Rapporto n. 5 2007
dmsia unibg.it

# A. Galántai and E. Spedicato ABS methods for nonlinear systems of algebraic equations

*Serie Ricerca*

Dipartimento
di Matematica, Statistica,
Informatica e Applicazioni
"Lorenzo Mascheroni"

UNIVERSITÀ DEGLI STUDI DI BERGAMO

# ABS methods for nonlinear systems of algebraic equations

A. Galántai and E. Spedicato

22-06-2007

**Abstract**

This paper gives a survey of the theory and practice of nonlinear ABS methods including various types of generalizations and computer testing. We also show three applications to special problems, two of which are new.

**Keywords.** Linear ABS methods, Huang method, Brent-Brown type methods, Petrov-Galerkin method, Newton method, quasi-Newton methods, bordered nonlinear systems, constrained optimization, primal dual interior point method

## 1 Introduction

The ABS algorithms have been introduced in 1984 by Abaffy, Broyden and Spedicato [3] to solve determined or underdetermined linear systems and have been later extended to linear least squares, nonlinear equations, optimization problems and integer (Diophantine) equations. The related literature consists presently of over 400 papers (see Abaffy and Spedicato's monograph [6] for a presentation of such techniques as available at the beginning of 1989 and Spedicato et al. [67] for a review dealing with later results). In the present paper we give a survey of the results concerning the nonlinear ABS methods and present some new applications as well.

For convenience, we recall the basic steps of the (unscaled) ABS class for solving the linear system

$$Ax = b \quad (A = [a_1, \ldots, a_m]^T \in \mathbb{R}^{m\times n},\ m \le n).$$

(a) $x_1 \in \mathbb{R}^n$, $H_1 \in \mathbb{R}^{n\times n}$ $(\det(H_1) \neq 0)$, $i = 1$.
(b) $s_i = H_i a_i$, $\tau_i = a_i^T x_i - b_i$. If $s_i \neq 0$ go to (c), the $i-$th equation is linearly independent from the previous equations. Otherwise, if $\tau_i \neq 0$ stop, the system has no solution, while if $\tau_i = 0$, remove equation $i$, being a linear combination of the previous ones, so set $x_{i+1} = x_i, H_{i+1} = H_i$ and go to (f).
(c) $p_i = H_i^T z_i,\ z_i \in \mathbb{R}^n$ $(z_i^T H_i a_i \neq 0)$

(d) $x_{i+1} = x_i - \alpha_i p_i$ ($\alpha_i = ({a_i}^T x_i - b_i)/a_i^T p_i$)
(e) $H_{i+1} = H_i - H_i a_i w_i^T H_i / w_i^T H_i a_i$ ($w_i \in \mathbb{R}^n$, $w_i^T H_i a_i \neq 0$).
(f) Stop, if $i = m$, $x_{m+1}$ solves the system, otherwise increment $i$ by one and go to (b).

The scaled ABS class can be obtained by applying the unscaled class to the scaled linear system $V^T Ax = V^T b$, where $V \in \mathbb{R}^{m \times m}$. For later use, we recall the so called Huang or implicit Gram-Schmidt algorithm, defined by the choices $H_1 = I$, $z_i = w_i = a_i$. This algorithm determines the unique solution of least Euclidean norm of an underdetermined system, if it is started with an arbitrary vector $x_1$ proportional to $a_1$, usually the zero vector. The solution is moreover approached, if $x_1 = 0$, monotonically in norm from below. Additionally, the search directions $p_i$ are orthogonal. The algorithm's stability can be improved in several ways. Usually one does a reprojection on the search direction, i.e., $p_i = H_i(H_i a_i)$, and then defines the update as

$$H_{i+1} = H_i - p_i p_i^T / p_i^T p_i.$$

This modification is called the modified Huang algorithm. Extensive testing indicates that this reprojected ABS method is very stable. Recently, Gáti [37] using a significantly improved version of Miller's automatic roundoff analyzer program [57] showed that the reprojected Huang is much more stable than the modified Gram-Schmidt method. For other special ABS methods, we refer to [6].

The nonlinear ABS-methods are Brown-Brent type projection methods. Their development was motivated by an idea of Stewart [72] and the fact that the Brent and Brown methods, when applied to linear systems of the form $Ax = b$, belong to the class of linear ABS methods ([5], [6]). The Brown-Brent type methods are known to be efficient and competitive to the Newton method in solving nonlinear systems of algebraic equations.

The first method of this type was published by Brown [8] in 1966 (see, also [61]). Brown [9] proved the local quadratic convergence in 1969. Brent [7] published his famous method in 1973. Gay [38] introduced the first class of Brown-Brent type methods in his Ph.D. thesis in 1975. Later Schmidt and Hoyer [64] developed a new generalization of the Brown-Brent type methods and proved various convergence theorems. They also showed the excellent theoretical efficiency properties of the Brown-Brent type methods [65]. Moré and Cosnard [58] developed an efficient FORTRAN implementation of the Brent method, which is available through Netlib as Algorithm TOMS 554 [59]. In 1979 Martinez [56] gave a block generalization of the Brown-Brent type methods, while Schwetlick [66] included one chapter on the theory and practice of the Brown-Brent type methods. In 1986 Frommer [20] proved the monotone convergence of the Brown method in the partial ordering. Two years later he also proved that Brown's method is faster than Newton's method in the partial ordering [21].

Next we explain the basic idea of the Brown-Brent type methods (see also [65], [66], [58]), which is essentially an extension of the idea of Petrov-Galerkin methods for linear systems (see, e.g. [62], [33]). Assume that the nonlinear

systems of algebraic equations are in the form

$$f(x) = 0 \quad (f : \mathbb{R}^m \to \mathbb{R}^m), \tag{1}$$

where $f(x) = [f_1(x), ..., f_m(x)]^T$. The solution of the system is denoted by $x^*$. Using the Newton-method one solves the linearized equation

$$f(x) + f'(x)(x^+ - x) = 0$$

for $x^+$ and expects a better approximation to the solution $x^*$. The Brown-Brent type methods replace this linearized system by

$$f_j(x) + w_j^T(x^+ - x) = 0 \quad (j = 1, \ldots, m),$$

where $w_j$ is an approximation to $\nabla f_j(x)$. Given an iterate $x$ one obtains the next (major) iterate $x^+$ by the following algorithm

$y_1 = x$
**for** $k = 1 : m$
  Let $y_{k+1}$ be the solution of $f_j(y_j) + w_j^T(y - y_j) = 0$ $(1 \le j \le k)$.
**end**
$x^+ = y_{m+1}$.

Since $y_{k+1}$ satisfies the first $k$ linearized equations, one gets the orthogonality relations

$$w_j^T(y_{k+1} - y_k) = 0 \quad (1 \le j \le k-1).$$

Select now directions $\{p_k\}_{k=1}^m$ such that $y_{k+1} - y_k = \lambda_k p_k$ $(k = 1, \ldots, m)$ holds with $w_j^T p_k = 0$ $(j < k)$ and $w_k^T p_k \neq 0$. Then one has

$$\lambda_k = -f_k(y_k) / w_k^T p_k$$

and

$$y_{k+1} = y_k - f_k(y_k)p_k/(w_k^T p_k) \quad (k = 1, \ldots, m).$$

The Brown-Brent type methods differ from each other in the selection of $p_i$'s and $w_i$'s. The Brown method is related to the Gaussian elimination because its direction matrix $P = [p_1, \ldots, p_m]$ is an approximation of the inverse of the upper triangular factor of the $LU$-factorization of $f'(x)$ (see also [61]). Brent's method chooses $p_k$ such that $y_{k+1}$ is the closest solution to $y_k$ in the Euclidean norm. It is an important feature of the Brown-Brent type methods that if $f_k(x)$ is linear function and $w_k = \nabla f_k(x)$, then the $k$th equation is solved exactly (see Schmidt, Hoyer [65], Moré, Cosnard [58] and Schwetlick [66]). Bus [10] and Ypma [73] suggest the elimination of the linear components of a nonlinear system by transforming it to an equivalent smaller nonlinear system with no linear components. It is clear however that no such transformation is necessary in the case of a properly chosen Brent-Brown type method.

## 2 Different forms of the nonlinear ABS methods

The present form of nonlinear ABS methods was developed in two steps by Abaffy, Galántai and Spedicato [5], and Abaffy and Galántai [4] (see also [6]). In the sequel $I_m$ and $A(x) = f'(x)$ denotes the $m \times m$ unit matrix and the Jacobian matrix of $f(x)$, respectively.

