Journal of Global Optimization manuscript No.
(will be inserted by the editor)

Global convergence and the Powell singular function

Trond Steihaug · **Sara Suleiman**[⋆]

Received: date / Accepted: date

Abstract The Powell singular function was introduced 1962 by M.J.D. Powell as an unconstrained optimization problem. The function is also used as nonlinear least squares problem and system of nonlinear equations. The function is a classic test function included in collections of test problems in optimization as well as an example problem in text books. In the global optimization literature the function is stated as a difficult test case. The function is convex and the Hessian has a double singularity at the solution. In this paper we consider Newton's method and methods in Halley class and we discuss the relationship between these methods on the Powell Singular Function. We show that these methods have global but linear rate of convergence. The function is in a subclass of unary functions and results for Newton's method and methods in the Halley class can be extended to this class. Newton's method is often made globally convergent by introducing a line search. We show that a full Newton step will satisfy many of standard step length rules and that exact line searches will yield slightly faster linear rate of convergence than Newton's method. We illustrate some of these properties with numerical experiments.

Keywords System of nonlinear equations · Global optimization · Higher order methods

Mathematics Subject Classification (2000) 49M15 · 58C15 · 65H10 · 90-08 · 90C06 · 90C30

1 Introduction

In 1962 M.J.D. Powell [37] introduced the objective function

$$f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4, \qquad (1)$$

⋆ Corresponding author

T. Steihaug · S. Suleiman
Department of Informatics, University of Bergen, P.O. Box 7803, N-5020 Bergen, Norway.
Fax: +47-55 58 41 99
E-mail: Sara.Suleiman@ii.uib.no

T. Steihaug
E-mail: Trond.Steihaug@ii.uib.no

to test an unconstrained optimization algorithm. The function goes under the names Powell Quartic Function or Powell Singular Function (PSF). Standard starting point is $x_0 = (3, -1, 0, 1)^T$ and the Hessian matrix at the standard starting point is nonsingular. PSF is convex and the unique unconstrained minimizer is $x^* = (0, 0, 0, 0)^T$. Moreover, the Hessian matrix at x^* is doubly singular thus PSF is a severe test problem [4, Page 141] and [16] and particularly for quadratically convergent methods [45].

PSF is a standard reference test for derivative free optimization methods [32,34], for heuristic methods [42] as well as global optimization algorithms [7,31,36,48]. The test function is also a standard reference in textbooks [4,9,38,46] and it is included in collections of test problems [3,29,39,41]. The function is extended to variably dimensioned problems [6,23,25,43]. PSF represents typical situations that are encountered in testing and is used to illustrate the behavior of the algorithms including Newton's method [28].

PSF is an easy problem for most global optimization methods that rely on heuristic techniques, see for example [21,22]. These methods converge to the global minimizer. This is also the case for gradient based methods like quasi-Newton methods, see [13,12] and the references therein. However, for quasi-Newton methods the super linear rate of convergence will be lost. For more complete collections of more challenging test problems in global optimization see [1] for unconstrained problems, and [17] for constrained problems.

PSF is a type of optimization problem where a linear transformation of the variables makes the optimization problem separable. A function $g(y) : \mathbb{R}^n \to \mathbb{R}$ is separable if

$$g(y) = \sum_{i=1}^{n} g_i(y_i), \tag{2}$$

where $g_i : \mathbb{R} \to \mathbb{R}$ and y_i is component i of y.

Let $S \in \mathbb{R}^{n \times n}$ be a nonsingular matrix and let g be a separable function. Define the function $f(x) = g(Sx)$. PSF can be written as $f(x) = g(Sx)$, where $g(y_1, y_2, y_3, y_4) = y_1^2 + 5y_2^2 + y_3^4 + 10y_4^4$ and S is 4×4 nonsingular matrix. Since the function g is separable, and Newton's method is invariant under a linear transformation of the variables [15, Section 3.3], Newton's method will behave like the one dimensional counterpart on $\min_\xi g_i(\xi)$. However, the rate of convergence will be determined by the slowest rate of convergence of the components [14]. For one dimensional problems Newton's method has a global convergence[1] under reasonable assumptions like the Fourier conditions [35, Chapter 9] or conditions in [47].

