International Journal of Recent Engineering Science (IJRES),
ISSN: 2349-7157, *Volume 3 Issue 1 January to February 2016*

A UDFO non-monotone Wedge Trust-Region Algorithm with Modified Self-Correcting Geometry

Weili Zheng, Qinghua Zhou*
College of Mathematics and Information Science, Hebei University,
Baoding, 071002, Hebei Province, China

Abstract

In this paper, we propose a UDFO non-monotone wedge trust region algorithm with modified self-correcting geometry. This method can be projected to substantially decrease the need of geometry improving steps by exploiting a self-correcting property of the interpolation set geometry, and the design of this algorithm depends on a self-correction mechanism resulting from the combination of the non-monotone wedge trust-region framework with the polynomial interpolation setting. The global convergence of this algorithm is proved under some mild conditions.

Keywords: *trust-region, non-monotone wedge technique, self-correcting geometry, unconstrained derivative-free optimization.*

I. INTRODUCTION

We consider the unconstrained optimization problem:

$$\min f(x), \quad x \in R^n, \qquad (1)$$

Where f is a nonlinear function from R^n into R. We first recall the general trust-region framework where derivatives of f are available before turning to the derivative-free case. At each iteration of an iterative trust-region method, a model of the form

$$m_k(x_k + s) = f(x_k) + g_k^T s + \frac{1}{2} s^T H_k s$$

(2)

(Where g_k and H_k are the function's gradient and Hessian) is minimized inside the following trust region

$$B_\infty(x_k, \Delta_k) = \left\{ x \in R^n \,\middle|\, \|x - x_k\|_\infty \le \Delta_k \right\}$$

(3)

Where $\|\cdot\|_\infty$ denotes the infinity norm.

In our derivative-free context, the model (2) is determined by the interpolation set Y_k, and the set Y_k of interpolation points must satisfy the conditions $m_k(y) = f(y)$ for all $y \in Y_k$. This derivative-free methods date back to the algorithm of Winfield [1, 2], which is one of the pioneering works in trust region methods. Excellent reviews on trust region algorithms for optimization without derivatives are given in [3, 4] and [5]. In order to keep the interpolation set from becoming degenerate, many methods use explicit "geometry-improving" steps, and then Scheinberg and Toint [6] exploited a self-correcting property of the interpolation set geometry with the trust region algorithm. Marazzi and Nocedal [7] proposed another strategy in which a "wedge constraint" is imposed to the trust region sub-problem so that the geometry of the interpolation set will not be destroyed after the trial point is included into the set. Considering the effectiveness of non-monotone strategy when coping with the problems, Zhang [8] proposed a new non-monotone method for trust-region algorithm. In this paper, we presented a hybrid algorithm which combined a non-monotone wedge trust region method and self-correcting geometry technique.

The paper is organized as follows. In section 2 we recommend some preliminaries about interpolation schemes, and introduced the wedge and non-monotone strategies in detail. We design the new non-monotone wedge trust-region algorithm with self-correcting geometry strategy fanaily. In section 3, we prove the global convergence of our algorithm. In section 4, we make some conclusions.

II. A UDFO NON-MONOTONE TRUST-REGION ALGORITHM WITH SELF-CORRECTING GEOMETRY

A. The Interpolation Models

Recently, many researchers proposed some different techniques to construct the interpolation model [9, 10, 11]. In the following, let us describe the definition and lemmas of the interpolation model minutely. Consider p_n^d, the space of polynomials of degree $\le d$ in R^n, and let $p_1 \square p+1$ be the dimension of the space.

