The Gradient Associated Conjugate Direction Method

Min Zhang

Deutsches Elektronen-Synchrotron (DESY) -MPY-
Notkestrasse 85, 22603 Hamburg, Germany

Abstract

A new conjugate direction (CD) based method, GaCD by name, is addressed with its full derivations. Comparison with some commonly used CD methods is made, concluding that the GaCD method is among the best of them, in terms of an assessment number. A practical example is presented in which a waveguide-microstrip transition section is optimized with respect to its transmission parameter S_{21}. The measurement of the transition agrees very well with the numerical results.

1 Introduction

A numerical optimization procedure can be divided into two distinctive parts: local extrema search and global area search, which is really necessary for most practical optimization problems. In its analytical counterpart, however, there will be no such clear classification. There, the determination of the global optimum is done relatively explicitly.

Here we will only concentrate on the local extrema search phase. As known, in a certain close vicinity of a local extremum, the goal function in question can be well approximated by an appropriate quadratic function, say, the 2nd order approximation of its Taylor's expansion series at that point.

For quadratic functions, there is a unique property that if there exist n conjugate directions, at most n 1d (one-dimensional) extrema searches along these directions are needed before the function's extremum is reached. This property is usually referred to as *quadratic convergence* or *quadratic cut-off*.

Based on this property, there developed a class of optimization methods, which are usually designated by the name of "Conjugate Direction" (CD) methods, e.g. DFP, Fletcher-Reeves, Powell, Daniel, Sorenson-Wolfe, etc [1].

In all the CD-derived methods, the key point is how to generate successively linearly independent search directions, which are conjugate to one another. In the GaCD method here, the gradient "G" of the goal function is used for providing an offset from current optimization point, which will never degenerate to zero unless a stagnation point has been reached. The offset serves as an indispensable ingredient for generating a next conjugate direction in this method. The above statements hold exactly true at least when the method is applied to the quadratic functions.

2 The GaCD Method

2.1 Preliminaries

Any n dimensional quadratic functions can be written in the following form:

$$\mathcal{G}(\mathbf{X}) = \frac{1}{2}\mathbf{X}^T\mathbf{H}\mathbf{X} + \mathbf{b}^T\mathbf{X} + c \quad (\mathbf{X} \in \Re^n). \tag{1}$$

$\mathbf{H}$ is an $n \times n$ positive-definite symmetric matrix. Therefore the function has one minimum point which is denoted by $\mathbf{X}^*$. $\mathbf{X}$ and $\mathbf{b}$ are n-dimensional vectors; c is a constant; T denotes transposition.

Conjugate: If there exist two n-dimensional vectors $\mathbf{P}_i$ and $\mathbf{P}_j$, and they fulfill:

$$\mathbf{P}_i^T\mathbf{H}\mathbf{P}_j = 0, \qquad \mathbf{P}_i^T\mathbf{H}\mathbf{P}_i \neq 0, \quad \textit{and} \quad \mathbf{P}_j^T\mathbf{H}\mathbf{P}_j \neq 0, \tag{2}$$

then the two vectors $\mathbf{P}_i$ and $\mathbf{P}_j$ are called *conjugate* to each other with respect to $\mathbf{H}$, or simply, $\mathbf{H}$-*conjugate*.

Definitions below will be used throughout the derivation:

$$\mathbf{Q}_i \equiv \nabla\mathcal{G}(\mathbf{X}_i) = \mathbf{H}\mathbf{X}_i + \mathbf{b}, \tag{3}$$

$$\mathbf{R}_i \equiv \mathbf{Q}_{i+1} - \mathbf{Q}_i, \tag{4}$$

$$\nabla^2\mathcal{G}(\mathbf{X}_i) \equiv \nabla \cdot \nabla\mathcal{G}(\mathbf{X}_i) = \mathbf{H}. \tag{5}$$

Let us assume that $\mathbf{P}_i$ $(i = 1, 2, \cdots, n)$ are n $\mathbf{H}$-conjugate vectors. Since $\mathbf{H}$-conjugate vectors are linearly independent of one another, $\mathbf{X}^*$ can be expressed with the n $\mathbf{H}$-conjugate vectors in the following form:

$$\mathbf{X}^* = \sum_{i=1}^{n} \lambda_i^*\mathbf{P}_i, \tag{6}$$

with λ_i^* $(i = 1, 2, \cdots, n)$ being coefficients to be determined. Since $\mathbf{X}^*$ is an extremum point, it satisfies

$$\nabla\mathcal{G}(\mathbf{X}^*) = \mathbf{0}, \quad \text{i.e.} \tag{7}$$

$$\mathbf{H}\mathbf{X}^* + \mathbf{b} = \mathbf{0}. \tag{8}$$

Substituting $\mathbf{X}^*$ in Eq. 8 with Eq. 6 yields

$$\mathbf{H}\left(\sum_{i=1}^{n} \lambda_i^*\mathbf{P}_i\right) + \mathbf{b} = \mathbf{0}. \tag{9}$$

Using the conjugate property (Eq. 2), we obtain

$$\lambda_i^* = -\frac{\mathbf{P}_i^T\mathbf{b}}{\mathbf{P}_i^T\mathbf{H}\mathbf{P}_i}. \tag{10}$$

Now the key point is how to find out $\mathbf{H}$ and $\mathbf{b}$. For large dimensional problems, finding them is almost impossible or too time-consuming. Thus such a direct solution is not practically usable. We have to figure out other ways. One of them is an iteration procedure using the so-called *accurate 1d search*. To illustrate this method in a clearer manner, we split it into two steps. In the first, the concept of *accurate 1d search* is introduced, and then derivation of the iteration procedure is presented.

2.2 Accurate 1d Search

Below is a 1d search problem

$$\mathcal{G}(\mathbf{X}_{k+1}) \equiv \mathcal{G}(\mathbf{X}_k + \lambda_k \mathbf{P}_k) \equiv \min_{\forall \lambda \in \Re} \{\mathcal{G}(\mathbf{X}_k + \lambda \mathbf{P}_k)\}, \tag{11}$$

with $\mathbf{X}_k$ being current point and $\mathbf{P}_k$ current search direction.

To solve this 1d problem, the following commonly used method can be employed:

$$\frac{d\mathcal{G}}{d\lambda} \mid_{\lambda=\lambda_k} = [\nabla \mathcal{G}(\mathbf{X}_k + \lambda \mathbf{P}_k)]^T \mid_{\lambda=\lambda_k} \mathbf{P}_k = 0. \tag{12}$$

Due to

$$\nabla \mathcal{G}(\mathbf{X}_{k+1}) = \mathbf{H}\mathbf{X}_{k+1} + \mathbf{b}, \tag{13}$$

inserting Eq. 13 into Eq. 12 yields

$$\begin{aligned} (\mathbf{H}\mathbf{X}_{k+1} + \mathbf{b})^T \mathbf{P}_k &= [\mathbf{H}(\mathbf{X}_k + \lambda_k \mathbf{P}_k) + \mathbf{b}]^T \mathbf{P}_k \\ &= [(\mathbf{H}\mathbf{X}_k + \mathbf{b}) + \lambda_k \mathbf{H}\mathbf{P}_k]^T \mathbf{P}_k \\ &= \nabla^T \mathcal{G}(\mathbf{X}_k) \mathbf{P}_k + \lambda_k \mathbf{P}_k{}^T \mathbf{H}\mathbf{P}_k \\ &= 0 \end{aligned} \tag{14}$$