Let $a_i \in \mathbb{R}^n$, $i = 1, \ldots, m$ be a constant vector and let $g(y) : \mathbb{R}^m \to \mathbb{R}$ be a separable function (2). A unary function [26] can be written on the form $f(x) = \sum_{i=1}^{m} g_i(y_i(x)),\ x \in \mathbb{R}^n$, where $y_i(x) = a_i^T x$. Let S be the $m \times n$ matrix where row i is a_i^T then $f(x) = g(Sx)$. PSF belongs to a class of unary functions with $m = n$ and a nonsingular matrix S. Other well known examples of functions in this class are Rosenbrock Cliff function [40, Chapter 4, Page 71], extended PSF [43], extended Rosenbrock Cliff function [18] and the variably dimensional function [18].

PSF is also used as a test function for nonlinear least squares [30] where $f(x) = F(x)^T F(x)$ or in a nonlinear system of equations $F(x) = 0$ where the nonlinear func-

[1] There is no assumption on x_0 to be sufficiently close to the solution.

tion $F : \mathbb{R}^4 \to \mathbb{R}^4$ is given by [5]

$$F(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \\ f_4(x) \end{pmatrix} \equiv \begin{pmatrix} x_1 + 10x_2 \\ \sqrt{5}(x_3 - x_4) \\ (x_2 - 2x_3)^2 \\ \sqrt{10}(x_1 - x_4)^2 \end{pmatrix}. \tag{3}$$

The Jacobian matrix of F at the solution x^* is singular, which makes the function a difficult test function [24]. Test functions for nonlinear least squares like the Cuyt and Cruyssen function [8] and Problem 201 in [39] are further examples of functions in the subclass of unary functions where we find PSF.

The classical iterative method to solve $\nabla f(x) = 0$, where $f : \mathbb{R}^n \to \mathbb{R}$ is at least three times continuously differentiable, is Newton's method. For a given iterate x^k the correction is given by

$$s_{(1)}^k = -\nabla^2 f(x^k)^{-1} \nabla f(x^k), \tag{4}$$

provided that the Hessian matrix at x^k is nonsingular and the new iterate is $x^{k+1} = x^k + s_{(1)}^k,\ k \geq 0$. If $\nabla^2 f(x^*)$ is nonsingular and $\nabla f(x^*) = 0$, it is well known that if the starting point x^0 is sufficiently close to the solution x^*, then Newton's method converges with a quadratic rate of convergence. The Halley class of methods [20] have local and third order rate of convergence. For a given value of α and a given iterate x^k, the new iterate of a method in the Halley class is given by $x^{k+1} = x^k + s_{(1)}^k + s_{(2)}^k,\ k \geq 0$, where $s_{(1)}^k$ is the Newton step (4) and $s_{(2)}^k$ is given by

$$s_{(2)}^k = -\frac{1}{2} \left(\nabla^2 f(x^k) + \alpha \nabla^3 f(x^k) s_{(1)}^k \right)^{-1} \nabla^3 f(x^k) s_{(1)}^k s_{(1)}^k. \tag{5}$$

For specific values of α in Halley class we get some well known methods. Chebyshev's method, Halley's method and super-Halley's method are obtained when $\alpha = 0, \frac{1}{2}$ and 1 respectively. Further, each method in the Halley class is invariant under linear transformation [44]. As for Newton's method on one dimensional problems, the Chebyshev and Halley methods have global convergence under reasonable assumptions [2,27].

A global strategy like a line search using the Newton direction and the PSF has been discussed in [10,16,19]. The new iterate using line search is $x^{k+1} = x^k + \lambda_k s_{(1)}^k,\ k \geq 0$ where λ_k is called the step length. For the PSF it is often observed that the full Newton step $\lambda_k = 1$ will be acceptable.

The paper is organized as follows. In section 2, we discuss solving PSF as an unconstrained optimization problem. We show that Newton's method is globally convergent with a linear rate of convergence. Further, we show that methods in the Halley class are globally convergent with a linear rate for α in an interval. In section 3, we take a closer look at Newton's method coupled with line search. We show that the step length $\lambda_k = 1$ and the Newton direction $s_1^{(k)}$ will satisfy standard line search conditions. Furthermore, we show that the Newton direction with exact step length has a faster convergence than with the step length $\lambda_k = 1$. This violates the folklore that taking the Newton step is the best choice close to a solution. Finally, the observation that alternating the step length between 1 and 3 reduces the number of iterations is proved by showing finite termination after two iterations. In section 4, we consider the Brent formulation (3) of PSF as system of nonlinear equations and we show global and linear rate of convergence for. In the final section 5, some of the properties are illustrated with numerical experiments.