Definition 2.1.[12, 13] Given a set of interpolation $Y = \{y^0, y^1, \cdots, y^p\} \subset R^n$, a basis of $p_1 = p+1$ polynomials $l_j(x)(\mathrm{j}=0,\cdots,\mathrm{p}) \in \mathrm{p}_n^d$ (4)

is called a basis of Lagrange polynomials if

$$l_j(\mathrm{y}^i) = \delta_{ij} = \begin{cases} 1, \mathrm{i} = \mathrm{j}, \\ 0, i \ne j. \end{cases} \quad (5)$$

If Y is poised, then Lagrange polynomials exist and are unique. Moreover, they have a lot of useful properties. Particularly, we are interested in the crucial fact that, if $m(x)$ interpolates $f(x)$ at the points in Y, then for all x,

$$m(x) = \sum_{j=0}^{p} f(\mathrm{y}^j) l_j(x).$$

It can also be shown that $\sum_{j=0}^{p} l_j(x) = 1, \forall x \in R^n$.

Lemma 2.1.[14, 15] Given the sphere $B(x, \sqrt{n}\Delta) \overset{def}{=} \{v \in R^n \,\|\, \|v - x\|_2 \le \sqrt{n}\Delta\}$, a poised set $Y \subset B(x, \sqrt{n}\Delta)$ and its associated basis of Lagrange polynomials $\{l_i(x)\}_{i=1}^{p}$, there exists constants $k_{ef} > 0$ and $k_{eg} > 0$ such that, for any interpolation polynomial $m(\mathrm{x})$ of degree one or higher of the form (6) and any point $y \in B(x, \sqrt{n}\Delta)$, one has

$$\|f(\mathrm{y}) - \mathrm{m}(\mathrm{y})\| \le k_{ef} \sum_{i=0}^{p} \|y^i - y\|_2^2 \, |l_i(\mathrm{y})| \quad \text{and}$$

$$\|\nabla f(\mathrm{y}) - \nabla\, \mathrm{m}(\mathrm{y})\|_2 \le k_{eg} \Lambda \Delta, \quad (7)$$

where $\Lambda = \max_{i=1,\cdots,p} \max_{y \in \mathrm{B}_2(x, \sqrt{n}\Delta)} |l_i(\mathrm{y})|$.

Lemma 2.2. For any given $\Lambda > 1$, a closed ball B, and a fixed polynomial basis ϕ, algorithm 2.1 terminates with a Λ-poised set Y after a finite number of iterations where the number of steps depends on Λ and ϕ.

B. Non-Monotone Technology

Currently, non-monotone technique has been studied by many scholars [8, 16, 17, 18]). Toint pointed out that the non-monotone technique can enhance the possibility of finding a global optimization. And it can improve the rate of convergence in cases where a monotone scheme is forced to creep along the bottom of a narrow curved valley. In this paper, we can use the ratio [8]

$$\rho_k = \frac{C_k - f(x_k + s_k)}{m_k(x_k) - m_k(x_k + s_k)} \quad (8)$$

where

$$C_k = \begin{cases} f(x_k), k = 0, \\ \frac{\eta_{k-1}Q_{k-1}C_{k-1} + f(x_k)}{Q_k}, k \le 1, \end{cases} \qquad Q_k = \begin{cases} 1, k = 0, \\ \eta_{k-1}Q_{k-1} + 1, k \ge 1, \end{cases} \quad (9)$$

$\eta_{\min} \in [0,1)$, $\eta_{\max} \in [\eta_{\min}, 1]$ and $\eta_{k-1} \in [\eta_{\min}, \eta_{\max}]$, are two given constants.

C. Wedge Trust Region Method.

The wedge trust region method was proposed by Marazzi and Nocedal [7]. The wedge constraint is added to the trust region sub-problem, and we have

$$\min_s m_k(x_k + s) \quad (10)$$

$$s.t. \|s\| \le \Delta_k \quad (11)$$

$$s \notin W_k , \quad (12)$$

Where W_k is a set which contains the "taboo region" area, and the purpose is to avoid the new point falling into it. As for solving the wedge trust region sub-problem (10)-(12), we usually first solve the standard trust region sub-problem without the wedge constraint and get a solution s_k^e at the k-th iteration. If s_k^e satisfies the wedge constraint, we set $s_k = s_k^e$ as the trail step. Otherwise, the wedge constraint is active. By rotating s_k, we find a vector satisfying the wedge constraint. Then we set the trail point $x_k^+ = x_k + s_k$. One point must be mentioned is that there is a measure to make sure the sufficient descent of the model function in the rotating process, i.e.,

$$m_k(x_k) - m_k(x_k^+) \ge \frac{1}{2}\left[m_k(x_k) - m_k(x_k + s_k^e)\right] \quad (13)$$

