Local convergence analysis of TR1 updates for solving nonlinear equations*

Sebastian Schlenkrich[†1], Andreas Griewank[2], and Andrea Walther[1]

[1]Institut für Wissenschaftliches Rechnen, TU Dresden
[2]Institut für Mathematik, HU Berlin

April 20, 2006

Abstract

For the solution of nonlinear equation systems, quasi-Newton methods based on low-rank updates are of particular interest. We analyze a class of TR1 update formulas to approximate the system Jacobian. The local q-superlinear convergence for nonlinear problems is proved for a particular subclass of updates. Moreover we give an estimate of the r-order of convergence. Numerical results comparing the TR1 method to Newton's and other quasi-Newton methods are presented.

Keywords: nonlinear equations, quasi-Newton methods, adjoint based update, Automatic Differentiation

1 Introduction

Many computational tasks involve the solution of a set of nonlinear simultaneous equations. This can be expressed as finding $x_* \in \mathbb{R}^n$ with $F(x_*) = 0 \in \mathbb{R}^n$ for a function $F : \mathbb{R}^n \to \mathbb{R}^n$. Since F is assumed to be nonlinear, frequently an iterative method has to be applied to compute a solution.

With reasonable assumptions on F a solution is found by Newton's method for any initial iterate sufficiently close to x_*, see e.g. [DS96]. The rate of convergence is quadratic. However Newton's method requires the repeated evaluation and factorization of the Jacobian F' at the state iterates. This can cause difficulties for example if a computational description of $F'(x)$ is not available. Furthermore for functions with a dense Jacobian, the effort for the computation of a new iterate is of cubic order in the dimension n, which is often unacceptable especially for large numbers of n.

*Partially supported by the DFG Research Center MATHEON "Mathematics for Key Technologies", Berlin

†Corresp. author: e-mail: sebastian.schlenkrich@tu-dresden.de, Fax: +49-351-463 37096

Quasi-Newton methods avoid the repeated evaluation and factorization of the Jacobian by maintaining a possibly factorized approximation of the Jacobian. One very simple approach is to freeze A_i and its factorization, for instance by setting $A_i = A := F'(x_0)$. This enables a very fast computation of iterates within a small multiple of n^2 operations (depending on the factorization used). However, the rate of convergence is at most linear as long as $A \neq F'(x_*)$. Another approach is to update A_i by a low-rank matrix and improve the approximation successively. Of particular interest is Broyden's update formula [Bro65] or the method of Gay and Schnabel using projected updates [GS78]. These rank-1 updates produce q-superlinear convergent iterations. Since they add a rank-1 matrix to an existing factorized approximation, the new factorization can be computed within $O(n^2)$ operations. Important features of these update methods are that they obey secant conditions and the so-called least change property.

The *forward* and *reverse mode* of *Automatic Differentiation (AD)* provide the possibility to compute $F'(x)u$ and $v^T F'(x)$ exact within machine accuracy for given vectors x, u, and v. The computational effort for each of these products is equal to the evaluation of F times a constant $c \leq 4$ independent of the dimension n of the state space. For further details on AD we refer to [Gri00] or *www.autodiff.org*. In this paper, the vector-Jacobian and Jacobian-vector products will be used, such that the considered rank-1 updates may also fulfill tangent conditions. For this purpose we analyze an alternative quasi-Newton update method proposed first in [GW02] in the context of equality constrained optimization. A similar update procedure was also considered in [Hab04] for a parameter estimation problem.

This paper has the following structure: Section 2 describes the update formula. In Section 2.1 we introduce the new class of *adjoint tangent rank-1 updates*. Based on the bounded deterioration property, we show local linear convergence in Theorem 8. Furthermore we can prove a transposed Dennis-Moré property in Lemma 11. For an important subclass, characterized by the *residual property*, we show q-superlinear convergence in Theorem 13. Section 2.3 illustrates the relation between the adjoint tangent rank-1 update and the TR1 update. With additional assumptions on the iteration we characterize the convergence more precisely and estimate the r-order of convergence in Theorem 17. Examples for adjoint tangent rank-1 updates are discussed in Section 2.4. In Section 3 some implementation details of the new methods are described. Furthermore numerical results comparing the adjoint tangent rank-1 updates to Newton's and other quasi-Newton methods are shown. Finally we give some concluding remarks in Section 4.