2 PSF as unconstrained optimization problem

For the unconstrained optimization problem where the objective function is smooth, we are searching for a stationary point $\nabla f(x) = 0$. Since (1) is convex, any stationary point will be a global minimizer. The determinant of the Hessian matrix $\nabla^2 f(x)$ is given by $12700800(x_1 - x_4)^2(x_2 - 2x_3)^2$. Therefore, the matrix $\nabla^2 f(x)$ is nonsingular if and only if $f_3(x) \neq 0$ and $f_4(x) \neq 0$ where $f_i, i = 1, 2, 3, 4$ are given in (3). For simplification, we will remove the iteration index and use that x is the current iterate and x^+ is the next iterate. The Newton step $s_{(1)}$ is given by (4) and can be written as

$$s_{(1)} = -\left(I - \frac{4}{3}N\right)x, \tag{6}$$

and the matrix N has eigenvalue decomposition $N = VDV^{-1}$ where

$$N = \frac{1}{42}\begin{bmatrix} 20 & -10 & 20 & -20 \\ -2 & 1 & -2 & 2 \\ -1 & -10 & 20 & 1 \\ -1 & -10 & 20 & 1 \end{bmatrix}, V = \begin{bmatrix} 0 & 1 & -10 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \tag{7}$$

and $D = \text{diag}(0, 0, 1/2, 1/2)$. The Newton iterate, $x^+ = x + s_{(1)}$, is given by

$$x^+ = \frac{4}{3}Nx. \tag{8}$$

Using the definition (3) at the iterate x^+ in (8), we have $f_i(x^+) = \frac{4}{9}f_i(x)$ for $i = 3, 4$, and then $\det(\nabla^2 f(x^+)) \neq 0$ if $\det(\nabla^2 f(x)) \neq 0$.

A Halley class iterate for a given α is defined by $x^+ = x + s_{(1)} + s_{(2)}$ where $s_{(1)}$ is the Newton step and $s_{(2)}$ is given by (5). In terms of the matrix N, $s_{(2)}$ is given by

$$s_{(2)} = \frac{2}{3(2\alpha - 3)}N x.$$

The determinant of $\nabla^2 f(x) + \alpha\nabla^3 f(x)s_{(1)}$ is $1411200(2\alpha - 3)^2(x_1 - x_4)^2(x_2 - 2x_3)^2$, which is nonsingular if $f_3(x) \neq 0$ and $f_4(x) \neq 0$, provided that $\alpha \neq 3/2$. Therefore, Halley class iterate can be written as

$$x^+ = \frac{2(4\alpha - 5)}{3(2\alpha - 3)}Nx. \tag{9}$$

From the definition (3) and (9), we have $f_i(x^+) = \frac{(4\alpha-5)^2}{9(2\alpha-3)^2}f_i(x)$ for $i = 3, 4$. Therefore, $\det(\nabla^2 f(x^+)) \neq 0$ if $\det(\nabla^2 f(x)) \neq 0$ provided $\alpha \neq 5/4$ and $\alpha \neq 3/2$.

We need a technical Lemma.

Lemma 1 *Let $|\beta| < 2$, and the matrix N be defined by (7). Consider the sequence of iterates $\{x^k\}$ defined by*

$$x^{k+1} = \beta N x^k,\ k \geq 0, \tag{10}$$

then $x^{k+1} = \frac{\beta}{2}x^k,\ k \geq 1$. Either $x^1 = x^*$ or $\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = \left|\frac{\beta}{2}\right|,\ k \geq 1$.

Proof If x^0 is in the null space of the matrix N i.e. $x^0 \in \text{Null}(N)$, then $x^1 = x^*$. Hence the iteration terminates with the solution after one iteration when $x^0 \in \text{Null}(N)$ or $\beta = 0$.

Consider the sequence of iterates $\{x^k\},\ k \geq 0$ given by (10). Since $N = VDV^{-1}$, then $N^2 = \frac{1}{2}N$. Now let $k \geq 1$, and using the iterate (10), then

$$x^{k+1} = \beta N x^k = \beta^2 N^2 x^{k-1} = \frac{\beta^2}{2} N x^{k-1} = \frac{\beta}{2}\left(\beta N x^{k-1}\right) = \frac{\beta}{2} x^k.$$

Then $\left\|x^{k+1} - x^*\right\| / \|x^k - x^*\| = |\beta/2|,\ k \geq 1$, using that $x^* = 0$ is the solution, and the sequence of iterates is converging provided $|\beta| < 2$. □

This lemma will be used to show global and linear rate of convergence for the methods considered. For simplicity, the case of reaching the solution in one iteration i.e. $x^1 = x^*$ is not considered in the theorem.