Algorithm 2.1 UDFO non-monotone trust region algorithm with self-correcting geometry

Step 0: initialization.

An initial trust-region radius Δ_0, initial accuracy threshold ε_0 are given. An initial poised interpolation set Y_0 that contains the starting point x_0 is known. An interpolation model m_0 around x_0 and associates Lagrange polynomial $\{l_{0,j}\}_{j=1}^{p}$ are computed. Constants η_1, $\Delta_{switch} \in (0,1)$, $\eta \in (0,1), 0 < \gamma_1 < \gamma_2 < 1, \mu \in (0,1), \theta > 0, \beta \ge 1, \varepsilon \ge 0, \Lambda > 1$, $p_{\max} \ge n+1$ are also given. Choose $v_0 \ne x_0$

Where v_i is variable introduced to keep track if the model at x_k is known to be well poised. Set $k = 0$ and $i = 0$.

Step 1: criticality test.

Step 1a: define $\hat{m}_i = m_k$.

Step 1b: if $\|\nabla_x \hat{m}_i(x_k)\| < \varepsilon_i$, set $\varepsilon_{i+1} = \mu\|\nabla_x \hat{m}(x_k)\|$, compute a Λ -poised model $\hat{m}_{i+1}$ in $B(x_k, \varepsilon_{i+1})$ and increment i by one. If $\|\nabla_x \hat{m}_i(x_k)\| < \varepsilon$, then return x_k, otherwise start step 1b again.

Step 1c: set $m_k = \hat{m}_i, \Delta_k = \theta\|\nabla_x m_k(x_k)\|$ and define $v_i = x_k$ (what indicates that the model at x_k is well poised and steps 4b and 4c need not to be visited in an unsuccessful iterate) if a new model has been computed.

Step 2: compute a trial point.

Solve the wedge trust region sub-problem (10) for getting s_k, so the trial point $x_k^+ = x_k + s_k$.

Step 3: evaluate the objective function at the trial point.

Compute $f(x_k^+)$ and ρ_k from (8).

Step 4: define the next iterate.

Step 4a: Augment interpolation set ($p_k < p_{max}$). If $p_k < p_{max}$, then: define $Y_{k+1} = Y_k \cup \{x_k^+\}$. If $\rho_k \geq \eta_1$, then define $x_{k+1} = x_k^+$ and choose $\Delta_{k+1} \geq \Delta_k$. If $\rho_k < \eta_1$, define $x_{k+1} = x_k$ and if $\Delta_k > \Delta_{switch}$, set $\Delta_{k+1} \in [\gamma_1 \Delta_k, \gamma_2 \Delta_k]$, otherwise $\Delta_{k+1} = \Delta_k$.

Step 4b: Successful iteration. If $\rho_k \geq \eta_1$ and $p_k = p_{max}$, define $x_{k+1} = x_k^+$, choose $\Delta_{k+1} \geq \Delta_k$ and define $Y_{k+1} = Y_k \setminus \{x_k^+\} \cup \{y_{k,\tau}\}$, where τ is the index j of any point $y_{k,j}$ in Y_k, for instance, such that $\tau = \arg\max_j \|y_{k,j} - x_k^+\|_\infty^2 \, |l_{k,j}(x_k^+)|$.