ALGORITHM N1 (block nonlinear ABS methods)
$x_1 \approx x^*$, $i := 1$.
**while** $i \geq 1$
$y_1 = x_i$, $H_1 = I_m$
Choose $V = [V_1, \ldots, V_{r(i)}] \in \mathbb{R}^{m \times m}$ ($V_j \in \mathbb{R}^{m \times m_j}$, $1 \leq r(i) \leq m$) such that $\det(V) \neq 0$.
**for** $k = 1 : r(i)$
Choose weights $\tau_{jk} \geq 0$ ($j = 1, \ldots, k$) such that $\sum_{j=1}^{k} \tau_{jk} = 1$.
$u_k = \sum_{j=1}^{k} \tau_{jk} y_j$
$P_k = H_k^T Z_k \quad (Z_k \in \mathbb{R}^{m \times m_k}, \det\left(P_k^T A^T(u_k) V_k\right) \neq 0)$
$y_{k+1} = y_k - P_k \left(V_k^T A(u_k) P_k\right)^{-1} V_k^T f(y_k)$
Choose $W_k \in \mathbb{R}^{m \times m_k}$ such that $W_k^T H_k A^T(u_k) V_k = I_{m_k}$.
$H_{k+1} = H_k - H_k A^T(u_k) V_k W_k^T H_k$
**end**
$x_{i+1} = y_{r(i)+1}$, $i = i + 1$
**end**

A particular method is defined by the parameter matrices $V = [V_1, \ldots, V_{r(i)}]$, $W = [W_1, \ldots, W_{r(i)}]$, $Z = [Z_1, \ldots, Z_{r(i)}]$ ($V, W, Z \in \mathbb{R}^{m \times m}$) and $T = [\tau_{ij}]_{i,j=1}^{r(i)}$, where $\tau_{11} = 1$ and $\tau_{ij} = 0$ for $i > j$. By definition $u_k = [y_1, \ldots, y_{r(i)}] T e_k$, where $e_k \in \mathbb{R}^{r(i)}$ is the $k$th unit vector. The partition and the parameter matrices may vary with the major iterates.

The weight matrix $T$ may represent different strategies for choosing the "stepsize" $(V_k^T A^T(u_k) P_k)^{-1} V_k^T f(y_k)$. The parameter set

$$u_k = y_k \quad (k = 1, \ldots, r(i);\ T = I_{r(i)})$$

corresponds to the Seidel principle and reevaluates the Jacobian matrix "row" by "row". The choice

$$u_k = y_1 \quad (k = 1, \ldots, r(i);\ T = [e_1, \ldots, e_1] \in \mathbb{R}^{r(i) \times r(i)})$$

keeps the Jacobian matrix fixed during the minor iterations. Stewart [72] suggested a similar approach for a general class of conjugate direction methods.

The Newton method corresponds to the choice $r(i) = 1$ ($i > 0$). If $r(i) = m$ for all $i$, then the form of Algorithm N1 is the following.

ALGORITHM N2 (scaled nonlinear ABS methods)
$x_1 \approx x^*$, $i := 1$.
**while** $i \geq 1$
$y_1 = x_i$, $H_1 = I_m$
Choose $V \in \mathbb{R}^{m \times m}$ such that $\det(V) \neq 0$.

**for** $k = 1 : m$
Choose the weights $\tau_{jk} \geq 0$ such that $\sum_{j=1}^{k} \tau_{jk} = 1$.
$u_k = \sum_{j=1}^{k} \tau_{jk} y_j$
$p_k = H_k^T z_k \quad (z_k \in \mathbb{R}^m,\, p_k^T A^T(u_k)\, v_k \neq 0)$
$y_{k+1} = y_k - p_k \left(v_k^T A(u_k)\, p_k\right)^{-1} v_k^T f(y_k)$
Choose $w_k \in \mathbb{R}^m$ such that $w_k^T H_k A^T(u_k)\, v_k \neq 0$.
$H_{k+1} = H_k - H_k A^T(u_k) v_k w_k^T H_k / w_k^T H_k A^T(u_k) v_k$
**end**
$x_{i+1} = y_{m+1},\, i = i + 1$
**end**

The Brown method is defined by the parameters $V = I_m$, $r(i) = m$, $W = Z = I_m$, $T = I_m$. The scaled Huang method is given by the selection $r(i) = m$, $z_k = w_k = A^T(u_k) v_k$ $(k = 1, \ldots, m)$. The symmetric conjugate direction (SCD) subclass [24], [26] is given by $W = Z = Q$, $v_k = C(u_k)\, p_k$ $(k = 1, \ldots, m)$. In the latter case it is assumed that $C(x) \in \mathbb{R}^{m \times m}$ is continuous and $G(x) = C(x)^T A(x)$ is a positive definite symmetric matrix in $B(x^*, \delta_0)$. Notation $B(x, \delta)$ stands for the open ball around $x$ with radius $\delta$.

# 3 Local convergence results

Several authors ([5], [4], [1], [22], [23], [75], [16], [17], [18], [14], [45], [23], [24], [26], [27], [28], [29]) proved local convergence results with various techniques under the following standard assumptions:

$$\exists A(x^*)^{-1}, \tag{2}$$

$$\exists K_0 \geq 0, \delta_0 > 0 : \|f(x) - f(y)\| \leq K_0 \|x - y\| \quad (x, y \in B(x^*, \delta_0)), \tag{3}$$

$$\exists K_1 \geq 0, 0 < \alpha \leq 1 : \|A(x) - A(y)\| \leq K_1 \|x - y\|^{\alpha} \quad (x, y \in B(x^*, \delta_0)). \tag{4}$$

The most general local convergence result [23] is given by

**Theorem 1** *Assume that conditions (2)-(4) hold, Algorithm N1 satisfies*

$$\det\left(P_k^T A^T(u_k) V_k\right) \neq 0, \quad W_k^T H_k A^T(u_k) V_k = I_{m_k} \quad (k = 1, \ldots, r(i)) \tag{5}$$

*and the matrices $V, P$ are such that*

$$\left\| P_k \left(V_k^T A(u_k) P_k\right)^{-1} V_k^T \right\| \leq K_2 \quad (k = 1, \ldots, r(i);\ i \geq 1) \tag{6}$$

*holds. Then there exists a number $\delta^* > 0$ $(\delta^* \leq \delta_0)$ such that for every $x_1 \in B(x^*, \delta^*)$, the sequence $\{x_i\}_{i=1}^{\infty}$ of the major iterates converges to $x^*$ with a speed of Q-order not less than $1 + \alpha$.*

Condition (6) is equivalent with the condition

$$\|R_k\| = \left\| P_k (V_k^T A(u_k) P_k)^{-1} V_k^T A(u_k) \right\| \leq \Gamma_2 \quad (k = 1, \ldots, r(i);\ i \geq 1) \tag{7}$$

provided that $A(u_k)$ and $A(u_k)^{-1}$ are uniformly bounded in $B(x^*, \delta^*)$. Hence condition (6) holds if and only if the projectors $R_k$ are uniformly bounded. The proof of Theorem 1 is also of some interest (see [23], [24], [26], [33]). It is not only revealing the behavior of the minor iterations but widely usable elsewhere for similar convergence analysis (see [26], [27], [28], [29], [33]).

The earlier convergence results follow from Theorem 1 (see [24], [26]). For the three special cases of the ABS methods the following result holds (for proof and related results, see [24], [26]).

**Theorem 2** *Assume that conditions (2)-(4) hold. Then*
*(i) The Brown method is locally convergent if $A(x)$ is strongly nonsingular;*
*(ii) The scaled Huang subclass is locally convergent if the scaling matrices $V$ satisfy $\|V - V_0\| \leq \frac{1}{2\|V_0^{-1}\|}$ $(\det(V_0) \neq 0)$;*
*(iii) Let $T = [e_1, ..., e_1] \in \mathbb{R}^{m\times m}$. The SCD ABS methods are locally convergent if $Q_i = Q$ $(i \geq 1)$ and $Q$ is nonsingular $(i \geq 1)$.*

# 4 Monotone convergence in partial ordering

We use the natural partial ordering for vectors and matrices; that is, $x \leq y$ $(x, y \in \mathbb{R}^m)$ if and only if $x_k \leq y_k$ $(k = 1, \ldots, m)$, and $A \leq B$ $(A, B \in \mathbb{R}^{m\times m})$ if and only if $a_{ij} \leq b_{ij}$ $(i, j = 1, \ldots, m)$. The function $F : \mathbb{R}^m \to \mathbb{R}^m$ is said to be *convex* on a convex set $D \subseteq \mathbb{R}^m$ if

$$F(\alpha x + (1-\alpha y)) \leq \alpha F(x) + (1-\alpha) F(y) \tag{8}$$

holds for all $x, y \in D$ and $\alpha \in [0, 1]$. We recall the following basic result [61]. Assume that $F : \mathbb{R}^m \to \mathbb{R}^m$ is differentiable on the convex set $D$. Then $F$ is convex on $D$ if and only if

$$F(y) - F(x) \geq F'(x)(y - x) \tag{9}$$

holds for all $x, y \in D$. The function $F : \mathbb{R}^m \to \mathbb{R}^k$ is said to be *isotone*, if $x \leq y$ implies $F(x) \leq F(y)$.