We can now state a global convergence result for Newton's method and methods in the Halley class.

Theorem 1 *Let x^0 be a starting point so that $\nabla^2 f(x^0)$ is nonsingular. Newton's method and methods in Halley class for $\alpha \neq 3/2, 5/4$ are well defined for all $k \geq 0$ and given by (8) and (9), respectively. Further, for any starting point the sequence of iterates $\{x^k\}$ defined by (8) is globally convergent and converges linearly to the solution x^* with quotient factor $\frac{2}{3}$. For any starting point the sequence of iterates $\{x^k\}$ given by (9) is globally convergent and converges linearly to the solution x^* with quotient factor $\left|\frac{4\alpha-5}{6\alpha-9}\right|$ provided $\alpha < 7/5$ or $\alpha > 2$.*

Proof First, we start with Newton's method. Assume that the Hessian $\nabla^2 f(x^k)$ is nonsingular. Then x^{k+1} is well defined and $f_i(x^{k+1}) = \frac{4}{9} f_i(x^k) \neq 0$, $i = 3, 4$. Then the Hessian $\nabla^2 f(x^{k+1})$ is nonsingular since $\det(\nabla^2 f(x^{k+1})) \neq 0$. By induction, Newton's method is well defined for all $k \geq 0$ and the iterates is given by (8). Using Lemma 1 for $\beta = \frac{4}{3}$ and (8) we have $\frac{\left\|x^{k+1}-x^*\right\|}{\|x^k-x^*\|} = \frac{2}{3},\ k \geq 1$. Thus Newton's method converges linearly to x^* with quotient factor $\frac{2}{3}$.

Similarly, using Halley class iterate (9), then from Lemma 1, the quotient factor of methods in Halley class is $\left|\frac{4\alpha-5}{6\alpha-9}\right|$ which is less than one for the values $\alpha < 7/5$ or $\alpha > 2$. □

3 Line search and Newton's direction

Define the matrix $M = -\left(I - \frac{4}{3}N\right)$, then from (6) the Newton direction is obtained by $s^k = Mx^k$. We introduce a step length λ which is a positive scalar and the new iterate is on the form

$$x^{k+1} = x^k + \lambda s^k = (I + \lambda M)\, x^k. \tag{11}$$

3.1 The Wolfe Conditions

Let the function $\phi(\lambda) = f(x + \lambda s)$, where s is the search direction. Wolfe conditions [33] are $\phi(\lambda) \leq \phi(0) + \lambda c_1 \phi'(0)$ and $\phi'(\lambda) \geq c_2 \phi'(0)$, which are to avoid too big or too small step lengths. The step length $\lambda = 1$ will satisfy the Wolfe conditions if $c_1 \leq \frac{1}{2}$ and

$c_2 \geq \frac{8}{27} \approx 0.296$. Grippo et al. [19] observed numerically that the full Newton step, $\lambda = 1$, satisfy Armijo's condition [33, Section 3.1]. Also the more stringent condition $\left|\phi'(\lambda)\right| \leq c_2 \left|\phi'(0)\right|$ is satisfied for $\lambda = 1$ provided $c_2 \geq \frac{8}{27}$. Typical choice for c_1 and c_2 in Newton's method is $c_1 = 10^{-4}$ and $c_2 = .9$ [9, Appendix A, Page 328] and [33, Section 3.1].

3.2 The Exact Line Search

Define the function $\phi_k(\lambda)$ at iterate k, as $\phi_k(\lambda) = a_k(1-\lambda)^2 + b_k\left(1-\frac{\lambda}{3}\right)^4$, where $a_k = f_1(x^k)^2 + f_2(x^k)^2$, and $b_k = f_3(x^k)^2 + f_4(x^k)^2$. The exact minimizer λ_k is defined by $\lambda_k = \arg\min \phi_k(\lambda)$. The function f is strictly convex so $\phi_k(\lambda)$ is strictly convex and the exact minimizer λ_k is the solution of the cubic equation $\phi_k'(\lambda) = 0$. The cubic equation has one real and two complex conjugate solutions. Let $w_k = \frac{3a_k}{2b_k}$ then the real root of the cubic polynomial is

$$\lambda_k = 3\left(\left(-w_k + w_k(1+w_k)^{\frac{1}{2}}\right)^{\frac{1}{3}} - \left(w_k + w_k(1+w_k)^{\frac{1}{2}}\right)^{\frac{1}{3}} + 1\right). \qquad (12)$$