Step 4c: replace a far interpolation point. If $\rho_k < \eta_1$, $p_k = p_{max}$, either $x_k \neq \upsilon_i$ or $\Delta_k \leq \varepsilon_i$, and the set

$$F_k \overset{\det}{=} \{ y_{k,j} \in Y_k \ such\ that \|y_{k,j} - x_k\| > \beta \Delta_k \ and\ l_{k,j}(x_k^+) \neq 0 \}$$

is non-empty, set $x_{k+1} = x_k$, $\Delta_{k+1} \in [\gamma_1 \Delta_k, \gamma_2 \Delta_k]$ and define $Y_{k+1} = Y_k \setminus \{x_k^+\} \cup \{y_{k,\tau}\}$ where τ is the index j of any point $y_{k,j}$ in F_k, for instance, such that

$$\tau = \arg\max_j \|y_{k,j} - x_k^+\|_\infty^2 \, |l_{k,j}(x_k^+)|.$$

Step 4d: replace a close interpolation point. If $\rho_k < \eta_1$, $p_k = p_{max}$, either $x_k \neq \upsilon_i$ or $\Delta_k \leq \varepsilon_i$, the set $F_k = \phi$ And

$$E_k \overset{\det}{=} \{ y_{k,j} \in Y_k \setminus \{x_k\} \ such\ that \|y_{k,j} - x_k\| \leq \beta \Delta_k \ and\ l_{k,j}(x_k^+) > \Lambda \}$$

is non-empty, then set $x_{k+1} = x_k$, $\Delta_{k+1} \in [\gamma_1 \Delta_k, \gamma_2 \Delta_k]$ and define $Y_{k+1} = Y_k \setminus \{x_k^+\} \cup \{y_{k,\tau}\}$ where τ is the index j of any point $y_{k,j}$ in E_k, for instance, such that

$$\tau = \arg\max_j \|y_{k,j} - x_k^+\|_\infty^2 \, |l_{k,j}(x_k^+)|.$$

Step 4e: reduce trust region radius. If $\rho_k < \eta_1$, $p_k = p_{max}$, and either $[x_k = v_i \ and\ \Delta_k > \varepsilon_i]$ or $F_k \cup E_k = \phi$, then set $x_{k+1} = x_k$, $\Delta_{k+1} \in [\gamma_2 \Delta_k, \gamma_1 \Delta_k]$ and define $Y_{k+1} = Y_k$.

Step 5: update the model and Lagrange polynomial.

If $Y_{k+1} \neq Y_k$, compute the interpolation model m_{k+1} around x_{k+1} using Y_{k+1} and the associated Lagrange polynomials $\{l_{k+1,j}\}_{j=0}^{p}$. Increment k by one and go to step 1.

III. GLOBAL CONVERGENCE

In this section, we prove the convergence results of this algorithm. It made use of the self-correcting property presented above and thus depended on the convergence results obtained by [6], and the addition of non-monotone method can be splendidly proved the convergence of new algorithm.

First, the assumptions are stated.

A1: the objective function f is continuously differentiable in an open set V containing all iterates generated by the algorithm, and its gradient $\nabla_x f$ is Lipschitz continuous in V with constant L ;

A2: there exists a constant k_{low} such that $f(x) \ge \mathrm{k}_{low}$ for every $x \in V$;

A3: there exists a constant $k_H \ge L$ such that $1 + \|H_k\| \le k_H$ for every $k \ge 0$;

A4: $|Y_k| \ge n+1$ for every $k \ge 0$.

Note that A1 only assumes the existence of first derivatives, not that they can be computed.

Lemma2.3. [6] Assume that, for some numbers $\{\alpha_i\}_{i=0}^{t}$ with $\sigma_{abs} \overset{def}{=} \sum_{i=0}^{t} |\alpha_i| > 2\sum_{i=0}^{t} \alpha_i \overset{def}{=} 2\sigma > 0$

If one defines $i^* = \arg\max_{i=0,\dots,t} |\alpha_i|$ *and* $j^* = \arg\max_{j=0,\dots,t,\ j \ne i^*} |\alpha_j|$ then

$$|\alpha_{j^*}| \ge \frac{\sigma_{abs} - 2\sigma}{2t} .$$

Lemma2.4. suppose that the sequence $\{x_k\}$ is generated by algorithm 2.1. Then the inequality $f(x_{k+1}) \le \mathrm{C}_{k+1} \le C_k, \forall k$ hold, where C_k is defined in (9).