Let $F(x) = V^T f(x)$ and $\tilde{A}(x) = V^T A(x)$. The following result holds for Algorithm N2 [25].

**Theorem 3** *Assume that*
*(i) $V$ is constant for $i \geq 1$, $F(x) = V^T f(x)$ is continuously differentiable and convex on $\mathbb{R}^m$, $\widetilde{A}(x)$ is a nonsingular $M$-matrix for all $x \in \mathbb{R}^m$ and $\widetilde{A} : \mathbb{R}^m \to \mathbb{R}^{m\times m}$ is isotone;*
*(ii) $Z = W = Q_i$ $(i \geq 1)$;*
*(iii) The matrices $Q_i$ are nonnegative such that $Q_i^{-1} \geq D_i \tilde{A}(x_i)$ holds for some nonsingular $D_i \geq 0$ $(i \geq 1)$;*
*(iv) There exist two nonsingular matrices $Q_\infty$ and $D_\infty$ such that $Q_i \geq Q_\infty \geq 0$ and $D_i \geq D_\infty \geq 0$ $(i \geq 1)$;*
*If $x_1 \in \mathbb{R}^m$ is an arbitrary point such that $F(x_1) \geq 0$ then Algorithm N2 satisfies*

$$x^* \leq x_{i+1} = y_{m+1} \leq y_m \leq \ldots \leq y_1 = x_i \quad (i \geq 1) \tag{10}$$

*and* $x_i \to x^*$ *as* $i \to +\infty$.

For arbitrary $x_0 \in \mathbb{R}^m$, the first Newton step $x_1 = x_0 - [F'(x_0)]^{-1} F(x_0)$ satisfies the condition $F(x_1) \geq 0$. Hence Theorem 3 is essentially a global convergence theorem. The choice of $Q_i$ is discussed in [25] and [26] (see also [33]).

It is possible to define a subordinate ABS algorithm, which provides a sequence $\{\widehat{x}_i\}$ of approximations such that $x^* \geq \widehat{x}_{i+1} \geq \widehat{x}_i$ $(i \geq 1)$ and $\hat{x}_i \to x^*$ as $i \to +\infty$ (see [25]). Thus we may have two-sided approximations to the solution $x^*$ in the form of inclusion intervals

$$\hat{x}_i \leq \hat{x}_{i+1} \leq x^* \leq x_{i+1} \leq x_i \quad (i = 1, \ldots) \tag{11}$$

with $x_i - \hat{x}_i \to 0$ $(i \to +\infty)$.

A comparison with the Newton method yields the following result [25].

**Theorem 4** *Assume that conditions of Theorem 3 are satisfied. Let* $T = [e_1, \ldots, e_1]$ *and* $Q = \tilde{A}^{-1}(y_1)D_1$, *where* $D_1 \geq 0$ *is diagonal. Then the corresponding ABS method N2 is faster than Newton's method in the partial ordering provided that they start from the same initial point* $x_1$.

Here we remind the similar result of Frommer [21] on the Brown method.

# 5 Modified nonlinear ABS methods

Although the convergence properties of the nonlinear ABS methods are similar to those of the Newton method these methods can also be quite expensive. Several modifications of the ABS methods were developed in order to decrease the computational cost per major iteration. The applied techniques include Shamanskii's idea of multistep methods [44],[26], [52], [15], [11], numerical differentiation [50], [52], [68], quasi-Newton updating [48], [47], [46], [34], [36], [39], truncation of the minor iteration loop [2], [13], [11] and ordering strategies for the equations [54]. Here we survey those modifications that are supported by sufficient numerical evidence as well.

## 5.1 Multistep ABS methods

Shamanskii's idea of multistep methods [61] was applied to the Brown-Brent type methods by Brent [7] and Hoyer [42]. It was Huang [44] who applied first Shamanskii's idea to the ABS methods.

ALGORITHM N3 (multistep block nonlinear ABS methods)

$x_1 \approx x^*$, $i = 1$.

**while** $i \geq 1$

$y_1 = x_i$, $H_1 = I_m$

Choose $V = [V_1, ..., V_{r(i)}] \in \mathbb{R}^{m \times m}$ $(V_j \in \mathbb{R}^{m \times m_j}, 1 \leq r(i) \leq m)$ such that $\det(V) \neq 0$.

**for** $k = 1 : r(i)$
Choose the weights $\tau_{jk} \geq 0$ such that $\sum_{j=1}^{k} \tau_{jk} = 1$.
$u_k = \sum_{j=1}^{m} \tau_{jk} y_j$
$P_k = H_k^T Z_k \qquad (Z_k \in \mathbb{R}^{m \times m_k}, \det\left(P_k^T A^T(u_k) V_k\right) \neq 0)$
$y_{k+1} = y_k - P_k \left(V_k^T A(u_k) P_k\right)^{-1} V_k^T f(y_k)$
Choose $W_k \in \mathbb{R}^{m \times m_k}$ such that $W_k^T H_k A^T(u_k) V_k = I_{m_k}$.
$H_{k+1} = H_k - H_k A^T(u_k) V_k W_k^T H_k$
**end**
$x^{(1)} = y_{r(i)+1}$
**for** $j = 1 : t$
$y_1^{(j)} = x^{(j)}$
**for** $k = 1 : r(i)$
$y_{k+1}^{(j)} = y_k^{(j)} - P_k \left(V_k^T A(u_k) P_k\right)^{-1} V_k^T f\left(y_k^{(j)}\right)$
**end**
$x^{(j+1)} = y_{r(i)+1}^{(j)}$
**end**
$x_{i+1} = x^{(t+1)}, \quad i = i + 1$
**end**

Using the proof of Theorem 1 we can show the $(t+2)$-order convergence of the multistep ABS methods [26] (see also [33]).

**Theorem 5** *Suppose that the conditions of Theorem 1 hold with $\alpha = 1$. Then Algorithm N3 is locally convergent with Q-order not less than $t + 2$.*

A kind of multistep formulation of the nonlinear ABS methods was also suggested by Deng and Chen [15].

## 5.2 Difference ABS methods

The difference Brown-Brent methods save half of the computational work to calculate the Jacobian matrix, which is a great advantage over the discretized Newton-method. We recall now the Schmidt-Hoyer generalization of the Brown-Brent methods.

ALGORITHM N4 (Schmidt-Hoyer-Brown-Brent class)
$x_1 \approx x^*$, $i = 1$.
**while** $i > 0$
Choose stepsize $h_i \neq 0$ and a regular $m \times m$ matrix $Q_i$.
$h = h_i$, $R_1 = Q_i$, $y_1 = x_i$
**for** $k = 1 : m$

$$a_k = \frac{1}{h}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ f_k(y_k + hR_k e_k) - f_k(y_k) \\ f_k(y_k + hR_k e_{k+1}) - f_k(y_k) \\ \vdots \\ f_k(y_k + hR_k e_m) - f_k(y_k) \end{bmatrix}$$

Find a regular $m \times m$ matrix

$$P_k = \begin{bmatrix} I_{k-1} & 0 \\ 0 & \tilde{P}_k \end{bmatrix}$$

that satisfies

$P_k^T a_k = s_k e_k$.

$R_{k+1} = R_k P_k$

$y_{k+1} = y_k - \frac{f_k(y_k)}{s_k} R_{k+1} e_k$

**end**

$x_{i+1} = y_{m+1}$, $i = i + 1$

**end**

If the stepsize $h_i$ is chosen in an appropriate way (e.g. $0 < |h_i| \leq \|x_i - x_{i-1}\|^2$ or $0 < |h_i| \leq \|F(x_i)\|$) then Algorithm N4 has local quadratic convergence [64]. In Brown's method $P_k$ is an upper triangular elementary Gauss matrix. In the Brent method $P_k$ is an orthogonal matrix (e.g. a Householder matrix).

Jeney [50] observed that if the parameter matrices $W$ and $Z$ of the ABS methods are lower triangular then half of the operations can be saved when computing $P$ and the update matrices $H_k$. He developed and analyzed the following methods.

ALGORITHM N5 (Jeney's discretized ABS class)

$x_1 \approx x^*$, $i = 1$.

**while** $i > 0$

Choose stepsize $h_i \neq 0$.

$h = h_i$, $H_1 = I_m$, $y_1 = x_i$

**for** $k = 1 : m$

$$a_k = \frac{1}{h}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ f_k\left(y_k + hH_k^T e_k\right) - f_k(y_k) \\ f_k\left(y_k + hH_k^T e_{k+1}\right) - f_k(y_k) \\ \vdots \\ f_k\left(y_k + hH_k^T e_m\right) - f_k(y_k) \end{bmatrix}$$

Select $z_k$ such that the first $k-1$ components are zero and $z_k^T a_k \neq 0$.

