Comparison of two methods for design centering

Luis M. Vidigal
Carnegie Mellon University

Stephen W. Director
NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS:
The copyright law of the United States (title 17, U.S. Code) governs the making of photocopies or other reproductions of copyrighted material. Any copying of this document without permission of its author may be prohibited by law.
COMPARISON OF TWO METHODS FOR DESIGN CENTERING

by

L. M. Vidigal and S.W. Director

DRC-18-31-81

September 1981
ABSTRACT

Two methods recently proposed to solve the design centering problem [2,9] are compared. Although the methods are formulated differently, they are shown, under general assumptions, to yield the same solutions. Simplifications required to make the methods efficiently implementable introduce, however, significant differences from an utilization point of view.

1. Introduction

Finding the best nominal design in order to maximize the yield is an important problem in IC design. Because of the unavoidable fluctuations m the manufacturing process, the actual values of the circuit parameters, denoted by the vector p, is characterized by the known joint density distribution <p> where p = represents the nominal values. Design centering methods [1,5,31 try to imbed in the so-called region of acceptability R the largest convex domain B_r(p) related to <p>, as shown below. The region of acceptability can for our purposes be defined as the point

\[ R_a = \{ p \mid f_j(y) \leq f_{\text{max}}, j = 1, \ldots, m \} \tag{1} \]

where the f_j represent the performance functions which characterize the circuit behavior. R_a is assumed to be simply connected. The contours of equal probability of <p> can be associated for all the distribution of interest with a norm n(p). See for example [3]. B_r(p), often referred to as a norm body, is defined relative to n(p)

\[ B_r(p) = \{ p \mid n(p-p^*) \leq r \} \tag{2} \]

and represents a body centered at p* whose size is proportional to r. The first method we will look into is the approach referred to as (VTP) in [2], and can be formulated as

\[ \text{maximize:} \quad f(y) \tag{3} \]

\[ \text{subject to:} \quad \max_{y \in B_r} f_j(y) \leq f_{\text{max}}, j = 1, \ldots, m \]

In this approach a maximally sized body is to be found, such that inside the body none of the performance functions will exceed their maximum allowable value. We note from the outset that the main difficulty in solving (3) derives from the maximization subproblem

\[ \text{maximize} \quad f_j(y) \tag{4} \]

\[ \text{subject to} \quad \max_{y \in B_r} f_j(y) \leq f_{\text{max}}, j = 1, \ldots, m \]

as it must be solved for all constraints i * 1, m and is likely to have local maxima, making gradient based methods unreliable. Thus (4) is likely to have local maxima can be inferred from the fact that quite often the f_j are convex functions and B_r(p) is a parallelepiped.

The second method, referred to as the (HP) method, is formulated as:

\[ (UP): \text{maximize} \quad \min_{W \in R_a} \max_{i \in I} f_i(y) \tag{5} \]

In this method the points on the boundary of the region of acceptability which are closest to the nominal design are located and then this distance is maximized. Searching for these near points is based upon the fact that they limit further expansion of the body B_r(p). It is possible to prove that if the performance functions are either quasiconvex or quasiconcave, their accumulation points are the points where the largest body touches the boundary of the region of acceptability.

The domain of the innermost minimization is the intersection of the region of acceptability with the surface f_j(y) = f^*, i.e. in the boundary of R_a, we will assume this domain to be nonempty. Otherwise the constraint f_j(y) = f^* is superfluous in the sense that if dropped from the set of constraints the region of acceptability remains unchanged. An algorithm can easily detect this situation by verifying that for the constraint f_j there is no feasible solution to the minimization problem. The main difficulty that arises in solving (5) is the subproblem

\[ \text{minimize} \quad n(y-p^*) \tag{6} \]

\[ \text{such that} \quad f_j(y) \leq f^*, j = 1, \ldots, m \]

\[ f_j(y) \geq f_{\text{max}} \]

The difficulty here is that (6) is a constrained minimization problem which is computationally intensive in itself, and one which must be solved repeatedly.

2. Equivalence of Methods

It is illustrative to recognize that if R_a is simply connected, the performance functions are differentiable and their gradients do no vanish at the boundary of R_a, then a locally optimal solution to (3) is also locally optimal to (5).
For convenience we define

\[ \text{bd}_i(R) = \{ p \in R_i | f_i(r - f_i) \leq 0 \} \]

and the states

\[ M_i(p, r) = \text{min} \left\{ f_i(p) \right\} \]  

Vincten center proved [10] that a solution \((0^*, \gamma)\) to problem (3) is locally optimal if

\[ f_i(\delta \cdot 0^*, \gamma) \leq f_{\max} \]  

and

\[ 0 \in \text{Co} \left\{ f_i(p) \right\} \]  

This last expression can be interpreted geometrically to mean that the convex hull defined by gradients of the constraint at the points where the body touches the boundary of \( R_0 \), contains the origin, intuitively, if this situation occurs, there is no direction in which the center could be moved such that the radius of the body could increase.