From (11) we have $f_i(x^{k+1}) = (1-\lambda)f_i(x^k)$, for $i = 1, 2$ and $f_i(x^{k+1}) = \left(1-\frac{\lambda}{3}\right)^2 f_i(x^k)$, for $i = 3, 4$. Thus, $w_{k+1} = \frac{(1-\lambda)^2}{\left(1-\frac{\lambda}{3}\right)^4} w_k$. The minimizer λ_k will satisfy $\phi_k'(\lambda_k) = 0$, and the expression for w_{k+1} simplifies $w_{k+1} = \frac{\lambda_k - 1}{1-\frac{\lambda_k}{3}}$. Let

$$\lambda(w) = 3\left(\left(-w + w(1+w)^{\frac{1}{2}}\right)^{\frac{1}{3}} - \left(w + w(1+w)^{\frac{1}{2}}\right)^{\frac{1}{3}} + 1\right).$$

The function $\lambda(w)$ is monotonically decreasing for $w \geq 0$, $\lambda(0) = 3$, and $\lambda(w) > 1$. Consider the function $g(w) = (\lambda(w) - 1)/(1 - \lambda(w)/3)$ which is also monotonically decreasing for $w > 0$ and has a fixed point $w^* = g(w^*)$. Since the sequence $\{w_k\}_{k\geq 0}$ satisfies the fixed point iteration, and $\lambda^* = \lambda(w^*)$ then $\lambda_k \to \lambda^*$ as $k \to \infty$ and $\lambda^* = \frac{15-3\sqrt{17}}{2}$. The quotient factor for the iteration (11) is given by $(\sqrt{17}-3)/2 \approx 0.5616$.

Table 1 is a comparison of the quotient factor of super-Halley, Halley, Chebyshev, Newton's methods in section 2 and exact line search. It is observed that Newton's method has the slowest convergence, after that Chebyshev, then Halley and then super-Halley has the fastest convergence. Table 1 shows that the quotient factor of super-Halley is half of the one given by Newton's method. However two steps of Newton is slower than one step of super-Halley. In addition, Newton's method is slower than exact line search.

Table 1 Quotient factor for different methods and for exact line search

Super-Halley	Halley	Chebyshev	Newton	Newton with exact line search
$1/3$	$1/2$	$5/9$	$2/3$	≈ 0.5616

The slow convergence of Newton's method was observed by Dixon [10]. A significant reduction in the number of iterations was achieved by using the dominating degree $D = 4$ [11], and a step length $\lambda = D - 1$ after a Newton with $\lambda = 1$. To see this, we have $f_i(x^{k+1}) = (1-\lambda)f_i(x^k)$, $i = 1, 2$ and $f_i(x^{k+1}) = (1-\frac{\lambda}{3})^2 f_i(x^k)$, $i = 3, 4$ from the definition (3) and the iterate (11). Moreover, after a step length $\lambda_k = 1$, we have $f_1(x^j) =$

$f_2(x^j) = 0$ for $j \geq k+1$. For $\lambda_k = 3$, the functions $f_3(x^j) = f_4(x^j) = 0$ for $j \geq k+1$. The Newton's method will therefore terminate after two iterations using step lengths $\lambda_0 = 1$ and $\lambda_1 = 3$.

4 PSF as a system of nonlinear equations

Consider the PSF as a nonlinear system of equations $F(x) = 0$ where F is given in (3), and the determinant of the Jacobian matrix is $84\sqrt{50}(x_1 - x_4)(x_2 - 2x_3)$. We immediately have that $F'(x)$ is nonsingular if and only if $f_3(x) \neq 0$ and $f_4(x) \neq 0$ using the definition (3). Assuming that $F'(x)$ is nonsingular, it can be shown that the Newton iterate is given by

$$x^+ = x - F'(x)^{-1}F(x) = Nx,$$

where N is given by (7). From definition (4), we have $f_i(x^+) = \frac{1}{4}f_i(x)$, $i = 3, 4$. We can conclude that $\det(F'(x^+)) \neq 0$ whenever $\det(F'(x)) \neq 0$, and Newton's method is well defined and from Lemma 1 will be globally convergent provided that the Jacobian at the initial point is nonsingular.