$p_k = H_k^T z_k$

$y_{k+1} = y_k - f_k(y_k) p_k / (z_k^T a_k)$

Select $w_k$ such that the first $k-1$ components are zero and $w_k^T a_k \neq 0$.

$H_{k+1} = H_k - a_k w_k^T H_k / (w_k^T a_k)$

**end**

$x_{i+1} = y_{m+1}, \quad i = i + 1$

**end**

Jeney [50] showed that Algorithm N5 is equivalent with a subclass of Algorithm N4. Using a very extensive numerical testing on the 100 test problems of [63] Jeney [49] identified several competitive discretized ABS methods. Jeney [52] also defined a multistep version of the discretized ABS class.

ALGORITHM N6 (Jeney's multistep discretized ABS class)

$x_1 \approx x^*$, $i = 1$.

**while** $i > 0$

Choose stepsize $h_i \neq 0$.

$h = h_i$, $H_1 = I_m$, $y_1 = x_i$

**for** $k = 1 : m$

$$a_k = \frac{1}{h} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ f_k\left(y_k + hH_k^T e_k\right) - f_k\left(y_k\right) \\ f_k\left(y_k + hH_k^T e_{k+1}\right) - f_k\left(y_k\right) \\ \vdots \\ f_k\left(y_k + hH_k^T e_m\right) - f_k\left(y_k\right) \end{bmatrix}$$

Select $z_k$ such that the first $k-1$ components are zero and $z_k^T a_k \neq 0$.

$p_k = H_k^T z_k$

$y_{k+1} = y_k - f_k(y_k) p_k / (z_k^T a_k)$

Select $w_k$ such that the first $k-1$ components are zero and $w_k^T a_k \neq 0$.

$H_{k+1} = H_k - a_k w_k^T H_k / (w_k^T a_k)$

**end**

$x^{(1)} = y_{m+1}$

**for** $j = 1 : t$

$y_1^{(j)} = x^{(j)}$

**for** $k = 1 : m$

$y_{k+1}^{(j)} = y_k^{(j)} - f_k\left(y_k^{(j)}\right) p_k / (z_k^T a_k)$

**end**

$x^{(j+1)} = y_{m+1}^{(j)}$

**end**

$x_{i+1} = x^{(t+1)}, \quad i = i + 1$

**end**

Jeney [52] showed that Algorithm N6 is equivalent with a subclass of the multistep version of the Schmidt-Hoyer-Brown-Brent methods [42]. Jeney's computational results indicate that by increasing $t$ the average CPU time of iterations decreases significantly [52].

Spedicato, Deng and Chen [68] also introduced a class of difference ABS-type methods, which can be considered as a variant of Algorithm N4 [50].

## 5.3 Quasi-Newton ABS methods

Huang [48], [47], [46], Galántai, Jeney [34], [36] and Ge [39] suggested different combinations of the quasi-Newton updates and the ABS methods. Here we discuss the approach due to Galántai and Jeney [34], [36]. In this approach the Jacobian matrix $A$ is kept fixed within and updated outside the minor iteration loop using quasi-Newton updates.

ALGORITHM N7 (Iteration $i$ of the quasi-Newton ABS methods)

$y_1 = x_i,\ H_1 = I_m$
**for** $k = 1 : m$
$p_k = H_k^T z_k$
$y_{k+1} = y_k - v_k^T f(y_k)\, p_k / \left(v_k^T A_i p_k\right)$
$H_{k+1} = H_k - H_k A_i^T v_k w_k^T H_k / \left(w_k^T H_k A_i^T v_k\right)$
**end**
$x_{i+1} = y_{m+1}$
$s_i = x_{i+1} - x_i$
$Y_i = f(x_{i+1}) - f(x_i)$
$A_{i+1} = \phi(A_i, s_i, Y_i)$

The matrix $A_0$ is given and $A_i$ is updated such that $A_{i+1} s_i = Y_i$ $(i \geq 0)$. This algorithm is still expensive in terms of memory space and cost. However using the results of [24], [25], [26], [31], [32] it is possible to derive two competitive special cases, when matrices $H_k$ can be omitted and $P = [p_1, \ldots, p_m]$ is computable in $O\left(n^2\right)$ arithmetic operations [34], [36]. These methods and their multistep versions are the following.

ALGORITHM N8
$y_1 = x_i$
Calculate $P = A_i^{-1} V^{-T}$
**for** $k = 1 : m$
$y_{k+1} = y_k - v_k^T f(y_k)\, p_k$
**end**
$x_{i+1} = y_{m+1}$
$s_i = x_{i+1} - x_i$
$Y_i = f(x_{i+1}) - f(x_i)$
$A_{i+1} = \phi(A_i, s_i, Y_i)$

ALGORITHM N9
$y_1 = x_i$
Calculate $QR = A_i^T V$
$P = Q$
**for** $k = 1 : m$
$y_{k+1} = y_k - v_k^T f(y_k)\, p_k / r_{kk}$
**end**
$x_{i+1} = y_{m+1}$
$s_i = x_{i+1} - x_i$
$Y_i = f(x_{i+1}) - f(x_i)$
$A_{i+1} = \phi(A_i, s_i, Y_i)$

Algorithms N8 and N9 require $O\left(m^2\right)$ arithmetic operations and two function evaluations per step, if the Sherman-Morrison-Woodbury inversion formula or the fast $QR$ update [19] is used for calculating $P$ and $V$ is diagonal. The multistep versions of Algorithm N8 and N9 which save on the calculation of $P$

are defined as follows.

ALGORITHM N10

$y_1^{(1)} = x_i$
Calculate $P = A_i^{-1} V^{-T}$
**for** $j = 1 : t$
  **for** $k = 1 : m$
  $y_{k+1}^{(j)} = y_k^{(j)} - v_k^T f\left(y_k^{(j)}\right) p_k$
  **end**
  $y_1^{(j+1)} = y_{m+1}^{(j)}$
**end**
$x_{k+1} = y_{m+1}^{(t)}$
$s_i = x_{i+1} - x_i$
$Y_i = f(x_{i+1}) - f(x_i)$
$A_{i+1} = \phi(A_i, s_i, Y_i)$

ALGORITHM N11

$y_1^{(1)} = x_i$
Calculate $QR = A_i^T V$
$P = Q$
**for** $j = 1 : t$
  **for** $k = 1 : m$
    $y_{k+1}^{(j)} = y_k^{(j)} - v_k^T f\left(y_k^{(j)}\right) p_k / r_{kk}$
  **end**
  $y_1^{(j+1)} = y_{m+1}^{(j)}$
**end**
$x_{k+1} = y_{m+1}^{(t)}$
$s_i = x_{i+1} - x_i$
$Y_i = f(x_{i+1}) - f(x_i)$
$A_{i+1} = \phi(A_i, s_i, Y_i)$

The cost of one iteration for both algorithms is $O(tm^2)$ arithmetic operations plus $t + 1$ function evaluations, if $V$ is a diagonal matrix.

The local linear convergence of these methods were proved for the Broyden update in [27], [28], [29], [33]. The computer experiments indicated that Algorithms N8 and N9 are almost as good as the Broyden method based on inversion or QR factorization, respectively. The QR factorization based multistep algorithms are somewhat better than the multistep Broyden method using QR factorization. The inversion based multistep algorithms and the multistep Broyden method are about the same. These experiments were published in [34], [36], [35], [51].

Spedicato and Huang [71] reports numerical experiments with Newton-like methods and Huang's quasi-Newton ABS method [46].

# 6 ABS methods on special nonlinear systems

The intrinsic properties and the freedom of parameter selection make the ABS methods applicable to several problems. Spedicato and Huang [69] developed an application of nonlinear ABS methods to nonlinear least squares problems. Spedicato and Huang [70] also showed that the nonlinear ABS methods can be applied to nonlinear underdetermined systems with Jacobians of maximum row rank. Recently, Ge and Xia proved the convergence of ABS methods for singular nonlinear systems with rank defects [40], [41].

Here we present two applications that effectively handle the special sparsity structure of the given problems. The third application outlines a new way for the possible handling of the very ill-conditioned cases of the central trajectory method of linear programming.