(Note: The property $\mathbf{H}^T = \mathbf{H}$ is utilized above.). Then we obtain

$$\lambda_k = -\frac{\mathbf{P}_k{}^T \nabla \mathcal{G}(\mathbf{X}_k)}{\mathbf{P}_k{}^T \mathbf{H}\mathbf{P}_k} = -\frac{\mathbf{P}_k{}^T \mathbf{Q}_k}{\mathbf{P}_k{}^T \mathbf{H}\mathbf{P}_k}. \tag{15}$$

The above procedure is called *accurate 1d search.*

2.3 The Derivation

Assume we have n $\mathbf{H}$-conjugate vectors $\mathbf{P}_i$ $(i = 1, 2, \cdots, n)$ and an initial point $\mathbf{X}_1$, performing accurate 1d search once along $\mathbf{P}_1$ leads to $\mathbf{X}_2$:

$$\mathbf{X}_2 = \mathbf{X}_1 + \lambda_1 \mathbf{P}_1, \tag{16}$$

(Note: λ_1 satisfies Eq. 15.). After n such searches which are performed along $\mathbf{P}_i$ $(i = 1, 2, \cdots, n)$ respectively, a point denoted by $\mathbf{X}_{n+1}$ is reached:

$$\mathbf{X}_{n+1} = \mathbf{X}_1 + \sum_{i=1}^{n} \lambda_i \mathbf{P}_i, \tag{17}$$

with λ_i $(i = 1, 2, \cdots, n)$ all fulfilling Eq. 15.

We will prove that $\mathbf{X}_{n+1}$ must be the minimum point $\mathbf{X}^*$, i.e.

$$\mathbf{Q}_{n+1} = \nabla \mathcal{G}(\mathbf{X}_{n+1}) = \mathbf{0}. \tag{18}$$

From Eqs. 3 and Eq. 17, we have

$$\begin{aligned}
\mathbf{Q}_{n+1} &= \nabla\mathcal{G}(\mathbf{X}_{n+1}) \\
&= \mathbf{H}\left(\sum_{i=1}^{n}\lambda_i\mathbf{P}_i + \mathbf{X}_1\right) + \mathbf{b} \\
&= \sum_{i=1}^{n}\lambda_i\mathbf{H}\mathbf{P}_i + (\mathbf{H}\mathbf{X}_1 + \mathbf{b}) \\
&= \sum_{i=1}^{n}\lambda_i\mathbf{H}\mathbf{P}_i + \mathbf{Q}_1. \qquad (19)
\end{aligned}$$

Multiplying both sides with $\mathbf{P}_1^T$:

$$\begin{aligned}
\mathbf{P}_1^T\mathbf{Q}_{n+1} &= \mathbf{P}_1^T\left(\sum_{i=1}^{n}\lambda_i\mathbf{H}\mathbf{P}_i + \mathbf{Q}_1\right) \\
&\overset{Eq.2}{=} \lambda_1\mathbf{P}_1^T\mathbf{H}\mathbf{P}_1 + \mathbf{P}_1^T\mathbf{Q}_1 \\
&\overset{Eq.15}{=} -\frac{\mathbf{P}_1^T\mathbf{Q}_1}{\mathbf{P}_1^T\mathbf{H}\mathbf{P}_1}\mathbf{P}_1^T\mathbf{H}\mathbf{P}_1 + \mathbf{P}_1^T\mathbf{Q}_1 \\
&= 0; \qquad (20)
\end{aligned}$$

And multiplying with $\mathbf{P}_2^T$:

$$\begin{aligned}
\mathbf{P}_2^T\mathbf{Q}_{n+1} &= \mathbf{P}_2^T\left(\sum_{i=1}^{n}\lambda_i\mathbf{H}\mathbf{P}_i + \mathbf{Q}_1\right) \\
&= \mathbf{P}_2^T\left(\sum_{i=2}^{n}\lambda_i\mathbf{H}\mathbf{P}_i + \lambda_1\mathbf{H}\mathbf{P}_1 + (\mathbf{H}\mathbf{X}_1 + \mathbf{b})\right) \\
&= \mathbf{P}_2^T\left(\sum_{i=2}^{n}\lambda_i\mathbf{H}\mathbf{P}_i + \mathbf{H}(\lambda_1\mathbf{P}_1 + \mathbf{X}_1) + \mathbf{b}\right) \\
&\overset{Eq.16}{=} \mathbf{P}_2^T\left(\sum_{i=2}^{n}\lambda_i\mathbf{H}\mathbf{P}_i + \mathbf{H}\mathbf{X}_2 + \mathbf{b}\right) \\
&= \mathbf{P}_2^T\left(\sum_{i=2}^{n}\lambda_i\mathbf{H}\mathbf{P}_i + \mathbf{Q}_2\right) \\
&\overset{Eq.15}{=} 0. \qquad (21)
\end{aligned}$$

Repeating the above for all other $\mathbf{P}_i$ $(i = 3, \cdots, n)$, we have

$$\mathbf{P}_i^T\mathbf{Q}_{n+1} = 0 \quad (\forall i = 1, 2, \cdots, n). \qquad (22)$$

As known, **H**-conjugate vectors are linearly independent of one another. It means that the n **H**-conjugate vectors ($\mathbf{P}_i$, $i = 1, 2, \cdots, n$) span a complete basis over $\Re^n$. A vector in $\Re^n$ which is orthogonal to all the base vectors is nothing but a zero vector. That is,

$$\mathbf{Q}_{n+1} = \mathbf{0}. \qquad (23)$$

Then Eq. 18 gets proved. From the above, following conclusions can be drawn:

For any function which can be expressed by Eq. 1, if there exist n **H***-conjugate n-dimensional vectors* $\mathbf{P}_i$ $(i = 1, 2, \cdots, n)$*, then its extremum point* $\mathbf{X}^*$ *can be obtained by doing at most n accurate 1d searches in the directions of* $\mathbf{P}_i$ $(i = 1, 2, \cdots, n)$*, respectively.*

This relieves us from the difficulty in using the explicit form for λ's (Eq. 15), since normally the matrix $\mathbf{H}$ is unknown. In the following, whenever we say "do accurate 1d search", it implies an evaluation of Eq. 15.

Next we will discuss how to generate $\mathbf{H}$-conjugate vectors. First, let us derive three important properties, from which the method is deduced.

$$\begin{aligned}
\mathbf{P}_i{}^T\mathbf{H}\mathbf{P}_j &= \frac{1}{\lambda_j}\mathbf{P}_i{}^T\mathbf{H}\lambda_j\mathbf{P}_j = \frac{1}{\lambda_j}\mathbf{P}_i{}^T\mathbf{H}(\mathbf{X}_{j+1} - \mathbf{X}_j) \\
&= \frac{1}{\lambda_j}\mathbf{P}_i{}^T[(\mathbf{H}\mathbf{X}_{j+1} + \mathbf{b}) - (\mathbf{H}\mathbf{X}_j + \mathbf{b})] = \frac{1}{\lambda_j}\mathbf{P}_i{}^T(\mathbf{Q}_{j+1} - \mathbf{Q}_j) \\
&= \frac{1}{\lambda_j}\mathbf{P}_i{}^T\mathbf{R}_j \begin{cases} = 0 & (i \neq j) \\ \neq 0 & (i = j) \end{cases} \quad (i, j = 1, 2, \cdots, n)
\end{aligned} \tag{24}$$

Eq. 24 can also be written in the following form:

$$\mathbf{P}_i{}^T\mathbf{Q}_i = \mathbf{P}_i{}^T\mathbf{Q}_{i-1} = \mathbf{P}_i{}^T\mathbf{Q}_{i-2} = \cdots = \mathbf{P}_i{}^T\mathbf{Q}_1 \quad (i = 1, 2, \cdots, n). \tag{25}$$