To establish the equivalence between the two problems it is enough to show that any point \( 0 \in M_i, (f(0), f) \), exists in \( \text{bd}_i(R) \) and is at a minimum distance from \( 0^* \). Assume \( p^* \in \text{bd}_i(R) \) such that \( 0 < f_{\max} \) then \( p^* \) exists in the interior of \( B_0, (f, i) \) and there is an \( \omega > 0 \) such that \( B_0 \times B = \omega \). By assumption \( V_i(p^*) \) and there is therefore a point \( P^* \in B_0(p^*) \) such that \( f_i(p^*) > Kp^* \). But this contradicts the hypothesis that \((0, \gamma)\) is optimal as for \( p^* \) we would have \( p^* \in B_0(0^*, \gamma) \).

3. Relation Between Subproblems (4) and (6)

The relation between problems (4) and (6) can be made clearer by dropping from (6) the requirement that \( f_i(r) \leq f_{\max} \). If the constraint is not superfluous, then constrained optimization problems for most cases are not active and the solution to (6) would remain unchanged. We rewrite therefore (6) in the following form

\[ \min_{n(y, p^*)} \]  

such that

\[ f_i(p^*) \geq f_{\max} \]  

On the other hand, using (2) subproblem (4) can also be rewritten as

\[ \max_{n(y, p^*)} \]  

such that

\[ f_i(p^*) \leq f_{\max} \]  

for \( p^* \) and \( r \) constants.

As illustrated in Fig. 1 for the norm \( n(x) = \| x \| = \max | x_1 |, ... | x_i |, \) in subproblem (10) for a constant sized parallelepiped we search for the point where \( f_i \) is maximized. In (11) the opposite problem is solved: given a fixed value of the function \( f_i \), the minimum value of \( r \) is found such that the resulting norm body will have at least a common point with the surface \( f_i(p) = f_{\max} \).

It is interesting to note that (11) and (12) are in some sense duals of each other since they have the objective function and the constraint interchanged. We saw before that the solution to these two subproblems for (5, 7) was the same. We further notice that a local solution to (11), assuming that \( n(x) \) is locally Lipschitz, is given by [4].

\[ 0 \in \text{Co} \left\{ f_i(y) \right\} \]  

where \( 3n(y, p^*) \) represents the generalized gradient of \( n(y, p^*) \). Similarly, for (12)

\[ 0 \in \text{Co} \left\{ n(y, p^*) \right\} \]  

The similarity of these problems is enhanced if we notice that at the solution of (3) and (5) we must have \( r^* \).

4. Implementation of the Methods

The implementation of the two methods is similar in the sense that the subproblems (4) and (6) are solved at each iteration and the function and gradient information gathered during this process is used by the outer maximization. In the first case the problem can be reduced to a constrained nonlinear problem which can be
solved by some constrained variable metric methods. Specifically, a variation of Powell's algorithm [7] is used, in the second method, the equivalent information is generated to produce a second order approximation to the constraining surfaces and the largest normbody is inscribed in that simplified approximation to the region of acceptability.

Serving subproblem (4) at each iteration would result in a computationally expensive adjustment. Therefore, in the second method, the equivalent information is used to generate a second order approximation to the region of different solutions obtained with both methods. (Note also the theoretical advantage of the (VTP) methods of not requiring a feasible starting point).

Example 1 In this example we have a single cone. Quadratic performance function

\[ f(p) = \| p \|^2 - 50(p^2_1 + p^2_2) \]

is constrained to be smaller than a 1.

<table>
<thead>
<tr>
<th>Starting Point</th>
<th>Final Point</th>
<th>Number of F.E.</th>
<th>Number of O.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(VTP) (2.0, 0.0)</td>
<td>(0.5, 1.0)</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>(NP) (0.5, 0.0)</td>
<td>(0.5, 1.0)</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

The (NP) method does not have this limitation.*

Although methods exist for maximizing problems when the objective function is inadmissible, for example [6], and this is presently an area of active research, all algorithms for this general problem tend to require a very large number of function and gradient (generalized gradients) evaluations making them unsuitable to be used in subproblem (6) where repeated solutions are required. In our implementation we limited the algorithm (NP) to the case where the region of acceptability is not one (formally convex) or the region of acceptability has a hole in its interior, as illustrated in Fig. 2.

![Figure 2:](image)

The (NP) method does not have this limitation.