## 6.1 ABS methods on bordered nonlinear systems

For special nonlinear systems the ABS methods can be quite effective. Consider the block bordered nonlinear systems of the form

$$\begin{array}{l} f_k(x_k, x_{q+1}) = 0, \quad k = 1, \ldots, q, \\ f_{q+1}(x_1, \ldots, x_{q+1}) = 0, \end{array} \tag{12}$$

where $x_k \in \mathbb{R}^{m_k}$, $f_k \in \mathbb{R}^{m_k}$ $(k = 1, \ldots, q+1)$, and $\sum_{k=1}^{q+1} m_k = m$. Let $\mathbf{x} = \left[x_1^T, \ldots, x_q^T, x_{q+1}^T\right]^T \in \mathbb{R}^m$ and

$$F(\mathbf{x}) = \left[f_1^T(x), \ldots, f_q^T(x), f_{q+1}^T(x)\right]^T \in \mathbb{R}^m. \tag{13}$$

Then the Jacobian matrix of system (12) has the block bordered or arrowhead structure

$$J(\mathbf{x}) = \begin{bmatrix} A_1 & & & & B_1 \\ & A_2 & & & B_2 \\ & & \ddots & & \vdots \\ & & & A_q & B_q \\ C_1 & C_2 & \cdots & C_q & D \end{bmatrix}, \tag{14}$$

where

$$A_k = \frac{\partial f_k(\mathbf{x})}{\partial x_k} \in \mathbb{R}^{m_k \times m_k} \ \ (k = 1, \ldots, q), \ D = \frac{\partial f_{q+1}(\mathbf{x})}{\partial x_{q+1}} \in \mathbb{R}^{m_{q+1} \times m_{q+1}}, \tag{15}$$

and

$$B_k = \frac{\partial f_k(\mathbf{x})}{\partial x_{q+1}} \in \mathbb{R}^{m_k \times m_{q+1}}, \ C_k = \frac{\partial f_{q+1}(\mathbf{x})}{\partial x_k} \in \mathbb{R}^{m_{q+1} \times m_k} \ \ (k = 1, \ldots, q). \tag{16}$$

Let

$$S(\mathbf{x}) = D(\mathbf{x}) - \sum_{k=1}^{q} C_k(\mathbf{x}) A_k^{-1}(\mathbf{x}) B_k(\mathbf{x}) \tag{17}$$

and $\mathbf{x}^i = \left[x_1^{iT}, \ldots, x_q^{iT}, x_{q+1}^{iT}\right]^T \in \mathbb{R}^m$. It can be shown [30] that Algorithm N1 takes the following form on problem (12).

ALGORITHM 12 (Iteration $i$ )
Step 1:
**for** $k = 1 : q$
  Solve $A_k\left(x^i\right) \triangle x_k = -f_k\left(x_k^i, x_{q+1}^i\right)$ for $\triangle x_k$.
  $x_k^{i,1} = x_k^i + \triangle x_k$
**end**
Step 2:
Calculate $S(x^i) = D(x^i) - \sum_{k=1}^{q} C_k\left(x^i\right) A_k^{-1}\left(x^i\right) B_k(x^i)$.
Solve $S\left(x^i\right) \triangle x_{q+1} = -f_{q+1}\left(x_1^{i,1}, \ldots, x_q^{i,1}, x_{q+1}^i\right)$ for $\triangle x_{q+1}$.

$x_{q+1}^{i+1} = x_{q+1}^{i} + \triangle x_{q+1}$
Step 3:
**for** $k = 1 : q$
$x_k^{i+1} = x_k^{i,1} - A_k^{-1}\left(x^i\right) B_k\left(x^i\right) \triangle x_{q+1}$
**end**

This form is identical with Method 2 of Hoyer and Schmidt [43] and the basic corrected implicit method of Zhang, Byrd and Schnabel [74]. The latter paper contains details of computer testing on a parallel computer.

## 6.2 Nonlinear ABS methods in constrained optimization

We investigate nonlinear optimization problems of the form

$$\begin{array}{l} f(x) \to \min \\ h_j(x) = 0, \quad j \in J = \{1, 2, \ldots, p\}, \\ g_i(x) \leq 0, \quad i \in I = \{1, 2, \ldots, m\}, \end{array} \tag{18}$$

where $f, g_i, h_j : \mathbb{R}^n \to \mathbb{R}$ $(i \in I,\, j \in J)$ are smooth enough. Let

$$L(x, \mu, \lambda) = f(x) + \sum_{j \in J} \mu_j h_j(x) + \sum_{i \in I} \lambda_i g_i(x)$$

and $z = \left[x^T, \mu^T, \lambda^T\right]^T$. A point $z^* = \left[x^{*T}, \mu^{*T}, \lambda^{*T}\right]^T$ is said to be a Kuhn-Tucker point (KTP) if it satisfies

$$\begin{array}{l} \nabla_x L(x, \mu, \lambda) = 0, \\ h_j(x) = 0 \quad (j \in J), \\ g_i(x) \leq 0 \quad (i \in I), \\ \lambda_i g_i(x) = 0 \quad (i \in I), \\ \lambda_i \geq 0 \quad (i \in I). \end{array} \tag{19}$$

Under a regularity assumption, conditions (19) are necessary for the optimality of $x^*$ in optimization problem (18). There are several methods to solve (19), especially in the case of $I = \emptyset$. For details and references we refer to [53].

**Definition 6** *A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is called NCP-function if it satisfies the complementarity condition*

$$\phi(a, b) = 0 \Leftrightarrow a \geq 0,\ ab = 0,\ b \leq 0.$$

Everywhere differentiable NCP-functions are

$$\phi(a, b) = ab + \frac{1}{2}\left(\max\{0, b - a\}\right)^2, \qquad \text{(Evtushenko-Purtov)}$$

$$\phi(a, b) = ab + \left(\max\{0, a\}\right)^2 + \left(\min\{0, b\}\right)^2, \qquad \text{(Evtushenko-Purtov)}$$

$$\phi(a, b) = (a + b)^2 + b|b| - a|a|. \qquad \text{(Mangasarian)}$$

Using any complementarity function one can write the Kuhn-Tucker conditions (19) as an equivalent nonlinear system $F_\phi(z) = 0$, where $F_\phi : \mathbb{R}^{n+p+m} \to \mathbb{R}$ is defined by

$$F_\phi(z) = \begin{bmatrix} \nabla_x L(x, \mu, \lambda) \\ h_1(x) \\ \vdots \\ h_p(x) \\ \phi(\lambda_1, g_1(x)) \\ \vdots \\ \phi(\lambda_m, g_m(x)) \end{bmatrix} = 0. \tag{20}$$

This kind of equivalence was first given by Mangasarian [55], who also gave the first technique to construct NCP-functions.

Under the following assumptions Kanzow and Kleinmichel [53] defined a class of Newton-type methods for solving the nonlinear system

$$F_\phi(z) = 0. \tag{21}$$

Let $z^* = \left[x^{*T}, \mu^{*T}, \lambda^{*T}\right]^T$ be a KTP of problem (18).

(A.1) The functions $f$, $g_i$ and $h_j$ ($i \in I$, $j \in J$) are twice differentiable with Lipschitz-continuous second derivatives in a neighborhood of $x^*$.

(A.2) The gradients $\nabla g_i(x^*)$ ($i \in I^{**}$) and $\nabla h_j(x^*)$ ($j \in J$) are linearly independent, where $I^{**} = \{i \in I \mid \lambda_i^* > 0\}$.

(A.3) $y^T \nabla^2_{xx} L(x^*, \mu^*, \lambda^*)\, y > 0$ for all $y \in \mathbb{R}^n$, $y \neq 0$, $y^T \nabla g_i(x^*) = 0$ ($i \in I^{**}$) and $y^T \nabla h_j(x^*) = 0$ ($j \in J$).

(A.4) $I^* = I^{**}$, where $I^* = \{i \in I \mid g_i(x^*) = 0\}$.

(A.5) The NCP-function $\phi$ satisfies

$$\begin{aligned} &\tfrac{\partial \phi}{\partial a}(\lambda_i^*, g_i(x^*)) = 0 \quad (i \in I^{**}), \\ &\tfrac{\partial \phi}{\partial b}(\lambda_i^*, g_i(x^*)) \neq 0 \quad (i \in I^{**}), \\ &\tfrac{\partial \phi}{\partial a}(\lambda_i^*, g_i(x^*)) \neq 0 \quad (i \notin I^{**}), \\ &\tfrac{\partial \phi}{\partial b}(\lambda_i^*, g_i(x^*)) = 0 \quad (i \notin I^{**}). \end{aligned} \tag{22}$$

Kanzow and Kleinmichel [53] proved the following result.