Below is given the derivation of another conclusion which reads as

$$\mathbf{P}_i{}^T\mathbf{Q}_j = 0 \quad (i < j). \tag{26}$$

Let $j = i + 1$,

$$\begin{aligned}
\mathbf{P}_i{}^T\mathbf{Q}_{i+1} &= \mathbf{P}_i{}^T(\mathbf{H}\mathbf{X}_{i+1} + \mathbf{b}) = \mathbf{P}_i{}^T[\mathbf{H}(\mathbf{X}_i + \lambda_i\mathbf{P}_i) + \mathbf{b}] \\
&= \mathbf{P}_i{}^T[(\mathbf{H}\mathbf{X}_i + \mathbf{b}) + \lambda_i\mathbf{H}\mathbf{P}_i] = \mathbf{P}_i{}^T\mathbf{Q}_i + \lambda_i\mathbf{P}_i{}^T\mathbf{H}\mathbf{P}_i \\
&\overset{Eq.15}{=} \mathbf{P}_i{}^T\mathbf{Q}_i - \frac{\mathbf{P}_i{}^T\mathbf{Q}_i}{\mathbf{P}_i{}^T\mathbf{H}\mathbf{P}_i}\mathbf{P}_i{}^T\mathbf{H}\mathbf{P}_i = 0.
\end{aligned} \tag{27}$$

Likewise, let $j > i + 1$,

$$\begin{aligned}
\mathbf{P}_i{}^T\mathbf{Q}_j &= \mathbf{P}_i{}^T(\mathbf{H}\mathbf{X}_j + \mathbf{b}) = \mathbf{P}_i{}^T[\mathbf{H}(\mathbf{X}_{j-1} + \lambda_{j-1}\mathbf{P}_{j-1}) + \mathbf{b}] \\
&= \mathbf{P}_i{}^T[(\mathbf{H}\mathbf{X}_{j-1} + \mathbf{b}) + \lambda_{j-1}\mathbf{H}\mathbf{P}_{j-1}] = \mathbf{P}_i{}^T\mathbf{Q}_{j-1} + \lambda_{j-1}\mathbf{P}_i{}^T\mathbf{H}\mathbf{P}_{j-1} \\
&= \mathbf{P}_i{}^T\mathbf{Q}_{j-1} = \cdots \\
&= \mathbf{P}_i{}^T\mathbf{Q}_{i+1} = 0.
\end{aligned} \tag{28}$$

Equation 26 gets proved.

This equation indicates that

The gradient $\mathbf{Q}_{k+1}$ *after k accurate 1d searches is orthogonal to all former search directions* ($\mathbf{P}_i$, $i = 1, 2, \cdots, k$).

Besides this, it also delivers us another deep connotation: after n accurate 1d searches, the last gradient $\mathbf{Q}_{n+1}$ satisfies ${\mathbf{P}_i}^T\mathbf{Q}_{n+1} = 0$, $(\forall i = 1, 2, \cdots, n)$. Since, as known, there are at most n linearly independent vectors in an n-dimensional linear space, the only solution to this equation is $\mathbf{Q}_{n+1} = \mathbf{0}$, which means that $\mathbf{X}_{n+1}$ is nothing but the extremum point $\mathbf{X}^*$.

A next search direction may be formed using $\mathbf{Q}_{k+1}$. It is this that the so-called *conjugate gradient method* is based on. Here we will use it for another purpose, namely, for finding an offset which does not lie in the subspace which is spanned by all previous search directions. The actual usefulness of the offset will be investigated right after.

To derive the GaCD method, one more conclusion is needed.

> *Known that* $\mathbf{P}_i$ $(i = 1, 2, \cdots, k)$ *are* $\mathbf{H}$*-conjugate.* k *successive accurate 1d searches in these directions from two different initial points, say* $\mathbf{X}_1'$ *and* $\mathbf{X}_1''$ *have been performed, yielding other two points* $\mathbf{X}_{k+1}'$ *and* $\mathbf{X}_{k+1}''$*. Then the difference vector* $(\mathbf{X}_{k+1}' - \mathbf{X}_{k+1}'')$ *is* $\mathbf{H}$*-conjugate to* $\mathbf{P}_i$ $(i = 1, 2, \cdots, k)$*, that is,*

$$(\mathbf{X}_{k+1}' - \mathbf{X}_{k+1}'')^T\mathbf{H}\mathbf{P}_i = 0 \quad (\forall i = 1, 2, \cdots, k). \tag{29}$$

Since

$$\mathbf{X}_{k+1}' = \mathbf{X}_1' + \sum_{i=1}^{k} \lambda_i'\mathbf{P}_i, \tag{30}$$

$$\mathbf{X}_{k+1}'' = \mathbf{X}_1'' + \sum_{i=1}^{k} \lambda_i''\mathbf{P}_i, \tag{31}$$

we have

$$\mathbf{X}_{k+1}' - \mathbf{X}_{k+1}'' = \mathbf{X}_1' - \mathbf{X}_1'' + \sum_{i=1}^{k}(\lambda_i' - \lambda_i'')\mathbf{P}_i. \tag{32}$$

For any $\mathbf{P}_j$ $(j = 1, 2, \cdots, k)$,

$$\begin{aligned}
(\mathbf{X}_{k+1}' - \mathbf{X}_{k+1}'')^T\mathbf{H}\mathbf{P}_j &= (\mathbf{X}_1' - \mathbf{X}_1'')^T\mathbf{H}\mathbf{P}_j + \sum_{i=1}^{k}(\lambda_i' - \lambda_i'')\mathbf{P}_i{}^T\mathbf{H}\mathbf{P}_j \\
&= (\mathbf{X}_1' - \mathbf{X}_1'')^T\mathbf{H}\mathbf{P}_j + (\lambda_j' - \lambda_j'')\mathbf{P}_j{}^T\mathbf{H}\mathbf{P}_j \\
&= (\mathbf{X}_1' - \mathbf{X}_1'')^T\mathbf{H}\mathbf{P}_j + \mathbf{P}_j{}^T\mathbf{Q}_j'' - \mathbf{P}_j{}^T\mathbf{Q}_j' \\
&= [\mathbf{H}(\mathbf{X}_1' - \mathbf{X}_1'')]^T\mathbf{P}_j + \mathbf{P}_j{}^T\mathbf{Q}_j'' - \mathbf{P}_j{}^T\mathbf{Q}_j' \\
&= [(\mathbf{H}\mathbf{X}_1' + \mathbf{b}) - (\mathbf{H}\mathbf{X}_1'' + \mathbf{b})]^T\mathbf{P}_j + \mathbf{P}_j{}^T\mathbf{Q}_j'' - \mathbf{P}_j{}^T\mathbf{Q}_j' \\
&= (\mathbf{Q}_1' - \mathbf{Q}_1'')^T\mathbf{P}_j + \mathbf{P}_j{}^T\mathbf{Q}_j'' - \mathbf{P}_j{}^T\mathbf{Q}_j' \\
&= \mathbf{P}_j{}^T\mathbf{Q}_1' - \mathbf{P}_j{}^T\mathbf{Q}_1'' + \mathbf{P}_j{}^T\mathbf{Q}_j'' - \mathbf{P}_j{}^T\mathbf{Q}_j' \\
&\overset{Eq.25}{=} 0.
\end{aligned} \tag{33}$$

Because the above assumption is made for an arbitrary j, Eq. 29 gets proved.

Equation 29 provides us an effective way to acquire search directions which are $\mathbf{H}$-conjugate while Eq. 26 may be utilized to get an offset point which is needed by Eq. 29. No care needs to be taken whether the offset will degenerate to zero or not, since, if true, it means that current point is already the extremum one, i.e. the $\mathbf{X}^*$, therefore, no more search is needed.