Proof. We prove $C_k \le f(x_0)$ by induction. Obviously, when $j = 0$, we have $C_0 = f(x_0)$. Assume that $C_j \le f(x_0)$ holds when $j = 1, \dots, k-1$. When $j = k$, from (9). We have

$$C_k = \frac{\eta_{k-1} Q_{ik-1} C_{k-1} + f(\mathrm{x}_k)}{Q_k} \le \frac{\eta_{k-1} Q_{k-1} f(x_0) + f(x_0)}{Q_k} = f(x_0) , \quad (14)$$

then we obtain that $f(x_{k+1}) \le C_{k+1} \le C_k \le f(x_0)$.

We complete the proof.

Lemma2.5. At the k-th iteration, the solution of the wedge trust region sub-problem (10)-(12) satisfies the fraction of Canchy decrease condition:

$$m_k(x_k) - m_k(x_k^+) \ge k_c \|g_k\| \min\left\{\Delta_k, \frac{\|g_k\|}{\|G_k\|}\right\}, \quad (15)$$

Where $k_c \in (0, \frac{1}{4})$ is a constant.

Proof. From Lemma 6.1.3 in [17], we know that if s_k^e is the exact solution of (10)-(11), then

$$m_k(x_k) - m_k(x_k + \mathrm{s}_k^e) \ge \frac{1}{2} \|g_k\| \min\left\{\Delta_k, \frac{\|g_k\|}{\|G_k\|}\right\} . \quad (16)$$

As the process of solving the sub-problem in the wedge trust region methods, under the safety management (16), we can deduce from (13) that there exists a constant $k_c \in (0, \frac{1}{4})$ such that

$$m_k(x_k) - m_k(x_k + \mathrm{s}_k) \ge \frac{1}{2}\left[m_k(x_k) - m_k(x_k + \mathrm{s}_k^e)\right] \ge \frac{1}{4}\|g_k\| \min\left\{\Delta_k, \frac{\|g_k\|}{\|G_k\|}\right\} \ge k_c \|g_k\| \min\left\{\Delta_k, \frac{\|g_k\|}{\|G_k\|}\right\} .$$

Lemma2.6. [6] Suppose that A1, A3 and A4 hold and assume that, for some $k_0 \ge 0$ and all $k \ge k_0, \|g_k\| \ge k_g$ for some $k_g > 0$. Then there exist a contain $k_\Delta > 0$ such that, for all $k \ge k_0$, $\Delta_k \ge k_\Delta$.

Lemma2.6. suppose that A1, A2 andA4 hold and that there is a finite number of successful iterations. Then $\lim_{k\to\infty}\inf \|g_k\| = 0$.

Proof. When the $\rho_k < \eta_1$, we can obtain that every iteration is unsuccessful, $x_k = x_*$ for some x_* and all k large. Then by Lemma 2.6, we have that $\Delta_k > k_\Delta > 0$ on all iterations. So the sequence $\{x_k\}$ is non-increasing and bounded below, and therefore convergent. Let $\Delta_\infty \overset{def}{=} \lim_{k\to\infty} \Delta_k \ge k_0$.

In the type 4a, $\Delta_{switch} > k_\Delta$, so Δ_k is decreased.

In the type 4c, the p_{max} can be necessary to ensure that all interpolation points belong to $B(x_*, \beta\Delta_\infty)$.

In the type 4d, the new trial point replaces some interpolation point from the set E_k, for all k large enough, the trial point x_k^+ replaces a previous interpolation point $y_{k,j}$ such that $|l_{k,j}(x_k^+)| \ge \Lambda$.