**Theorem 7** *(Kanzow and Kleinmichel). Let $z^* = \left[x^{*T}, \mu^{*T}, \lambda^{*T}\right]^T$ be a KTP of problem (18). Suppose that the assumptions (A.1)-(A.5) hold at $z^*$. Let $\phi : \mathbb{R}^2 \to \mathbb{R}$ be a continuously differentiable NCP-function. Then the Jacobian matrix $F'_\phi(z^*)$ is nonsingular.*

The Jacobian matrix $F'_\phi(z)$ is given by

$$\left[\begin{array}{c|ccc|ccc} \nabla^2_{xx}L(x,\mu,\lambda) & \nabla h_1(x) & \dots & \nabla h_p(x) & \nabla g_1(x) & \dots & \nabla g_m(x) \\ \hline \nabla h_1(x)^T & & & & & & \\ \vdots & & 0 & & & 0 & \\ \nabla h_p(x)^T & & & & & & \\ \hline \left(\frac{\partial\phi}{\partial b}\right)_1 \nabla g_1(x)^T & & & & \left(\frac{\partial\phi}{\partial a}\right)_1 & & \\ \vdots & & 0 & & & \ddots & \\ \left(\frac{\partial\phi}{\partial b}\right)_m \nabla g_m(x)^T & & & & & & \left(\frac{\partial\phi}{\partial a}\right)_m \end{array}\right],$$

where

$$\begin{aligned} \left(\frac{\partial\phi}{\partial a}\right)_i &= \frac{\partial\phi}{\partial a}(\lambda_i, g_i(x)) \quad (i \in I), \\ \left(\frac{\partial\phi}{\partial b}\right)_i &= \frac{\partial\phi}{\partial b}(\lambda_i, g_i(x)) \quad (i \in I). \end{aligned} \tag{23}$$

This Jacobian matrix has the block structure

$$F'_\phi(z) = \begin{bmatrix} A & C & B \\ C^T & 0 & 0 \\ D & 0 & E \end{bmatrix}, \tag{24}$$

where the diagonal matrix $E$ may become singular. The Newton-type methods suggested by Kanzow and Kleinmichel [53] take special care of this singularity by first solving a reduced linear system and then making a correction step. Note that in [53] the inequality constraints precede the equations $h_i(x) = 0$ and the Jacobian $F'_\phi(z)$ is permutationally similar to (24).

Here we suggest another approach using Algorithm N1 with parameters $V = Z = W = I$. For this special case it can be shown (similarly to Theorem 2) that it is locally convergent if the Jacobian matrix at the solution is strongly block nonsingular (i.e., block LU-decomposable). The following result indicates that it is indeed the case and by applying the special ABS method we can avoid the handling of the zero diagonals in $E$.

**Lemma 8** *If $F'_\phi(z^*)$, $A$ and $C^TA^{-1}C$ are nonsingular, then $F'_\phi(z^*)$ is block strongly nonsingular and its LU-decomposition is given by*

$$\begin{bmatrix} I & 0 & 0 \\ C^TA^{-1} & I & 0 \\ DA^{-1} & DA^{-1}C\left(C^TA^{-1}C\right)^{-1} & I \end{bmatrix} \begin{bmatrix} A & C & B \\ 0 & -C^TA^{-1}C & -C^TA^{-1}B \\ 0 & 0 & W \end{bmatrix},$$

*where $W = E - DA^{-1}B + DA^{-1}C\left(C^TA^{-1}C\right)^{-1}C^TA^{-1}B$ is also nonsingular.*

The proof of the Lemma is easy by direct calculation. Coleman and Fenyes [12] points out that many authors assume that $A = \nabla^2_{xx}L(z^*)$ is positive definite. In such a case, $C^TA^{-1}C$ is also positive definite, provided that $C$ has maximum column rank. For other assumptions, see [60], [12] and [53].

Consider now one step of Algorithm N1 ($V = Z = W = I$) on the partitioned system

$$F_\phi(z) = \begin{bmatrix} F_1(x,\mu,\lambda) \\ F_2(x) \\ F_3(x,\lambda) \end{bmatrix},$$

where $F_1(z) = F_1(x,\mu,\lambda) = \nabla_x L(x,\mu,\lambda)$,

$$F_2(x) = \begin{bmatrix} h_1(x) \\ \vdots \\ h_p(x) \end{bmatrix}, \quad F_3(x,\lambda) = \begin{bmatrix} \phi(\lambda_1, g_1(x)) \\ \vdots \\ \phi(\lambda_m, g_m(x)) \end{bmatrix}$$

and $r = 3$ ($n_1 = n$, $n_2 = p$, $n_3 = m$). We also partition $u^i$ accordingly, that is

$$u^i = \begin{bmatrix} u_1^i \\ u_2^i \\ u_3^i \end{bmatrix},$$

where $u_1^i \in \mathbb{R}^n$, $u_2^i \in \mathbb{R}^p$ and $u_3^i \in \mathbb{R}^m$. By direct calculation and the choice $\tau_{ji} = 0$ ($j > 1$) we have the following special algorithm:
Step 1:
$u^1 = z^k$,

$$\begin{bmatrix} A & C & B \\ C^T & 0 & 0 \\ D & 0 & E \end{bmatrix} = F'_\phi(u^1);$$

Step 2:

$$\begin{aligned} u_1^2 &= u_1^1 - A^{-1}F_1(u^1), \\ u_2^2 &= u_2^1, \\ u_3^2 &= u_3^1; \end{aligned}$$

Step 3:

$$\begin{aligned} u_1^3 &= u_1^2 - A^{-1}C\left(C^TA^{-1}C\right)^{-1}F_2(u_1^2), \\ u_2^3 &= u_2^2 + \left(C^TA^{-1}C\right)^{-1}F_2(u_1^2), \\ u_3^3 &= u_3^2; \end{aligned}$$

Step 4:

$$W = E - DA^{-1}B + DA^{-1}C\left(C^TA^{-1}C\right)^{-1}C^TA^{-1}B,$$

$$\begin{aligned} u_1^4 &= u_1^3 - \left(-A^{-1}B + A^{-1}C\left(C^TA^{-1}C\right)^{-1}C^TA^{-1}B\right)W^{-1}F_3\left(u_1^3, u_3^3\right), \\ u_2^4 &= u_2^3 + \left(C^TA^{-1}C\right)^{-1}C^TA^{-1}BW^{-1}F_3\left(u_1^3, u_3^3\right), \\ u_3^4 &= u_3^3 - W^{-1}F_3\left(u_1^3, u_3^3\right), \end{aligned}$$

Step 5:
$z^{k+1} = u^4$

It is clear that we can avoid inversions when implementing the algorithm. Preliminary numerical testing in MATLAB indicates that the algorithm works well if $C^T A^{-1} C$ is well-conditioned.

### 6.3 An ABS application to the primal-dual interior programming method of linear programming

Consider the standard linear programming problem

$$c^T x \to \min, \quad Ax = b,\ x \geq 0,\ (A \in \mathbb{R}^{m \times n})$$

and its dual problem

$$b^T y \to \max, \quad A^T y + s = c,\ s \geq 0.$$

The central trajectory method is related to the solution of the penalized LP:

$$c^T x - \mu \sum_{j=1}^{n} \ln(x_j) \to \min, \quad Ax = b,\ x > 0,$$

where $\mu > 0$ is a parameter ($\mu \to 0$). The Karush-Kuhn-Tucker conditions for this problem are

$$\begin{aligned} &Ax = b, \quad x > 0 \\ &c - \mu X^{-1} e = A^T y, \end{aligned}$$

where $X = \text{diag}(x_1, \ldots, x_n)$ and $e = [1, 1, \ldots, 1]^T$. Defining $s = \mu X^{-1} e$ and $S = \text{diag}(s_1, \ldots, s_n)$ one can rewrite the KTP equations in the form

$$\begin{aligned} &Ax = b, \quad x > 0 \\ &A^T y + s = c, \quad s > 0 \\ &XSe - \mu e = 0. \end{aligned} \tag{25}$$

Denoting by $(x(\mu), y(\mu), s(\mu))$ the solution of (25) for a given $\mu > 0$ we call the set $\Gamma = \{(x(\mu), y(\mu), s(\mu)) \mid \mu > 0\}$ the central trajectory (path) of the primal linear program. Equation (25) has a special structure, where nonlinearity occurs only in the third subset of equations in the form $x_i s_i = \mu$ $(i = 1, \ldots, n)$. In primal-dual algorithms, steps are generated by applying the Newton method to the three equalities of (25) with obvious restrictions on the iterates $\left(x^k, y^k, s^k\right)$, that is with $\left(x^k, y^k\right) > 0$. The search direction of the Newton step at the point $(x, y, s)$ is defined by

$$\begin{bmatrix} 0 & A & 0 \\ A^T & 0 & I \\ 0 & S & X \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \\ \Delta s \end{bmatrix} = - \begin{bmatrix} r_P \\ r_D \\ XSe - \sigma\mu e \end{bmatrix},$$

where $r_P = Ax - b$, $R_D = A^T y + s - c$, $\mu = x^T s/n$ and $\sigma \in [0, 1]$ is a governing parameter. Eliminating both $\Delta s$ and $\Delta x$, one has the so called normal equation

$$AD^2 A^T \Delta y = t \quad \left(D^2 = S^{-1} X\right),$$

which is often solved by a Cholesky algorithm or a sparse Cholesky algorithm if $A$ is sparse. Note that $A$ is fixed, while $D$ and $t$ change with the Newton iterations.