2.4 The Algorithm

Based upon all the arguments above, the GaCD method can be formed as follows. $a \Longleftarrow b$ stands for *assign b to a*. ϵ is a given stop-criterion. $\mathbf{X}_1$ is an initial point.

1. $j \Longleftarrow 0$;
2. $j \Longleftarrow j+1$; $\mathbf{X}_j^0 \Longleftarrow \mathbf{X}_j$;
3. Find $\mathbf{Q}_j^0 = \nabla \mathcal{G}(\mathbf{X}_j^0)$ and do *accurate 1d search* in the $\mathbf{Q}_j^0$, i.e.
$$\mathbf{X}_j^0 \Longleftarrow \mathbf{X}_j^0 + \lambda_j^0 \mathbf{Q}_j^0;$$
4. If $j > 1$, then for $i = 1, 2, \cdots, j-1$, do *accurate 1d searches*:
$$\mathbf{X}_j^i \Longleftarrow \mathbf{X}_j^{i-1} + \lambda_j^i \mathbf{P}_i;$$
5. $\mathbf{X}_{j+1} \Longleftarrow \mathbf{X}_j^{j-1}$; $\mathbf{P}_j \Longleftarrow \mathbf{X}_{j+1} - \mathbf{X}_j$;
6. If $j > 1$, then do one more *accurate 1d search* in $\mathbf{P}_j$:
$$\mathbf{X}_{j+1} \Longleftarrow \mathbf{X}_{j+1} + \lambda_j^j \mathbf{P}_j;$$
7. If $j < n$, then go back to 2; Else,
if $|\mathbf{X}_{n+1} - \mathbf{X}_n| < \epsilon$, then stop; Else
$$\mathbf{X}_1 \Longleftarrow \mathbf{X}_{n+1},$$
and go back to 1.

2.5 A 2d Verification

Assume we have a 2d quadratic function as follows:

$$\mathcal{G}(x_1, x_2) = x_1^2 + 2x_2^2 - 4x_1 - 2x_1x_2. \tag{34}$$

It is found that

$$\mathbf{H} = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} -4 \\ 0 \end{pmatrix}, \quad c = 0.$$

It reaches its minimum at point $\mathbf{X}^* = (x_1^*, x_2^*)^T = (4, 2)^T$.

By selecting an initial point

$$\mathbf{X}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

and following the algorithm above step by step, we obtain the final result $\mathbf{X}^\wedge = (4, 2)^T = \mathbf{X}^*$ after two search loops, which verifies the *quadratic cut-off* property of the method. Note that the

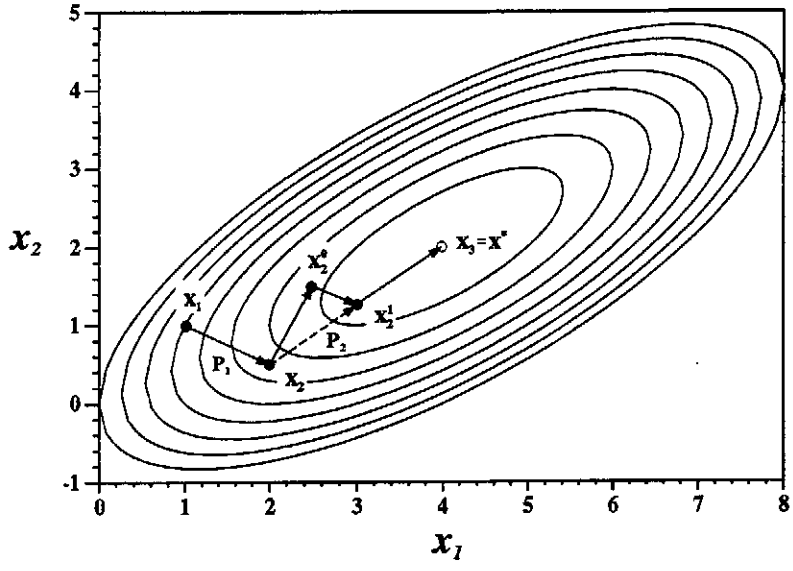

Figure 1: Search Path in Minimizing $\mathcal{G}(x_1, x_2)$

above verification holds true both analytically and numerically. The search path for the numerical verification is shown in Fig. 1.

Starting from $\mathbf{X}_1$, evaluating the gradient at that point, using the gradient as a search direction $\mathbf{P}_1$, it reaches $\mathbf{X}_2$ after the first loop. Then a search in the current gradient is performed, yielding $\mathbf{X}_2^0$, which serves as the offset point. Searching in $\mathbf{P}_1$, we get a new point $\mathbf{X}_2^1$, which is a counterpart to $\mathbf{X}_2$, for they are both the extremum points in the same search direction ($\mathbf{P}_1$) but from two different initial points, $\mathbf{X}_2^0$ and $\mathbf{X}_1$ respectively. From Eq. 29, it is known that the difference vector $\mathbf{P}_2 = (\mathbf{X}_2^1 - \mathbf{X}_2)$ must be $\mathbf{H}$-conjugate to $\mathbf{P}_1$. Now down to this direction $\mathbf{P}_2$, the algorithm finally converges to $\mathbf{X}_3 = \mathbf{X}^*$.

3 Comparison with Other CD Methods

The comparison is made by optimizing nine "standard" test functions [2], which are listed in Appendix A. The methods to be checked are GaCD, DFP, Fletcher-Reeves, M-Powell (modified Powell), of which DFP and Fletcher-Reeves' are among the best and most popularly used CD ones. In addition, the Evolution Strategy (ES) [3] is also tested and presented here for the readers who may be interested in such stochastic methods.

In order that a uniform assessment of optimization performance of the methods can be achieved, a parameter η is introduced, which is defined as

$$\eta \equiv \frac{|\log(\mathcal{G}^\wedge/\mathcal{G}^0)|}{N_\mathcal{G}} \times 100,$$

where $\mathcal{G}^0$ are $\mathcal{G}^\wedge$ are the initial and final goal values, respectively; $N_\mathcal{G}$ is the number of function evaluations. The greater the number, the more effective the method is supposed to be. In the definition, we took the number of evaluations of goal function as a factor which reflects the computation time, not directly the time itself. It is because for numerical optimization procedures,

the evaluation of a goal function means a complete solution of an electromagnetic problem, which takes, normally, the most significant part of the whole computing time including the optimization strategy and other auxiliary operations.

The test runs are performed on a SUN SPARC-1 workstation. Each run takes normally less than one second. Table 1 shows the statistic summary and rankings of the methods, which are derived from the raw calculation results in Appendix B. It is seen that the GaCD method is the most robust (σ_η) while its effectiveness (η) comes to the second after the DFP method.