2 Local convergence with the TR1 update

A general framework for solving $F(x) = 0$ using a quasi-Newton approximation of $F'(x)$ is given in the following algorithm.

Algorithm 1 (Quasi-Newton Algorithm) *Suppose* $F : \mathbb{R}^n \to \mathbb{R}^n$, $\{A_i\}_{i\in\mathbb{N}_0} \subset \mathbb{R}^{n\times n}$ *with* A_i *non-singular,* $x_0 \in \mathbb{R}^n$, *and* $x_* \in \mathbb{R}^n$ *with* $F(x_*) = 0$ *are given, then the algorithm*

$$\begin{array}{rcl} s_i & = & -A_i^{-1}F(x_i) \\ x_{i+1} & = & x_i + s_i \end{array} \quad \text{for} \quad i = 0, 1, 2, \ldots$$

is called quasi-Newton method to find x_*.

2.1 The TR1 update formula

The *two-sided rank-1 (TR1)* update formula was introduced in [GW02] as a generalization of the symmetric rank-1 (SR1) update formula [CGT91] which is used in optimization to approximate Hessians. In a general form, the TR1 update formula is given by:

Definition 2 (TR1 update) *Let* $F : \mathbb{R}^n \to \mathbb{R}^n$ *be differentiable,* $x_{i+1} \in \mathbb{R}^n$. *For a given matrix* $A_i \in \mathbb{R}^{n\times n}$ *and given directions* $s_i, \sigma_i \in \mathbb{R}^n$, *the formula*

$$A_{i+1} = A_i + \frac{(F'(x_{i+1}) - A_i)s_i\sigma_i^T(F'(x_{i+1}) - A_i)}{\sigma_i^T(F'(x_{i+1}) - A_i)s_i} \tag{1}$$

is called 'two-sided rank-1 update' (TR1 update) of A_i.

This general definition is valid for any pair of primal and dual directions s_i and σ_i for which the denominator $\sigma_i^T(F'(x_{i+1}) - A_i)s_i$ is nonzero. Even though it is rarely exactly equal to zero ([CGT91]), the choice of s_i and σ_i should guard against numerical instability.

We wish to remark that the definition of the TR1 formula in this paper differs slightly from that in [GW02]. In [GW02] a combination of a secant and a tangent condition is considered applying for example AD to provide exact derivative information. While most other quasi-Newton update methods fulfill the secant condition

$$A_{i+1}s_i = F(x_{i+1}) - F(x_i) = F'(x_{i+1})s_i + O(\|s_i\|^2),$$

the approximation with the TR1 formula considered here satisfies a direct and adjoint *tangent* condition and thus agrees exactly with the new Jacobian $F'(x_{i+1})$ in certain directions.

Remark 3 *The TR1 update* A_{i+1} *as given in Definition 2 fulfills the 'direct tangent condition'*

$$A_{i+1}s_i = F'(x_{i+1})s_i$$

and the 'adjoint tangent condition'

$$\sigma_i^T A_{i+1} = \sigma_i^T F'(x_{i+1}).$$

Similar to the SR1 update method the TR1 update maintains the validity of previous tangent conditions, if the function F is affine, i.e., we have

$$A_i s_j = F'(x_i) s_j \quad \text{and} \quad \sigma_j^T A_i = \sigma_j^T F'(x_i)$$

for all $j < i$. This property is called *heredity* In contrast the good and bad Broyden formulas like most least change updates do not share this property. The heredity yields convergence of the TR1 update for affine problems F after at most n steps, provided none of the denominators happen to vanish exactly. This can be proved in analogy to the SR1 update method [NW99].