In the type 4e, the iterations can not happen infinitely often because Δ_k is bounded below by Δ_∞ and $\gamma_2 < 1$.

Lemma2.7. [6] Suppose that A1-A4 holds and that the number of successful iterations is infinite, then $\lim_{k\to\infty}\inf \|g_k\| = 0$.

Lemma2.8. Suppose that A1, A3 and A4 hold. Then, for any constant $\Lambda > 1$, if iteration k is unsuccessful, $p = p_k = p_{max}, F_k = \phi$ and

$$\Delta_k \le \min\left[\frac{1}{k_H}, \frac{(1-\eta_1)k_c}{2k_{ef}(\beta+1)^2(p_k\Lambda+1)}\right]\|g_k\| \overset{def}{=} k_\Lambda \|g_k\|, \quad (17)$$

then $E_k \ne \phi$.

Proof . As we assume that the maximum number of points in the set $p_k = p_{max}$ is reached, and assume iteration k is unsuccessful, which is to say that $\frac{C_k - f(x_k + s_k)}{m_k(x_k) - m_k(x_k + s_k)} < \eta$. Lemma2.4. guarantee that $C_k \ge f(x_k) \ge f(x_{k-1})$, then we have $\frac{f(x_k) - f(x_k + s_k)}{m_k(x_k) - m_k(x_k + s_k)} < \eta$. Now, because of the identity $f(x_k) = m_k(x_k)$, this in turn means that

$$|f(x_k + s_k) - m_k(x_k + s_k)| > (1-\eta)|m_k(x_k) - m_k(x_k + s_k)|. \quad (18)$$

We may now deduce from Lemma2.1. that

$$|f(x_k + s_k) - m_k(x_k + s_k)| \le k_{ef} \sum_{j=0}^{p} \|y_{k,j} - (x_k + s_k)\|^2 |l_{k,j}(x_k + s_k)|. \quad (19)$$

Observe now that $F_k = \phi$ ensures that $\|y_{k,j} - x_k\| \le \beta\Delta_k$ whenever $l_{k,j}(x_k + s_k) \ne 0$.

This observation and the trust-region bound then imply that

$\|y_{k,j} - x_k - s_k\| \le \|y_{k,j} - x_k\| + \|s_k\| \le (\beta+1)\Delta_k$ for j such that $l_{k,j}(x_k + s_k) \ne 0$, so that (18) and (19) then imply that

$$|f(x_k + s_k) - m_k(x_k + s_k)| \le k_{ef}(\beta+1)^2 \Delta_k^2 \sum_{j=0}^{p} |l_{k,j}(x_k + s_k)|. \quad (20)$$

On the other hand, the Cauchy condition [6] and (17) together imply that

$\left|m_k(x_k) - m_k(x_k + s_k)\right| \geq k_c \left\|g_k\right\|\Delta_k$, hence (20) gives that

$$\sum_{j=0}^{p} \left|l_{k,j}(x_k + s_k)\right| \geq \frac{(1-\eta)\, k_c \left\|g_k\right\|}{k_{ef}(\beta+1)^2 \Delta_k} . \quad (21)$$

As a consequence, we have, using (21) and (17) successively, that $\sum_{j=0}^{p} \left|l_{k,j}(x_k + s_k)\right| \geq s(p\Lambda + 1)$. Moreover, (6) also implies that $\sum_{j=0}^{p} l_{k,j}(x_k + s_k) = 1$. We may use this equality, and Lemma 2.1. to deduce that, if $j^* = \arg\max_{j=0,\dots,p} \left|l_{k,j}(x_k + s_k)\right|$, then $\left|l_{k,\tau}(x_k + s_k)\right| \geq \Lambda$ for $\tau = \arg\max_{\substack{j=0,\dots,p \\ j\neq j^*}} \left|l_{k,j}(x_k + s_k)\right|$, which together with (19) and (20), implies $E_k \neq \phi$.