Comparison of Various Optimization Methods via Common Test Functions (Statistics&Rankings)							
		GaCD	**BFGS**	**DFP**	**Fletcher**	**M-Powell**	**ES**
Rankings	$<\eta_o>$	0.86 **(1)**	0.84 **(2)**	0.77 **(3)**	0.65 **(4)**	0.35 **(5)**	0.17 **(6)**
	σ_{η_o}	0.17 **(2)**	0.24 **(5)**	0.19 **(3)**	0.22 **(4)**	0.28 **(6)**	0.10 **(1)**
BEALE		1.0	0.6284	0.7339	0.8970	0.0527	0.1085
CRAGG		0.9053	0.2524	1.0	0.3437	0.2660	0.1003
ENGVALL-2d		0.9170	1.0	0.8207	0.6696	0.7281	0.1540
ENGVALL-3d		0.8453	1.0	0.4120	0.3170	0.0	0.2465
POWELL		0.4286	1.0	0.7788	0.4342	0.2757	0.0896
WHITE		0.7663	1.0	0.5217	0.6521	0.3339	0.2352
WOOD		1.0	0.7594	0.7213	0.8710	0.5734	0.3802
ZANGWILL-2d		0.9070	0.9785	1.0	0.8722	0.0873	0.1795
ZANGWILL-3d		1.0	0.9357	0.8963	0.8220	0.8046	0.0534

Table 1: Rankings of the Optimization Methods Tested. The η's associated with respective test functions are normalized with the highest η value of the different methods. For instance, η's for the BEALE function are normalized with that of the M-Powell's, i.e. 1.923, which is the largest in that row (Appendix B). $\eta = 0$ means that the corresponding methods failed for the related functions, e.g. the Fletcher-Reeves method failed in handling the WHITE function. In principle, the larger the number η, the more effective the method is considered to be; while the smaller the number σ_η (deviation of η), the more versatile or stabler the method is regarded.

4 Optimization of a Waveguide-Microstrip Transition

The object to be optimized is a *pair* of transitions between a rectangular waveguide and a microstrip line. Scattering parameter S_{21} is used to assess the performance of the component. It is maximized at the center frequency 10GHz with a ±400MHz bandwidth.

4.1 The Structure

A basic structure of such Waveguide-Microstrip Transition Pairs (WMTP) is shown in Fig. 2. Such layout with two identical back-to-back transitions will accommodate a more accurate mea-

surement.

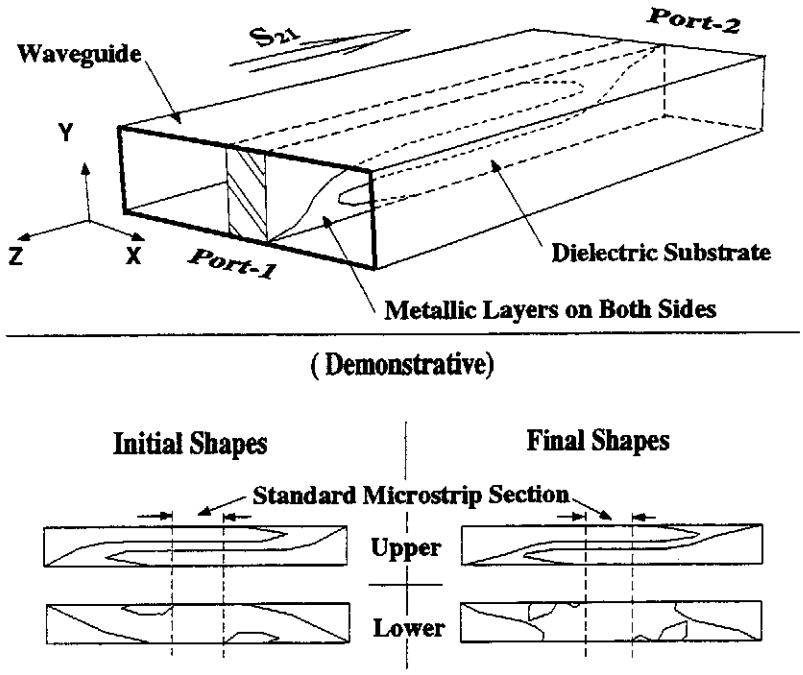

Figure 2: Basic Structure of a WMTP. The standard X-band rectangular waveguide is used, i.e. $a \times b = 22.8\text{mm} \times 10.0\text{mm}$. The dielectric is RT/duroid 5880 from Rogers with a thickness of $1.57 \pm 0.05\text{mm}$, $\epsilon_r = 2.20$.

4.2 Definition of the Problem

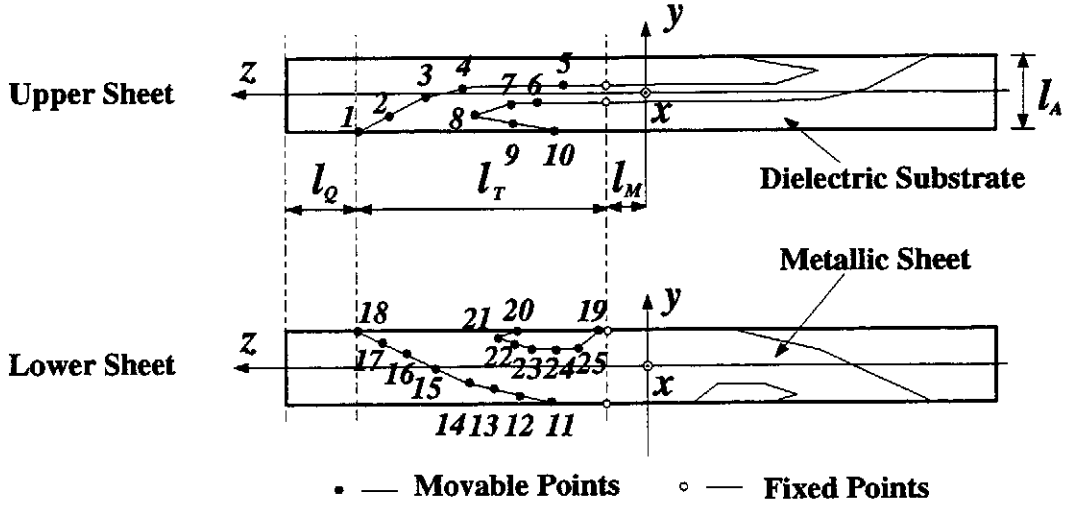

Figure 3: Optimization Vector for WMTP

Total 25 points along the profiles of the upper and the lower metallic sheets are used as the optimization vector $\mathbf{X}$ (Fig. 3). Among them, points #21 to #25 are changeable in both y and z

directions while the rest only in z. It is therefore a problem of 30 degrees of freedom. The definition range is: $l_M \leq z \leq (l_M + l_T)$ for points #1 to #20 and $l_M \leq z \leq (l_M + l_T)$, $-l_A/2 \leq y \leq l_A/2$ for points #21 to #25. The whole set of the definition ranges for $\mathbf{X}$ is denoted by $\mathcal{X} \subset \Re^{30}$. A standard microstrip section in the middle of the structure is maintained by imposing the constraint that points #5, #6, #11, and #19 are not allowed to pass the limit of $z < l_M$.

The goal function is defined as

$$\mathcal{G}(\mathbf{X}) \equiv \overline{S_{21}}(\mathbf{X}), \tag{35}$$

with $\overline{S_{21}}$ being an average of $|S_{21}|$ over the desired bandwidth, that is,

$$\overline{S_{21}} \equiv \frac{1}{n-m} \sum_{i=m}^{n} |S_{21}(f_i)|. \tag{36}$$

f_m to f_n covers the desired bandwidth, i.e. 9.6GHz - 10.4GHz.

The optimization problem can then be expressed as

$$\mathcal{G}^* \equiv \max\{\mathcal{G}(\mathbf{X})\} \quad (\forall \mathbf{X} \in \mathcal{X}). \tag{37}$$

Note: The mode in question is the fundamental one, i.e. H_{10}.

4.3 The Optimization

The optimization is performed by using the optimization driver which is implemented in the general purpose electromagnetic software package – MAFIA [4], which is a finite difference solver. With this driver, the whole optimization procedure is carried out fully automatically, from data initialization to printout of final results.