2.2 Local convergence for nonlinear problems

So far, we have not specified the relation between the directions s_i and σ_i. In analogy to Broyden's and the SR1 method we define the directions s_i for the direct tangent condition by the current step $s_i = x_{i+1} - x_i$. However, the choice of the direction σ_i used in the adjoint tangent condition is not that obvious. Two approaches for the choice of σ_i are discussed in [SWG05]. One choice in [SWG05] ensures that the directions of steepest descent of the sum of squares residual of the function and the linear model generated by the quasi-Newton method are equal. The other choice in [SWG05] ensures invariance with respect to linear transformations of the state space. Unfortunately so far we could not prove local convergence for these approaches.

Broyden, Dennis and Moré [BDM73] showed that a quasi-Newton method is locally linear convergent if it exhibits the property of *bounded deterioration*. Considering this property we define the following class of updates:

Definition 4 (Adjoint tangent rank-1 update) *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be differentiable, $x_{i+1} \in \mathbb{R}^n$. For a given matrix $A_i \in \mathbb{R}^{n\times n}$ and a given direction $\sigma_i \in \mathbb{R}^n$, the formula*

$$A_{i+1} = A_i + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(F'(x_{i+1}) - A_i) \tag{2}$$

is called 'adjoint tangent rank-1 (ATR1) update' of A_i.

In this general form the update has less in common with the the TR1 update in Definition 2. However, the methods, which are of particular interest, are indeed closely related to the TR1 update. Here we want to present three approaches for the choice of σ_i:

(A) $\sigma_i = (F'(x_{i+1}) - A_i)s_i$ (*transposed tangent Broyden update*),

(B) $\sigma_i = F(x_{i+1})$, and

(C) $\sigma_i = (F(x_{i+1}) - F(x_i))/\alpha_i - A_i s_i$ for a sequence $\{\alpha_i\} \subset (0, 1]$ with $\lim_{i\to\infty} \alpha_i = 1$ (*transposed secand Broyden update*).

Method (A) is the TR1 update and Method (B) is an approximation to (A) of order $o(\|s_i\|)$. Method (C) can be interpreted as a generalization of (B). Here the factor α_i may be given by a line search strategy. For a closer discussion of these approaches we refer to Section 2.4.

The formula (2) has the advantageous property that as long as $\sigma_i \neq 0$, the update is well defined since $\sigma_i^T \sigma_i > 0$. If $\sigma_i = 0$, the update can be skipped. Similar to Broyden's method we have for the adjoint tangent rank-1 update (2) and an arbitrary $B \in \mathbb{R}^{n\times n}$ with $\sigma_i^T B = \sigma_i^T F'(x_{i+1})$

$$A_{i+1} - A_i = \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(F'(x_{i+1}) - A_i) = \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(B - A_i).$$

Thus for any two matrix norms $\|\cdot\|_a$ and $\|\cdot\|_b$ with $\|A \cdot B\|_a \leq \|A\|_b \cdot \|B\|_a$ for $A, B \in \mathbb{R}^{n\times n}$ and $\|\frac{vv^T}{v^T v}\|_b = 1$ for $0 \neq v \in \mathbb{R}$ we have

$$\|A_{i+1} - A_i\|_a \leq \left\| \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i} \right\|_b \|(B - A_i)\|_a = \|(B - A_i)\|_a.$$

Choosing for instance for $\|\cdot\|_a$ the Frobenius norm and for $\|\cdot\|_b$ the l_2 norm gives

$$\|A_{i+1} - A_i\|_F \leq \|B - A_i\|_F. \tag{3}$$

Defining $Q := \{B \in \mathbb{R}^{n\times n} : \sigma_i^T B = \sigma_i^T F'(x_{i+1})\}$ yields that A_{i+1} is the solution of $\min_{B\in Q} \|B - A_i\|_F$. This solution is unique since Q is an affine subset of $\mathbb{R}^{n\times n}$ and $\|\cdot\|_F$ is strictly convex.