Assume that $m < n$ and $\text{rank}(A) = m$. The normal equation $AD^2 A^T x = t$ can be written in the form

$$ADy = t, \quad DA^T x = y. \tag{26}$$

The minimal least squares solution of the first equation $y = (AD)^+ t$ belongs to $\text{Range}(AD)^T$. Hence the second equation has a unique solution.

So the suggested solution scheme is to solve $ADy = t$ first, then solve $DA^T x = y$. The first equation can be solved with the Huang method starting from the 0 vector. It gives the requested minimal least squares solution at the cost of at least $mn^2$ flops (1 flop=1 arithmetic operations of the type $+, -, \times, /$).

The system $DA^T x = y$ can be written as $A^T x = D^{-1} y$. The latter can be solved with any ABS method. The ABS method also detects $m$ independent equations. Since the latter system matrix is constant we can reuse the search vectors $p_i$ of the first iteration in the subsequent Newton steps. Thus the cost of solving $A^T x = D^{-1} y$ is $4m^2 + m$ flops per iteration.

Altogether the cost of the suggested solution scheme is at least $mn^2 + 4m^2 + m$ flops per iteration. Since $n > m$, this cost is at least $m^3 + 4m^2 + m$ flops per iteration. The forming of $AD^2 A^T$ requires $O\left(m^2 n/2\right)$ flops per iteration. The cost of (dense) Cholesky is $(1/3)\, m^3 + 2m^2$ flops. The Cholesky decomposition is the winner by a factor 2 in complexity terms if $m$ is small relative to $n$. If $m$ is close to $n$ the difference is small. However there are three reasons in favor of the suggested approach. First, it uses the matrix $AD$ ($A$) the condition number of which is the square root of the condition number of $AD^2 A^T$ ($AA^T$). Second, as we pointed out in the first section, the Huang method is extremely stable in floating point arithmetic. For very ill-conditioned and numerically unstable cases the suggested approach can be useful although a full scale computational testing is yet to be performed. The third reason is that the Huang method proved to be very competitive on vector computers.

# References


[1] Abaffy J.: A Superlinear Convergency Theorem in the ABSg Class for Nonlinear Algebraic Equations, JOTA, 59, 1988, pp. 39–43

[2] Abaffy J.: Superlinear Convergence Theorems for Newton-Type Methods for Nonlinear Systems of Equations, JOTA, 73, 1992, pp. 269–277

[3] Abaffy, J., Broyden, C.G., Spedicato, E.: A class of direct methods for linear systems, *Numer. Math.* 45, 1984, pp. 361-376

[4] Abaffy J., Galántai A.: Conjugate Direction Methods for Linear and Nonlinear Systems of Algebraic Equations, Colloquia Mathematica Soc. János Bolyai, 50. Numerical Methods, Miskolc (Hungary) 1986. (ed. P. Rózsa) North-Holland, Amsterdam, 1987, pp. 481–502

[5] Abaffy J., Galántai A., Spedicato, E.: The Local Convergence of ABS Methods for Nonlinear Algebraic Equations, Numer. Math. 51, 1987, pp. 429–439

[6] Abaffy J., Spedicato, E.: ABS-projection Algorithms: Mathematical Techniques for Linear and Nonlinear Algebraic Equations, Ellis Horwood, Chichester, 1989

[7] Brent, R.P.: Some Efficient Algorithms for Solving Systems of Nonlinear Equations, SIAM J. Numer. Anal. 10, 1973, pp. 327–344

[8] Brown, K.M.: A Quadratically Convergent Method for Solving Simultaneous Nonlinear Equations, Ph.D. thesis, Purdue University, Lafayette, Indiana, 1966

[9] Brown, K.M.: A Quadratically Convergent Newton-Like Method Based upon Gaussian-Elimination, SIAM J. Numer. Anal., 6, 1969, pp. 560–569

[10] Bus, J.C.P.: Numerical Solution of Systems of Nonlinear Equations, Matematisch Centrum, Amsterdam, 1980

[11] Chen, Z., Deng, N., Spedicato, E.: Truncated Difference ABS-Type Algorithm for Nonlinear Systems of Equations, Optimization Methods and Software, 13, 2000, pp. 263–274

[12] Coleman, T.F., Fenyes, P.A.: Partitioned quasi-Newton methods for nonlinear equality constrained optimization, Mathematical Programming 53, 1992, 17–44

[13] Deng, N., Chen, Z.: Truncated Nonlinear ABS Algorithm and Its Convergence Property, Computing 45, 1990, pp. 169–173

[14] Deng, N.Y., Chen, Z.: A Unified Convergence Theorem for Nonlinear ABS-Type Algorithms, Proceedings of the First International Conference on ABS Algorithm, Luoyang, China, September 2-6, 1991 (ed. Spedicato, E.), University of Bergamo, Bergamo, 1992, pp. 22-30

[15] Deng N. and Chen Z.: A New Nonlinear ABS-Type Algorithm and its Efficiency Analysis, Optimization Methods and Software 10, 1998, pp. 71–85

[16] Deng, N., Spedicato, E., Zhu, M.: A Local Convergence Theorem for Nonlinear ABS Algorithms, Computational and Applied Mathematics 13, 1994, pp. 49–59

[17] Deng, N., Zhu, M.: Further Results on the Local Convergence of the Nonlinear ABS Algorithms, Optimization Methods and Software, 1, 1992, pp. 211–221

[18] Deng, N., Zhu, M.: The Local Convergence Property of the Nonlinear Voevodin-ABS Algorithm, Optimization Methods and Software, 1, 1992, pp. 223–231

[19] Dennis, J.E., Schnabel, R.B.: Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice-Hall, Englewood Cliffs, New Jersey, 1983

[20] Frommer, A.: Monotonie und Einschließung beim Brown-Verfahren, doctoral thesis, Inst. für Angewandte Mathematik, University of Karlsruhe, 1986

[21] Frommer, A.: Comparison of Brown's and Newton's Method in the Monotone Case, Numer. Math. 52, 1988, pp. 511–521

[22] Galántai A.: Block ABS Methods for Nonlinear Systems of Algebraic Equations, Computational Solution of Nonlinear Systems of Equations, (eds.: E.L. Allgower, K. Georg) SIAM Lectures in Applied Mathematics 26, 1990, pp. 181–187

[23] Galántai A.: Convergence Theorems for the Nonlinear ABS-Methods, Colloquia Mathematica Soc. János Bolyai, 59. Numerical Methods, Miskolc (Hungary) 1990. (ed. P. Rózsa), North-Holland, Amsterdam-New-York, 1991, pp. 65–79.

[24] Galántai A.: Generalized Implicit LU Algorithms in the Class of ABS Methods for Linear and Nonlinear Systems of Algebraic Equations, Quaderno DMSIA 93/5, University of Bergamo, Bergamo, 1993

[25] Galántai A.: The Global Convergence of the ABS Methods for a Class of Nonlinear Problems, Optimization Methods and Software, 4, 1995, pp. 283–295

[26] Galántai A.: ABS Methods for the Solution of Linear and Nonlinear Systems of Algebraic Equations, Habilitation dissertation, Technical University of Budapest, Budapest, 1994, pp. 115

[27] Galántai A.: A Fast Method for Solving Systems of Nonlinear Algebraic Equations, Publ. Univ. of Miskolc, Series C. Mechanical Engineering, 45, 1, 1995, pp. 85–92

[28] Galántai A.: Hybrid Quasi-Newton ABS Methods for Nonlinear Equations, in F.J. Cobos, J.R. Gómez, F. Mateos (eds.): EAMA97 Meeting on Matrix Analysis and Applications, Sevilla, September 10-12, 1997, pp. 221–228

[29] Galántai A.: The Local Convergence of Quasi-Newton ABS Methods, Publ. Univ. of Miskolc, Series D. Natural Sciences, 37., Mathematics, 1997, pp. 31–38

[30] Galántai A.: Parallel ABS Projection Methods for Linear and Nonlinear Systems with Block Arrowhead Structure, Computers and Mathematics with Applications 38, 1999, 11–17

[31] Galántai A.: Rank Reduction and Conjugation, Acta Technica Acad. Sci. Hung. 108, 1-2, 1997-99, pp. 107–130

[32] Galántai A.: Rank Reduction, Factorization and Conjugation, Linear and Multilinear Algebra, 49, 2001, pp. 195–207

