492, 1998.
7. T.G. Kolda, R.M. Lewis, V. Torcson: *Optimization by direct search: new perspectives on some classical and modern methods*, SIAM Review, Vol. 45-3, pp. 385–482, 2003.
8. X. Li, J. Le, P. Gopalakrishnan, L.T. Pileggi: *Asymptotic probability extraction for non-normal distributions of circuit performance*, Proc. ICCAD 2004.
9. M. Pronath, H. Graeb, K. Antreich: *On parametric test design for analog integrated circuits considering error in measurement and stimulus*, In: K. Antreich, R. Bulirsch, A. Gilg, P. Rentrop (Eds.): *Modeling, simulation and optimization of integrated circuits*, In. Series of Numer. Maths. Vol. 146, pp. 283–301, 2003.
10. J.F. Rodriguez, J.E. Renaud, L.T. Watson: *Convergence of trust region augmented Lagrangian methods using variable fidelity approximation data*, Structural Optim., Vol. 15, pp. 141–156, 1998.
11. J.F. Rodriguez, J.E. Renaud, L.T. Watson: *Trust region augmented Lagrangian methods for sequential response surface approximation and optimization*, Proc. of ASME Design Engineering Technical Conferences DETC 97, Sacramento CA, DET97DAC3773, 1997.