To analyze the local convergence properties we consider a general nonlinear function $F : \mathbb{R}^n \to \mathbb{R}^n$ that complies the following two general assumptions.

Assumption 5 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be differentiable and F' Lipschitz-continuous at $x_* \in \mathbb{R}^n$ with Lipschitz-constant $L < \infty$.*

Assumption 6 *Suppose $F(x_*) = 0$ and $F'(x_*)$ is non-singular.*

From now on, $\|\cdot\|$ will be used to denote an arbitrary norm, where as $\|\cdot\|_2$ denotes the l_2-norm and $\|\cdot\|_F$ the Frobenius norm. The bound on the approximation is described in the following lemma.

Lemma 7 (Bounded deterioration) *Suppose the Assumption 5 holds for the function F. Then the adjoint tangent rank-1 update (2) has the property*

$$\|A_{i+1} - F'(x_*)\|_2 \leq \|A_i - F'(x_*)\|_2 + L\|x_{i+1} - x_*\|_2.$$

Proof:

$$\begin{aligned} A_{i+1} - F'(x_*) &= A_i - F'(x_*) + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(F'(x_{i+1}) - A_i + F'(x_*) - F'(x_*)) \\ &= \left[I - \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}\right][A_i - F'(x_*)] + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}[F'(x_{i+1}) - F'(x_*)]. \end{aligned} \tag{4}$$

With the identities
$$\left\|I - \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}\right\|_2 = \left\|\frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}\right\|_2 = 1$$
we obtain the estimate
$$\|A_{i+1} - F'(x_*)\|_2 \leq \|A_i - F'(x_*)\|_2 + \|F'(x_{i+1}) - F'(x_*)\|_2.$$
Finally, since F' is assumed to be Lipschitz-continuous, we have
$$\|A_{i+1} - F'(x_*)\|_2 \leq \|A_i - F'(x_*)\|_2 + L\|x_{i+1} - x_*\|_2.$$
□

This gives immediately the following local convergence result.

Theorem 8 (Local linear convergence) *Suppose the Assumptions 5 and 6 hold for the function F. Then for the quasi-Newton Algorithm 1 with the adjoint tangent rank-1 update (2) and any $r \in (0,1)$ there exist $\epsilon(r) > 0$ and $\delta(r) > 0$ such that if $\|x_0 - x_*\| < \epsilon(r)$ and $\|A_0 - F'(x_*)\| < \delta(r)$ the sequence $\{x_i\}$ is well defined and converges to x_*. Furthermore, one has*
$$\|x_{i+1} - x_*\| \leq r\|x_i - x_*\| \quad \textit{for} \quad i = 0, 1, 2, \ldots \tag{5}$$
and $\|A_i\|$ and $\|A_i^{-1}\|$ are uniformly bounded.

Proof: The result follows by applying the general convergence result of [BDM73, Theorem 3.2.] in combination with Lemma 7. □

To prove that the method is even q-superlinear convergent we have to show that the *Dennis-Moré property*
$$\lim_{i\to\infty} \frac{\|(A_i - F'(x_*))s_i\|}{\|s_i\|} = 0 \tag{6}$$
holds, see e.g. [DM74]. To verify this property for the rank-1 update (2) we need the following two technical lemmas.

Lemma 9 *Suppose the Assumptions 5 and 6 hold for the function F. If the adjoint tangent rank-1 update (2) is applied in the quasi-Newton Algorithm 1 and if $\{x_i\}$ converges linearly as in (5), there is a $\tau > 0$ with $\|A_i - F'(x_*)\|_2 \leq \tau$ for all $i \in \mathbb{N}_0$.*