Consider any method in the Halley class where the iterate is given by

$$\begin{aligned} s_{(1)} &= -F'(x)^{-1}F(x), \\ s_{(2)} &= -\frac{1}{2}\left(F'(x) + \alpha F''(x)s_{(1)}\right)^{-1} F''(x)s_{(1)}s_{(1)}, \\ x^+ &= x + s_{(1)} + s_{(2)}. \end{aligned}$$

For a given scalar α, the determinant of the matrix $F'(x) - \alpha F''(x)F'(x)^{-1}F(x)$ is $105\sqrt{2}(\alpha - 2)^2(x_1 - x_4)(x_2 - 2x_3)$. It follows that, if $F'(x)$ is nonsingular and $\alpha \neq 2$, then $F'(x) - \alpha F''(x)F'(x)^{-1}F(x)$ is also nonsingular. The next iterate x^+ is

$$x^+ = \frac{2\alpha - 3}{2(\alpha - 2)}Nx.$$

The last two components of F at x^+ are given by $f_i(x^+) = \frac{(2\alpha-3)^2}{16(\alpha-2)^2}f_i(x)$, $i = 3, 4$, then $\det(F'(x^+)) \neq 0$ when $\det(F'(x)) \neq 0$, $\alpha \neq 3/2$ and $\alpha \neq 2$. We can conclude that if the Jacobian matrix is nonsingular at x then for $\alpha \neq 2$ the new iterate x^+ is well defined. If in addition $\alpha \neq 3/2$ the Jacobian at x^+ is nonsingular and the methods in the Halley class are well defined and globally convergent for $\alpha < 11/6$ or $\alpha > 5/2$. The quotient factors for some of these methods described in Table 2.

For a quadratic function two steps of Newton's method is one step of super-Halley method ($\alpha = 1$) [20]. The function (4) is quadratic and to see that two steps of Newton's method is one step of super-Halley, recall $N = VDV^{-1}$

$$x^+ = NNx = N^2x = VD^2V^{-1}x = \frac{1}{2}VDV^{-1}x = \frac{1}{2}Nx.$$

Table 2 Quotient factor for different methods on $F(x) = 0$

Super-Halley	Halley	Chebyshev	Newton
$1/4$	$1/3$	$3/8$	$1/2$

5 Numerical illustrations

In this section numerical comparison between the methods with the variation of PSF described in Section 2 is introduced. In addition to some experiments on the exact line search explained in details in Section 3. All experiments conducted from MATLAB 7.11 release R21010b on a personal computer with processor Intel Core 2 Duo 2.4 GHz and 3.8 GB memory.

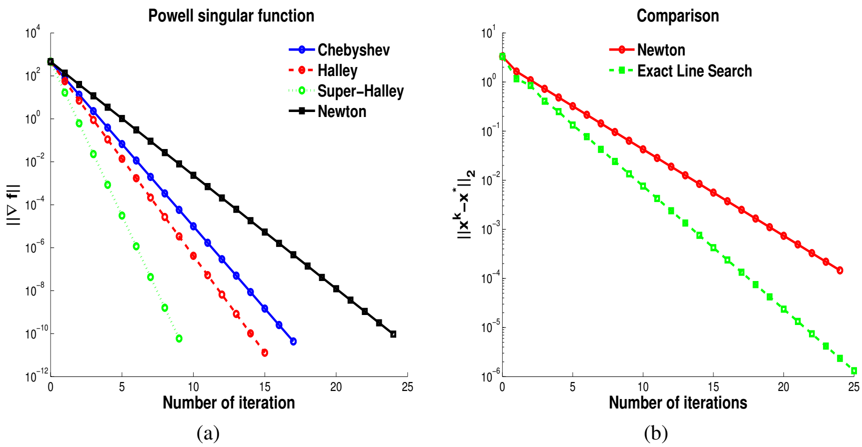

Fig. 1 (a) Comparing the number of iterations of the used method to solve PSF as unconstrained optimization problem. (b) Comparing Newton's method to the exact line search in the perspective of the iterate.

Consider PSF problem $\nabla f(x) = 0$, where the objective function is given by (1). We solved this problem by Newton, Chebyshev, Halley, and super-Halley methods. The standard starting point $x^0 = (3, -1, 0, 1)^T$ used with the stopping criteria $||\nabla f(x)|| \leq 10^{-10}$. Figure 1(a) shows that all methods are linearly converging and super-Halley has the smallest quotient factor of the other methods. Figure 1(b) is the comparison between the exact line search and Newton's method with respect to the iterate x^k. It can be shown from Figure 1(b) that the exact line search has the least values of $||x^k - x^*||_2$ comparing to Newton's method. We also see that the quotient factor stabilizes after 4 iterations.