From the designer's point of view, the (NP) method yields interesting information on how well the constraints are formulated. At early stages of a design, it is very often the case that a large number of constraints is not quite specified. If the solution to (6) corresponds to a value of the constraints, the constraints can be modified. Further, the relative distances to the final nominal give an indication on how strongly the constraint affects the circuits yield.

A limitation of the NP method which could be serious is that the initial design is very far from the final solution, the quadratic approximations might have to be updated, thereby significantly increasing the computation time.

5. Examples

We tried both methods on a group of examples to compare the behavior of both methods. For the (VTP) we used the infinite norm, while for the (NP) we used the L_1 norm, this accounts for the different solutions obtained with both methods. (Note also the theoretical advantage of the (VTP) methods of not requiring a feasible starting point).

Example 2 This example is taken from [8]

\[ f_1(p) = e^p_1 + 1 (p_2^2 - 1) \]
\[ f_2(p) = e^p_1 - p_2 + 1 \]
\[ f_3(p) = p_2^2 + p_2^2 + 1 \]

The constraints are \( x_i \leq 1.5, i = 1, 2, 3 \)

<table>
<thead>
<tr>
<th>Starting Point</th>
<th>Final Point</th>
<th>Number of F.E.</th>
<th>Number of O.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(VTP) (2.0, 0.0)</td>
<td>(0.5, 1.0)</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>(NP) (0.1, 0.1)</td>
<td>(0.3, 1.0)</td>
<td>15</td>
<td>46</td>
</tr>
</tbody>
</table>

![Figure 3:](image)

![Figure 4:](image)
Example 3 This example was taken from [8].

\[
\begin{align*}
 f_1(p) &= 1.5p_1(1-p_2) \\
 f_2(p) &= 2.25p_1(1-p_2^2) \\
 f_3(p) &= 2.625p_1(1-p_2^2) \\
\end{align*}
\]

Constrained to \( f_1 \leq 1.5 \), \( j = 1, \ldots, 8 \).

<table>
<thead>
<tr>
<th>Starting Point</th>
<th>Final Point</th>
<th>Number of F.E.</th>
<th>Number of G.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(VTP)</td>
<td></td>
<td>33</td>
<td>67</td>
</tr>
<tr>
<td>(NP)</td>
<td></td>
<td>24</td>
<td>20</td>
</tr>
</tbody>
</table>

Starting Point Final Point Number of F.E. Number of G.E.
(VTP) (2.0,2.0) (1.961,0.134) 33 67
(NP) (1.3,0.3) (1.822,0.142) 24 20

Conclusion

We have shown that two apparently very different problems can be seen as different formulations of the same body imbedding problem, yielding under general conditions an identical result. With reasonable assumptions the problems lend themselves to very different implementations: one (VTP) suitable for the worst case problem, the other (NP) for a general body center problem when the norm is differentiable. It is interesting to note further that when applied to similar problems they seem to have computational requirements of the same order, when measured in terms of the number of function and gradient evaluations.

References

[1] Bandler, John W., Liu, Peter C. and Chen, James H. K.
Worst Case Network Tolerance Optimization.
IEEE Transactions on Microwave Theory and Techniques

[2] Brayton, Robert K., Director, S. W., Hachtel, Gary D. and
Vidigal, L. M.
A New Algorithm for Statistical Circuit Design Based on
Quasi-Newton Methods and Function Splitting.
IEEE Transactions on Circuits and Systems CAS-26(9):754-764,
September, 1979.

Yield Maximization and Worst-Case Design with Arbitrary
Statistical Distributions.
IEEE Transactions on Circuits and Systems CAS-27(9):755-764,
September, 1980.

[4] Clarke, Frank H.
A Nonlinear Approach to Lagrange Multipliers.
Mathematics of Operations Research 1(2):165-174, May,
1978.

[5] Director, Stephen W. and Hachtel, Gary D.
The Simplicial Approximation Approach to Design
Centering.
IEEE Transactions on Circuits and Systems CAS-24(7):363-372,

An Algorithm for Constrained Optimization with
Semismooth Functions.
1977.

[7] Powell, M.I.D.
A Fast Algorithm for Nonlinearly Constrained Optimization
Calculations.
In A. Dold and B. Eckmann (editor), #630 Lect. Notes in

Algorithms for Worst-Case Tolerance Optimization.
IEEE Transactions on Circuits and Systems CAS-26(9):775-783,
September, 1979.

[9] Vidigal, L.M. and Director, S.W.
Design Centering: The Quasi-Convex, Quasi-Concave
Performance Function Case.
In International Symposium on Circuits and Systems. IEEE,
1980.

Personal Communication.
1983.