IV. CONCLUSIONS

In this paper, we proposed a non-monotone self-correcting wedge trust-region method for unconstrained derivative-free optimization and analyzed the properties of the new algorithm. It is efficient for solving these unconstrained derivative-free optimization problems. The global convergence result of the new proposed method is proved under some mild conditions. In the near future, we will learn and seek more efficient non-monotone strategies for the wedge trust region methods to settle the unconstrained derivative-free optimization problem.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (61473111) and the Natural Science Foundation of Hebei Province (Grant No. A2014201003, A2014201100).

REFERENCES

[1] S. M. Wild. Derivative-free Optimization Algorithm for Computationally Expensive Functions. Phd thesis, Cornell University, Ithaca, NY, USB (2008).

[2] S. M. Wild. Mnh: a derivative-free optimization algorithm using minimal norm Hessians. In Tenth Copper Mountain Conference on Iterative Methods,(2008).

[3] R. M. Lewis and V. Torczon. A globally convergence augmented lagrangian pattern search algorithm for optimization with general constraints and simple bounds. SIAM J. on Optimization, 12:1075-1089,(2002).

[4] M. Marazzi and J. Nocedal: Wedge trust region methods for derivative free optimization, Math. Program, Series A, 91, p. 289-305 (2002).

[5] B. Colson and ph. L. Toint. A derivative-free algorithm for sparse unconstrained optimization problems. In A. H. Siddiqi and M. Kocvara, editors, Trends in Industrial and Applied Mathematics, Applied Optimization, pages 131-149, Dordrecht, The Netherlands,(2002).

[6] K. Scheinberg and Ph. L. TOINT, Self-correcting geometry in model-based algorithms for derivative-free unconstrained optimization, SIAM Journal on Optimization, 20: 3512-3532, (2010).

[7] M. Marazzi and J. Nocedal, Wedge trust region methods for derivative free optimization, Mathematical Programming, 91: 289-305,(2002).

[8] H.C. Zhang, W.W. Hager, A nonmonotone line search technique and its application to unconstrained optimization, SIAM J. Optim. 14(4) 1043-1056, (2004).

[9] P. G. Ciarlet and P. A. Raviart, General Lagrange and Hermite interpolation in R^n with applications to finite element methods, Archive for Rational Mechanics and Analysis, 46: 178-199, (1972).

[10] M. J. D. Powell, On the Lagrange function of quadratic models that are defined by interpolation, Optimization Methods, and Software, 16: 289-309,(2001).

[11] R. Oeuvray, Trust-Region Methods Based on Radial Basis Function with Application to Biomedical Imaging, PhD thesis, Institut de Mathematiques, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland, (2005).

[12] A. R. Conn, K. Scheinberg and L. N. Vicente, Introduction to Derivative-Free Optimization, MPS-SIAM, Series on

Optimization, SIAM, Philadephia, PA, USA, (2008).

[13] W. Sun, Q. K. Du and J. R. Chen Computational Methods, Science Press, Beijing, (2007).

[14] P. G. Ciarlet and P. A. Raviart. General Lagrange and Hermite interpolation in R^n with applications to finite element methods. Archive for Rational Mechanics and Analysis, 46(3):177-199, (1972).

[15] A. R. Conn, K. Scheinberg, and L,N, Vicente. Introduction to Derivative-free Optimization. MPS-SIAM Series on Optimization. SIAM,philadephia, PA, USA,(2008).

[16] N.Z. Gu, J. T. Mo, Incorporating nonmonotone strategies into the trust region for unconstrained optimization, Computers and Mathmatics with Applications 55: 2158-2172, (2008).

[17] Sun W, Yuan Y. Optimization Theory and Methods: Nonlinear Programming. New York: Spinger, (2006).

[18] Andrei N. Scaled conjugate gradient algorithms for unconstrained optimization. Comput Optim Appl, 38: 401-416,(2007).