References

1. Addis, B., Locatelli, M.: A new class of test functions for global optimization. Journal of Global Optimization **38**(3), 479–501 (2007)
2. Amat, S., Busquier, S., Gutiérrez, J.M., Hernández, M.A.: On the global convergence of Chebyshev's iterative method. Journal of Computational and Applied Mathematics **220**(1-2), 17–21 (2008)
3. Bongartz, I., Conn, A.R., Gould, N., Toint, P.L.: CUTE: Constrained and unconstrained testing environment. ACM Transactions on Mathematical Software **21**(1), 123–160 (1995)
4. Brent, R.P.: Algorithms for Minimization Without Derivatives. Prentice-Hall Inc., Englewood Cliffs, New Jersey (1973)
5. Brent, R.P.: Some efficient algorithms for solving systems of nonlinear equations. SIAM Journal on Numerical Analysis **10**(2), 327–344 (1973)
6. Conn, A.R., Gould, N.I.M., Toint, P.L.: Testing a class of methods for solving minimization problems with simple bounds on the variables. Mathematics of Computation **50**(182), 399–430 (1988)

7. Csendes, T., Ratz, D.: Subdivision direction selection in interval methods for global optimization. SIAM Journal on Numerical Analysis **34**(3), 922–938 (1997)
8. Cuyt, A., Cruyssen, P.: Abstract Padé-approximants for the solution of a system of nonlinear equations. Computers & Mathematics with Applications **9**(4), 617–624 (1983)
9. Dennis, J.E., Schnabel, R.B.: Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice-Hall, Englewood Cliffs, New Jersey (1983)
10. Dixon, L.C.W.: Nonlinear optimisation: A survey of the state of the art. In: D.J. Evans (ed.) Software for Numerical Mathematics: Proceedings of the Loughborough University of Technology Conference of the Institute of Mathematics and Its Applications Held in April 1973, pp. 193–216. Academic Press, Inc., Orlando, USA (1974)
11. Dixon, L.C.W., Biggs, M.C.: Meander–a Newton based procedure for n-dimensional function minimization. Tech. Rep. 9, The Haffield Polytechnic, Numerical Optimization Center (1970)
12. Du, D.Z., Pardalos, P.M., Wu, W.: Mathematical Theory of Optimization. Kluwer Academic Publishers (2001)
13. Du, D.Z., Pardalos, P.M., Wu, W.: Rosen's method, global convergence, and Powell's conjecture. In: C.A. Floudas, P.M. Pardalos (eds.) Encyclopedia of Optimization, 2nd edn., pp. 3345–3354. Springer (2009)
14. El-Bakry, A., Steihaug, T.: On the component-wise convergence rate. Tech. Rep. TR 98-16, Department of Computational and Applied Mathematics, Rice University (2000)
15. Fletcher, R.: Practical Methods of Optimization, 2nd edn. John Willey & Sons Ltd. (1987)
16. Fletcher, R., Freeman, T.L.: A modified Newton method for minimization. Journal of Optimization Theory and Applications **23**(3), 357–372 (1977)
17. Floudas, C.A., Pardalos, P.M., Adjiman, C., Esposito, W.R., Gümüs, Z.H., Harding, S.T., Klepeis, J.L., Meyer, C.A., Schweiger, C.A.: Handbook of test problems in local and global optimization. Kluwer Academic Publishers, Dordrecht (1999)
18. Goldfarb, D., Wang, S.: Partial-update Newton methods for unary, factorable, and partially separable optimization. SIAM Journal on Optimization **3**(2), 382–397 (1993)
19. Grippo, L., Lampariello, F., Lucidi, S.: A nonmonotone line search technique for Newton's method. SIAM Journal on Numerical Analysis **23**(4), 707–716 (1986)
20. Gundersen, G., Steihaug, T.: On large-scale unconstrained optimization problems and higher order methods. Optimization Methods and Software **25**(3), 337–358 (2010)
21. Hirsch, M.J., Meneses, C.N., Pardalos, P.M., Resende, M.G.C.: Global optimization by continuous grasp. Optimization Letters **1**(2), 201–212 (2007)
22. Hirsch, M.J., Pardalos, P.M., Resende, M.G.C.: Solving systems of nonlinear equations with continuous GRASP. Nonlinear Analysis Series, Real World Applications **10**(4), 2000–2006 (2009)
23. Hu, J., Fu, M.C., Marcus, S.I.: A model reference adaptive search method for global optimization. Operations Research **55**(3), 549–568 (2007)
24. Kearfott, R.B.: Some tests of generalized bisection. ACM Transactions on Mathematical Software **13**(3), 197–220 (1987)
25. Lukšan, L.: Computational experience with improved variable metric methods for unconstrained minimization. Kybernetika **26**(5), 415–431 (1990)
26. McCormick, G.P., Sofer, A.: Optimization with unary functions. Mathematical Programming **52**(1), 167–178 (1991)
27. Melman, A.: Geometry and convergence of Euler's and Halley's methods. SIAM Review **39**(4), 728–735 (1997)
28. Moré, J.J., Cosnard, M.Y.: Numerical solution of nonlinear equations. ACM Transactions on Mathematical Software **5**(1), 64–85 (1979)
29. Moré, J.J., Garbow, B.S., Hillstrom, K.E.: Algorithm 566: Fortran subroutines for testing unconstrained optimization software. ACM Transactions on Mathematical Software **7**(1), 136–140 (1981)
30. Moré, J.J., Garbow, B.S., Hillstrom, K.E.: Testing unconstrained optimization software. ACM Transactions on Mathematical Software **7**(1), 17–41 (1981)
31. Morgan, A.P.: A method for computing all solutions to systems of polynomials equations. ACM Transactions on Mathematical Software **9**(1), 1–17 (1983)
32. Nelder, J.A., Mead, R.: A simplex method for function minimization. Computer Journal **7**, 308–313 (1965)
33. Nocedal, J., Wright, S.J.: Numerical Optimization, 2nd edn. Springer-Verlag (2006)
34. O'Neill, R.: Algorithm AS 47: Function minimization using a simplex procedure. Journal of the Royal Statistical Society. Series C (Applied Statistics) **20**(3), 338–345 (1971)
35. Ostrowski, A.M.: Solution of Equations in Euclidean and Banach Spaces. Academic Press, New York (1973). Third edition of Solution of Equations and Systems of Equations, Pure and Applied Mathematics, Vol. 9