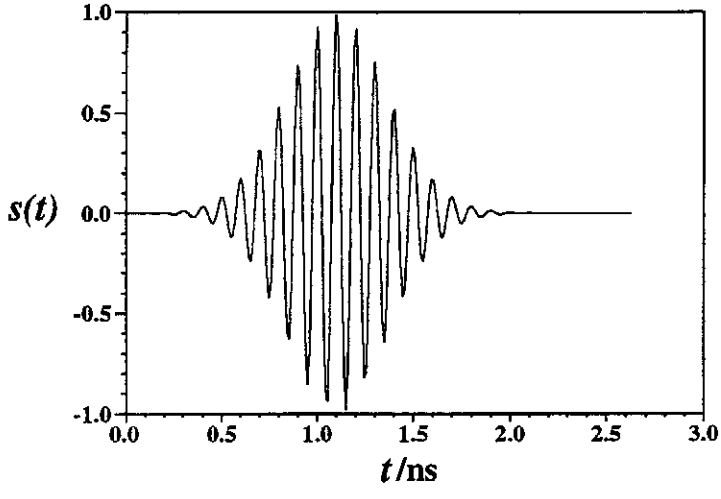

Figure 4: The excitation Signal

The optimization method used is the GaCD presented here. In addition to this local search method, a global search scheme is employed that uses the so-called Evolution Strategy. It should be pointed out that no satisfactory result could have been achieved if this global search phase had not been implemented.

The evaluation of the goal function is achieved by exciting the structure at the port-1 with a Gaussian modulated carrier signal (Fig. 4) and picking up the response at the port-2 of the same mode (i.e. mode H_{10}). After an FFT of the two signals, the goal function can be evaluated. The simulation is the common FDTD procedure.

The simulation is carried out in three dimensions. In the whole process of the optimization, the mesh is fixed to avoid discretization noises, which may cause wrong actions in the optimization strategy. Figure 5 describes the mesh in the area of the microstrip in the y-z plane. In x, an equi-distance mesh is employed. Total number of mesh points is $n_x \times n_y \times n_z = 12 \times 11 \times 59 = 7788$. The resolution for the optimization points is about 1 mm. With this being known, the stop criteria for this optimization should be accordingly set (see "Error Analysis" later).

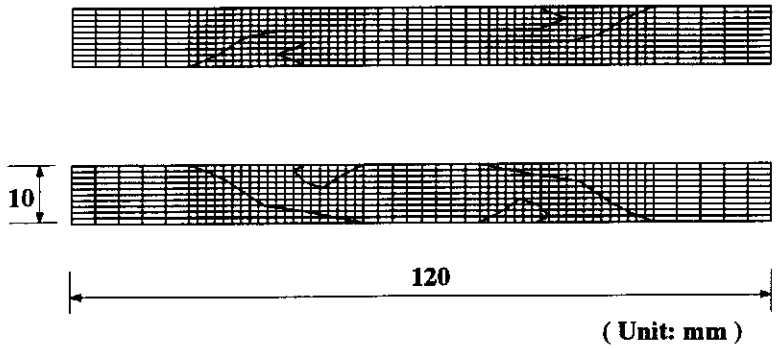

Figure 5: Mesh in the yz Plane

4.4 Results Acquired

After total 338 field calculations (i.e. goal function evaluations), which took about 103 hours on an IBM550 workstation, the final goal value of $\overline{S_{21}} = 0.996$ over the desired bandwidth was achieved, with the initial $\overline{S_{21}}$ being 0.501! The initial and final patterns are shown in Fig. 6.

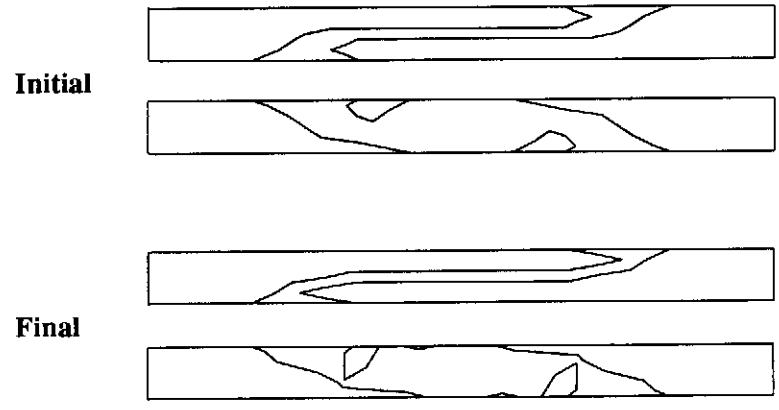

Figure 6: Initial and Final Patterns of the Transition Pair

4.5 The Measurement

Both the initial and the final transition pairs were fabricated and measured (Fig. 7). The grooves in the left block of the mounting waveguide were used for inserting the transition strips.

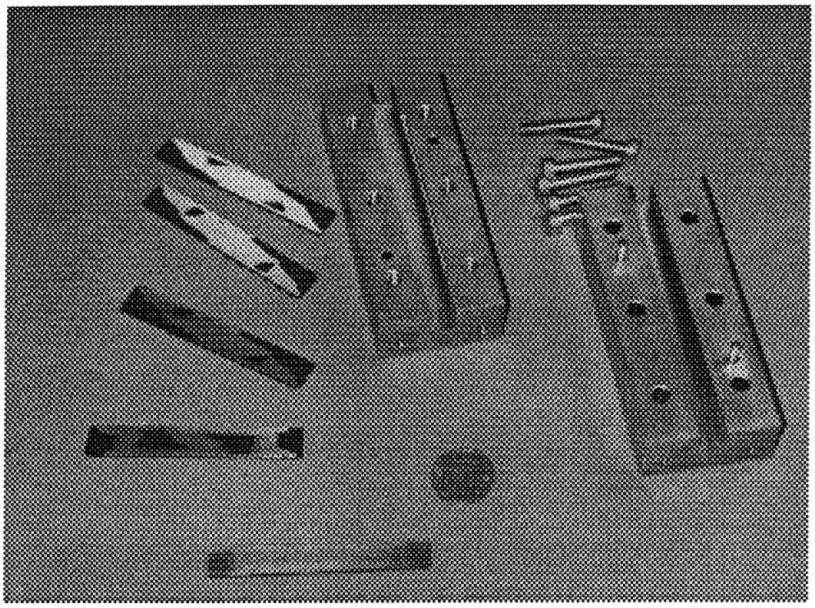

Figure 7: Transition Microstrips and Mounting Waveguide Blocks

The measurement showed a good agreement with the calculated results. The measured S_{21} together with the calculated one can be found in Fig. 8. It is found that the measured S_{21} is generally poorer than the calculated one. It is, most probably, because ohmic losses and the mounting grooves of the structure were not taken into account in the simulation.

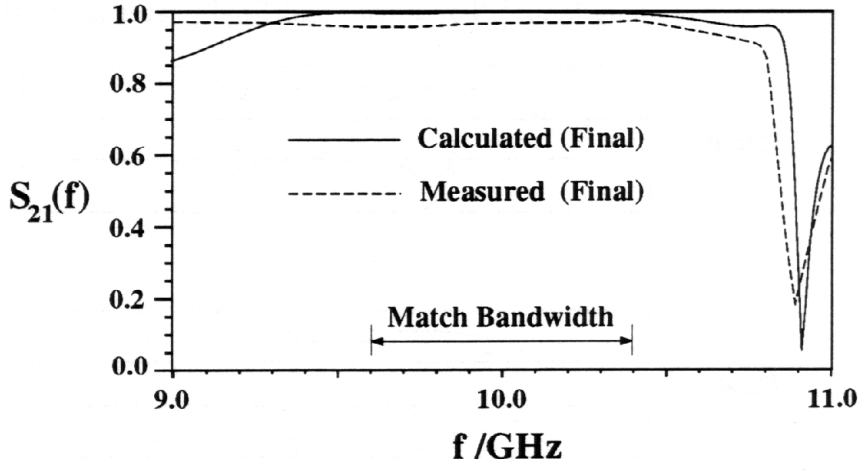

Figure 8: Measured and Calculated S_{21} for the Final Structure

5 Error Analysis

There are mainly three types of error sources. The first is coming from the evaluation of the goal function, or in other words, errors from the numerical solvers of the electromagnetic problems. This part of error has a strong relation with mesh sizes; Computer digital length is another error source, which limits the sensitivity of the optimization algorithms to the fine changes in goal functions. If the changes are too small to be digitized with the available effective length, the algorithm will do nothing and treat the values as unchanged. This can be influenced also by the definition of the goal function, which is the third source of errors. An inappropriate definition of the same problem will surely lead to either a premature stop of the algorithm or a wrong action by it.