[33] Galántai A.: Projectors and Projection Methods, Kluwer, 2004

[34] Galántai A., Jeney A.: Quasi-Newton ABS Methods, In MicroCAD'93 Conference, Section Modern Numerical Methods, Miskolc, Hungary, 1993, University of Miskolc, 1993, pp. 63–68

[35] Galántai, A., Jeney, A., Spedicato, E.: Testing of ABS-Huang Methods on Moderately Large Nonlinear Systems, Quaderno DMSIA 93/6, University of Bergamo, 1993

[36] Galántai A., Jeney, A.: Quasi-Newton ABS Methods for Solving Nonlinear Algebraic Systems of Equations, JOTA, 89, 1996, pp. 561–573

[37] Gáti A.: Automatic Error Analysis with Miller's Method, Miskolc Mathematical Notes, 5, 1, 2004, pp. 25–32

[38] Gay, D.M.: Brown's Method and Some Generalizations with Application to Minimization Problems, Cornell University, Comput. Sci. Techn. Rep., 75-225, 1975

[39] Ge, R.: A Family of Broyden-ABS Type Methods for Solving Nonlinear Systems, Quaderno DMSIA 97/1, University of Bergamo, Bergamo, 1997

[40] Ge, R.-D., Xia, Z.-Q.: An ABS Algorithm for Solving Singular Nonlinear Systems with Rank One Defect, The Korean Journal of Computational & Applied Mathematics, 9, 1, 2002, pp. 167–183

[41] Ge, R.-D., Xia, Z.-Q.: An ABS Algorithm for Solving Singular Nonlinear Systems with Rank Defects, The Korean Journal of Computational & Applied Mathematics, 12, 1-2, 2003, pp. 1–20

[42] Hoyer, W.: Zur Effektivität mehrstufiger Brown-Brent-Verfahren, Beiträge zur Numer. Math. 10, 1981, 57–69

[43] Hoyer, W., Schmidt, J.W.: Newton-Type Decomposition Methods for Equations Arising in Network Analysis, ZAMM, 64, 9, 1981, pp. 397–405

[44] Huang Z.: Multi-step Nonlinear ABS Methods and Their Efficiency Analysis, Computing, 46, 1991, pp. 143–153

[45] Huang, Z.: Convergence Analysis of the Nonlinear Block Scaled ABS Methods, JOTA, 75, 2, 1992, pp. 331–344

[46] Huang Z.: Row Update ABS Methods for Solving Sparse Nonlinear Systems of Equations, Optimization Methods and Software, 2, 1993, pp. 297–309.

[47] Huang Z.: Restart Row Update ABS Methods for Systems of Nonlinear Equations, Computing, 50, 1993, pp. 229–239.

[48] Huang Z.: Row Update ABS Methods for Systems of Nonlinear Equations, Optimization Methods and Software, 3, 1994, pp. 41–60

[49] Jeney A.: Nemlineáris egyenletrendszerek megoldására szolgáló ABS-módszerek numerikus vizsgálata, Alkalmazott Matematikai Lapok, 15, 3–4, 1990-91, 353–364 (in Hungarian)

[50] Jeney A.: Nemlineáris egyenletrendszerek megoldására szolgáló ABS-módszerek egy részosztályának diszkretizációja, Alkalmazott Matematikai Lapok, 15, 3–4, 1990-91, 365–379 (in Hungarian)

[51] Jeney A.: Numerical Comparison of Multistep Quasi-Newton ABS Methods with the Multistep Broyden Method for Solving Systems of Nonlinear Equations, Publ. Univ. of Miskolc, Series D. Natural Sciences. 36, 2, Mathematics, 1996, pp. 47–54

[52] Jeney A.: Multistep Discrete ABS Methods for Solving Systems of Nonlinear Equations, Publ. Univ. of Miskolc, Series D. Natural Sciences. 37. Mathematics, 1997, 39–46

[53] Kanzow, C., Kleinmichel, H.: A class of Newton-type methods for equality and inequality constrained optimization, Optimization Methods and Software, 5, 1995, 173–198

[54] Li, Z. Deng, N., Wang, L.: ABS-Type Methods with Ordering Strategies for Solving Nonlinear Equations, Quaderno DMSIA 96/12, University of Bergamo, Bergamo, 1996.

[55] Mangasarian, O.L.: Equivalence of the complementarity problem to a system of nonlinear equations, SIAM J. Appl. Math., 31, 1976, 89–92

[56] Martinez, J.M.: Generalization of the Methods of Brent and Brown for Solving Nonlinear Simultaneous Equations, SIAM J. Numer. Anal., 16, 1979, pp. 434–448

[57] Miller, W.: Software for Roundoff Analysis, ACM TOMS 1, 1975, pp. 108–128

[58] Moré, J.J., Cosnard, M.Y.: Numerical Solution of Nonlinear Equations, ACM TOMS 5, 1979, pp. 64–85

[59] Moré, J.J., Cosnard, M.Y.: Algorithm 554, BRENTM, A Fortran Subroutine for the Numerical Solution of Systems of Nonlinear Equations, ACM TOMS 6, 1980, pp. 240–251

[60] Nocedal, J., Overton, M.L.: Projected Hessian updating algorithms for nonlinearly constrained optimization, SIAM J. Numer. Anal. 22, 1985, 821–850

[61] Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970

[62] Saad, Y.: Iterative Methods for Sparse Linear Systems, PWS Publishing Company, 1996

[63] Roose, A., Kulla, V., Lomp, M., Meressoo, T.: Test Examples of Systems of Nonlinear Equations, Version 3-90, Estonian Software and Computer Service Company, Tallin, 1990

[64] Schmidt, J.W., Hoyer, W.: Ein konvergenzsatz für Verfahren vom Brown-Brent-typ, ZAMM, 57, 1977, pp. 397–405

[65] Schmidt, J.W., Hoyer, W.: Die Verfahren vom Brown-Brent-typ bei gemischt linearen-nichtlinearen Gleichungssystemen, ZAMM, 58, 1978, pp. 425–428

[66] Schwetlick, H.: Numerische Lösung nichtlinearer Gleichungen, VEB Deutscher Verlag der Wissenschaften, Berlin, 1979

[67] Spedicato, E., Bodon, E., Del Popolo, A., Xia, Z.: ABS algorithms for linear systems and optimization: a review and a bibliography, Ricerca Operativa 89, 2000, pp. 39-88

[68] Spedicato, E., Chen, Z., Deng, N.: A Class of Difference ABS-type Algorithms for a Nonlinear System of Equations, Numerical Linear Algebra with Applications 13, 1994, 313–329

[69] Spedicato, E., Huang, Z.: An Orthogonally Scaled ABS Method for Nonlinear Least Squares Problems, Optimization Methods and Software, 1, 1992, pp. 233–242

[70] Spedicato, E., Huang, Z.: Optimally Stable ABS Methods for Nonlinear Underdetermined Systems, Optimization Methods and Software, 5, 1995, pp. 17–26

[71] Spedicato, E., Huang, Z.: Numerical Experience with Newton-like Methods for Nonlinear Algebraic Systems, Computing, 58, 1997, 69–89

[72] Stewart, G.W.: Conjugate Direction Methods for Solving Systems of Linear Equations, Num. Math., 21, 1973, pp. 285–297

[73] Ypma, T.J.: On Nonlinearity and Eliminating Linear Components from Nonlinear Systems, Nonlinear Analysis, Theory, Methods and Applications, 8, 1984, 757–763

[74] Zhang, X., Byrd, R.H., Schnabel, R.B.: Parallel Methods for Solving Nonlinear Block Bordered Systems of Equations, SIAM J. Scientific and Statistical Comput. 13, 1992, pp. 841–859

[75] Zhu, M.: A Note on the Condition for Convergence of Nonlinear ABS Method, Calcolo 28, 1991, pp. 307–314

**Acknowledgements**

Work partially supported by GNCS Visiting Professors Program (A. Galantai) and by 2007 University of Bergamo Academic Funds (E. Spedicato)

**The authors' addresses:**

A. Galántai
Budapest Tech, John von Neumann Faculty of Informatics,
1034 Budapest, Bécsi u. 96/b, Hungary
e-mail: galantai.aurel@nik.bmf.hu

E. Spedicato
University of Bergamo, Department of Mathematics,
24127 Bergamo, Via dei Caniana, 2, Italy
e-mail: emilio.spedicato@unibg.it

**Redazione**

Dipartimento di Matematica, Statistica, Informatica ed Applicazioni
Università degli Studi di Bergamo
Via dei Caniana, 2
24127 Bergamo
Tel. 0039-035-2052536
Fax 0039-035-2052549

La Redazione ottempera agli obblighi previsti dall'art. 1 del D.L.L. 31.8.1945, n. 660 e successive modifiche

Stampato nel 2007
presso la Cooperativa
Studium Bergomense a r.l.
di Bergamo