36. Özdamar, L., Demirhan, M.: Experiments with new stochastic global optimization search techniques. Computers & Operations Research **27**(9), 841–865 (2000)
37. Powell, M.J.D.: An iterative method for finding stationary values of a function of several variables. The Computer Journal **5**, 147–151 (1962)
38. Rao, S.S.: Engineering Optimization: Theory and Practice, 3rd edn. John Wiley & Sons, Inc. (1996)
39. Roose, A., Kulla, V., Lomp, M., Meressoo, T.: Test Examples of Systems of Nonlinear Equations. Estonian Software and Computer Service Company, Tallinn (1990)
40. Rosenbrock, H.H., Storey, C.: Computational Techniques for Chemical Engineers. Pergamon Press (1966)
41. Schittkowski, K.: More Test Examples for Nonlinear Programming Codes, *Lecture Notes in Economics and Mathematical Systems*, vol. 282. Springer-Verlag New York, Inc. (1987)
42. Schwefel, H.P.: Numerical Optimization of Computer Models. John Wiley & Sons, Inc. (1981)
43. Spedicato, E.: Computational experience with quasi-Newton algorithms for minimization problems of moderately large size. In: L.C.W. Dixon, G.P. Szegö (eds.) Towards Global Optimisation 2, pp. 209–219. North-Holland Publishing Company, New York, USA (1978)
44. Steihaug, T., Suleiman, S.: Notes on invariance of higher order methods under linear transformation. Tech. Rep. (under preparation), Department of Informatics, University of Bergen (2012)
45. Stewart, G.W.: A modification of Davidon's minimization method to accept difference approximations of derivatives. Journal of the Association for Computing Machinery **14**(1), 72–83 (1967)
46. Sun, W., Yuan, Y.: Optimization Theory and Methods: Nonlinear Programming. Springer Optimization and Its Applications. Springer (2006)
47. Thorlund-Petersen, L.: Global convergence of Newton's method on an interval. Mathematical Methods of Operations Research **59**, 91–110 (2004)
48. Walster, G.W., Hansen, E.R., Sengupta, S.: Test results for a global optimization algorithm. In: Proceedings of the SIAM Conference on Numerical Optimization, Boulder, Colorado June 12–14,1984, pp. 272–287. SIAM (1985)