5.1 More on Mesh-related Errors

For mesh-supported solvers, like the orthogonal mesh used here, attention should be paid to the selection of mesh. If the mesh is to be fixed throughout whole optimization processes, the stop criteria for the algorithm should be chosen accordingly. Figure 9 describes two different choices of mesh: one is fixed, the other adaptive.

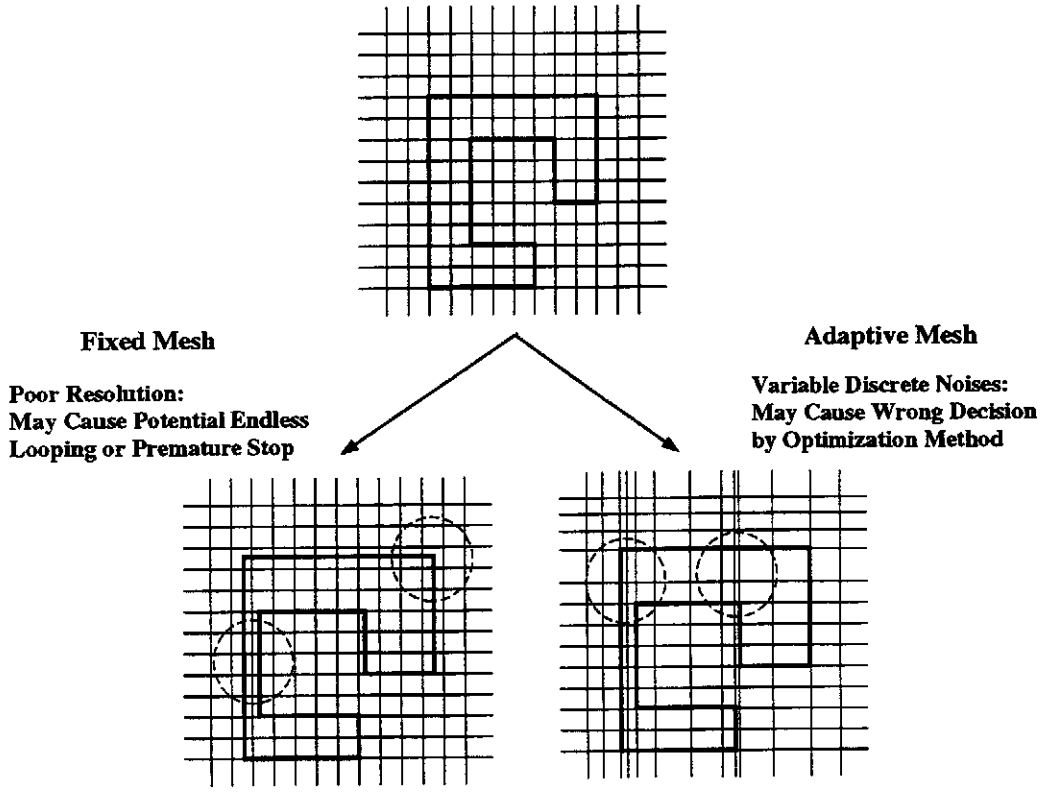

Figure 9: Two Types of Error Source Related to Mesh

For the fixed mesh case, an endless optimization looping is most likely to occur, if the stop criteria are set too fine or bear no account of the repeated failures due to that any fine tuning of the geometry will be impossible if the tune step is smaller than half of the local mesh size. This is

really very hard for practical design problems, since one should have some preknowledges about the relationship between the stop criteria and the mesh sizes. The best solution for this would be to impose a stop criterion which checks such non-improvement failures.

For the adaptive case, there comes another problem where due to computer discretization noises, the algorithms may make wrong decisions if the noises are comparable to the actual changes in the goal function. Such kind of noises is normally solver-type dependent. It becomes severer only when a very accurate result is to be achieved. For FDTD, it is relatively loose compared to eigenvalue cases, where the solution has a relatively high sensitivity to the local mesh size ratios.

6 Conclusions

A full derivation of the GaCD method was presented. In terms of the newly introduced assessment number η, the effectiveness of the method was compared with other commonly used conjugate direction methods against the test functions. It has been shown that the GaCD method is among the best of the methods tested. Further more, a practical optimization was successfully carried out using this new method. A hardware was built based on the numerical results, which delivered a very good performance as numerically predicted.

As a final remark, we would like to point out that for multipeak goal functions, a global search phase is absolutely necessary, since otherwise, starting from different initial points will, most probably, lead to different final results.

References

[1] Donald A. Pierre, "Optimization Theory with Applications", Dover Publications Inc. (New York) 1986, pp314-322

[2] F. A. Lootsma (Ed.), "Numerical Methods for Non-linear Optimization", Academic Press Inc. (London), 1972, pp69-97

[3] M. Zhang, "Numerical Optimization of Electromagnetic Components", PhD Thesis, Technische Hochschule Darmstadt, 1995, Darmstadt, Germany

[4] The MAFIA Collaboration, User's Guide MAFIA Version 3.x, CST GmbH, Lauteschlägerstraße 38, 64289 Darmstadt, Germany

Appendix A

The test functions used are defined below.

1. **BEALE**

$$\mathcal{G}(\mathbf{X}) = [1.5 - x_1(1 - x_2)]^2 + [2.25 - x_1(1 - x_2^2)]^2 + [2.625 - x_1(1 - x_2^3)]^2$$

2. **CRAGG**

$$\mathcal{G}(\mathbf{X}) = [\exp(x_1) - x_2]^4 + 100(x_2 - x_3)^6 + \tan^4(x_3 - x_4) + x_1^8 + (x_4 - 1)^2$$

3. **ENGVALL-2d**

$$\mathcal{G}(\mathbf{X}) = x_1^4 + x_2^4 + 2x_1^2x_2^2 - 4x_1 + 3$$

4. **ENGVALL-3d**

$$\mathcal{G}(\mathbf{X}) = \sum_{i=1}^{5} g_i^2(\mathbf{X})$$

$$\begin{aligned}
g_1(\mathbf{X}) &= x_1^2 + x_2^2 + x_3^2 - 1 \\
g_2(\mathbf{X}) &= x_1^2 + x_2^2 + (x_3^2 - 2)^2 - 1 \\
g_3(\mathbf{X}) &= x_1 + x_2 + x_3 - 1 \\
g_4(\mathbf{X}) &= x_1 + x_2 - x_3 + 1 \\
g_5(\mathbf{X}) &= x_1^3 + 3x_2^2 + (5x_3 - x_1 + 1)^2 - 36
\end{aligned}$$

5. **POWELL**

$$\mathcal{G}(\mathbf{X}) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$$

6. **WHITE**

$$\mathcal{G}(\mathbf{X}) = 100(x_2 - x_1^3)^2 + (1 - x_1)^2$$

7. **WOOD**

$$\begin{aligned}
\mathcal{G}(\mathbf{X}) = {} & 100(x_2 - x_1^2)^2 + 90(x_4 - x_3^2)^2 + (1 - x_1)^2 + (1 - x_3)^2 + \\
& 10.1[(1 - x_2)^2 + (1 - x_4)^2] + 19.8(1 - x_2)(1 - x_4)
\end{aligned}$$

8. **ZANGWILL-2d**

$$\mathcal{G}(\mathbf{X}) = \frac{1}{15}(16x_1^2 + 16x_2^2 - 8x_1x_2 - 56x_1 - 256x_2 + 991)$$

9. **ZANGWILL-3d**

$$\mathcal{G}(\mathbf{X}) = (-x_1 + x_2 + x_3)^2 + (x_1 - x_2 + x_3)^2 + (x_1 + x_2 - x_3)^2$$

Appendix B

$\mathbf{X}^0$ and $\mathcal{G}^0$ are the initial point and the initial function value; $\mathbf{X}^*, \mathcal{G}^*$ the analytical ones; $\mathbf{X}^\wedge, \mathcal{G}^\wedge$ the ones found by the optimization methods; $N_{\mathcal{G}}$ the number of function evaluations.

Raw Results for Test Optimization Runs

Comparison of Various Optimization Methods via Common Test Functions (Test Runs Summary)								
	Init/Final	**Item**	**GaCD**	**BFGS**	**DFP**	**Fletcher**	**M-Powell**	**ES**
BEALE	X0: (10,-10)	X^	(6.3668,0.8117)	(6.3442,0.8142)	(2.9883,0.4959)	(6.3426,0.8155)	(-12.5481,1.0742)	(2.9861,0.4963)
	G0: 1.0114e8	G^	2.1017e-1	2.0256e-1	5.4101e-5	2.0166e-1	5.7019e-1	3.2355e-5
	X*: (3,0.5)	NG	350	558	674	391	6308	4641
	G*: 0	η,η_0	5.7119,1.0	3.5894,0.6284	4.1924,0.7339	5.1236,0.8970	0.3011,0.0527	0.6199,0.1085
CRAGG	X0:(5,5, 5,5)	X^	(-0.0102,0.9445, 0.9861,0.9992)	(0.2806,1.2365, 1.1070,1.0068)	(0.1747,1.1523, 1.0520,0.9976)	(0.1692,1.1400, 1.0608,0.9954)	(0.1313,1.1028, 1.0404,1.0008)	(0.1326,1.1128, 1.0583,0.9997)
	G0:4.2340e8	G^	5.3529e-6	7.1738e-4	1.1994e-4	6.9229e-5	1.1058e-5	1.5332e-5
	X*:(0,1,1,1)	NG	761	2312	622	1844	2531	6641
	G*:0	η,η_0	4.2052,0.9053	1.1723,0.2524	4.6451,1.0	1.5966,0.3437	1.2357,0.2660	0.4661,0.1003
ENG-VALL (2d)	X0:(10,10)	X^	(0.9949,0.0007)	(0.9951,-0.0024)	(1.0009,-0.0058)	(0.9964,0.0049)	(0.9950,-1.91e-6)	(0.9986,-0.0021)
	G0:39963.0	G^	1.5187e-4	1.5593e-4	7.2479e-5	1.2493e-4	1.4830e-4	2.0504e-5
	X*:(1,0)	NG	475	435	551	657	599	3121
	G*:0	η,η_0	4.0817,0.9170	4.4510,1.0	3.6530,0.8207	2.9807,0.6696	3.2407,0.7281	0.6854,0.1540
ENG-VALL (3d)	X0:(5,-5,-5)	X^	(-0.5195,-0.3588, 0.8870)	(1.5646,-2.8408, 0.0365)	(0.0972,-0.1124, 1.0180)	(0.3715,-0.9582, 1.0313)	**Failed**	(0.5986,-0.6582, -1.2334)
	G0:1025177.0	G^	2.0974	335.6351	7.1720e-3	3.0082		118.2426
	X*:(0,0,1)	NG	506	262	1488	1312		1201
	G*:0	η,η_0	2.5889,0.8453	3.0627,1.0	1.2619,0.4120	0.9710,0.3170	0.0,0.0	0.7550,0.2465
POWELL	X0:(5,-5, 5,-5)	X^	(0.0601,-0.0101, -0.0489,-0.0420)	(-0.0466,0.1047, -0.6756,0.1404)	(-0.1104,0.0116, -0.1024,-0.1044)	(0.0296,-0.0014, -0.1072,-0.1256)	(-0.8386,0.0628, -0.3853,-0.4822)	(-0.1806,0.0217, -0.1053,-0.1027)
	G0:153150.0	G^	3.0639e-3	8.8360	2.2369e-3	9.8247e-3	7.3507e-1	4.6038
	X*:(0,0,0,0)	NG	1339	316	750	1235	1438	3761
	G*:0	η,η_0	1.3239,0.4286	3.0887,1.0	2.4056,0.7788	1.3411,0.4342	0.8517,0.2757	0.2768,0.0896
WHITE	X0:(-5,-5)	X^	(-1.7196,-5.0435)	(-1.7088,-4.9851)	(-1.5010,-3.3602)	(-1.7192,-5.0762)	(-1.7146,-5.0458)	(-1.6492,-4.4786)
	G0:1440036.0	G^	7.5649	7.3401	6.3007	7.3969	7.3716	7.0226
	X*:(1,1)	NG	220	169	328	259	506	721
	G*:0	η,η_0	5.5258,0.7663	7.2111,1.0	3.7620,0.5217	4.7024,0.6521	2.4076,0.3339	1.6964,0.2352
WOOD	X0:(-3,-1, -3,-1)	X^	(1.0180,1.0407, 0.9709,0.9406)	(0.7458,0.5275, -0.7379,0.5027)	(1.3441,1.8087, 0.0541,0.0246)	(1.0740,1.1627, 0.8236,0.6697)	(0.8279,0.7390, 0.7766,0.6563)	(-1.0139,1.0307, -0.7823,0.6445)
	G0:19192.0	G^	7.9214e-3	12.7312	1.6522	3.5732e-1	4.2780	8.3979
	X*:(1,1,1,1)	NG	810	531	715	689	808	1121
	G*:0	η,η_0	1.8149,1.0	1.3782,0.7594	1.3091,0.7213	1.5808,0.8710	1.0407,0.5734	0.6899,0.3802
ZANG-WILL (2d)	X0:(0,0)	X^	(3.9936,8.9948)	(4.0049,9.0002)	(4.0049,8.9999)	(3.9950,8.9979)	(3.9379,8.9928)	(4.0337,9.0108)
	G0:66.0667	G^	-18.2000	-18.2000	-18.2000	-18.2000	-18.1961	-18.1989
	X*:(4,9)	NG	301	279	273	313	3128	1521
	G*:-18.2	η,η_0	*,0.9070	*,0.9785	*,1.0	*,0.8722	*,0.0873	*,0.1795
ZANG-WILL (3d)	X0:(100,-1, 2.5)	X^	(-0.0246,-0.0045, -0.0067)	(-0.0282,-0.0104, -0.0161)	(-0.0208,-0.0245, -0.0248)	(0.0246,0.0169, 0.0172)	(-0.0305,0.0005, -0.0171)	(-0.0004,-0.0009, 0.0010)
	G0:29726.75	G^	1.3991e-3	1.6628e-3	1.6782e-3	1.3041e-3	2.6743e-3	8.4492e-6
	X*:(0,0,0)	NG	830	878	916	1014	992	20241
	G*:0	η,η_0	2.0327,1.0	1.9019,0.9357	1.8220,0.8963	1.6708,0.8220	1.6355,0.8046	0.1086,0.0534