Proof: With $E_i := A_i - F'(x_*)$ and $e_i := x_i - x_*$ we have from Lemma 7
$$\|E_{i+1}\|_2 - \|E_i\|_2 \leq L\|e_{i+1}\|_2.$$
Taking the sum over i gives
$$\|E_{k+1}\|_2 - \|E_0\|_2 = \sum_{i=0}^{k} \|E_{i+1}\|_2 - \|E_i\|_2 \leq L \sum_{i=0}^{k} \|e_{i+1}\|_2.$$

Since $\|e_{i+1}\|_2 \leq r\|e_i\|_2$ with $r \in (0,1)$ and thus $\sum_{i=0}^{k} \|e_{i+1}\|_2 \leq \frac{r}{1-r}\|e_0\|_2$ we obtain

$$\|E_{k+1}\|_2 \leq \|E_0\|_2 + \frac{r\|e_0\|_2}{1-r} =: \tau \quad \text{for all} \quad k \in \mathbb{N}_0.$$

□

Lemma 10 *Let $\sigma \in \mathbb{R}^n$ be nonzero, and $E \in \mathbb{R}^{n\times n}$. Then*

$$\left\| E \left(I - \frac{\sigma\sigma^T}{\sigma^T\sigma} \right) \right\|_F \leq \|E\|_F - \frac{1}{2\|E\|_F} \left(\frac{\|E\sigma\|_2}{\|\sigma\|_2} \right)^2.$$

Proof: [DS96, Lemma 8.2.5]. □

Using the bounds in the last two lemmas, we can now prove the following result:

Lemma 11 (Transposed Dennis-Moré property) *Suppose the Assumptions 5 and 6 hold for the function F. If the adjoint tangent rank-1 update (2) is applied in the quasi-Newton Algorithm 1 and if $\{x_i\}$ converges linearly as in (5), then the update fulfills the transposed Dennis-Moré property*

$$\lim_{i\to\infty} \frac{\|(A_i - F'(x_*))^T\sigma_i\|_2}{\|\sigma_i\|_2} = 0.$$

Proof: Defining $E_i := A_i - F'(x_*)$ and $e_i := x_i - x_*$, we can conclude from (4) that

$$\begin{aligned}
\|E_{i+1}\|_F &\leq \left\| \left(I - \frac{\sigma_i\sigma_i^T}{\sigma_i^T\sigma_i} \right) E_i \right\|_F + \left\| \frac{\sigma_i\sigma_i^T}{\sigma_i^T\sigma_i} (F'(x_{i+1}) - F'(x_*)) \right\|_F \\
&\leq \left\| E_i^T \left(I - \frac{\sigma_i\sigma_i^T}{\sigma_i^T\sigma_i} \right) \right\|_F + \|F'(x_{i+1}) - F'(x_*)\|_2 \\
&\leq \|E_i^T\|_F - \frac{1}{2\|E_i^T\|_F} \left(\frac{\|E_i^T\sigma_i\|_2}{\|\sigma_i\|_2} \right)^2 + L\|e_{i+1}\|_2 \\
&= \|E_i\|_F - \frac{1}{2\|E_i\|_F} \left(\frac{\|E_i^T\sigma_i\|_2}{\|\sigma_i\|_2} \right)^2 + L\|e_{i+1}\|_2.
\end{aligned}$$

Because of the norm equivalence in $\mathbb{R}^n$ and Lemma 9 there is a $\tilde{\tau} > 0$ such that $\|E_i\|_F \leq \tilde{\tau}$. This yields

$$\frac{1}{\tilde{\tau}} \left(\frac{\|E_i^T\sigma_i\|_2}{\|\sigma_i\|_2} \right)^2 \leq \|E_i\|_F - \|E_{i+1}\|_F + L\|e_{i+1}\|_2$$

and thus

$$S_k := \frac{1}{\tilde{\tau}} \sum_{i=0}^{k} \left(\frac{\|E_i^T\sigma_i\|_2}{\|\sigma_i\|_2} \right)^2 \leq \|E_0\|_F - \|E_{k+1}\|_F + L \sum_{i=0}^{k} \|e_{i+1}\|_2 \leq \tilde{\tau} + \frac{Lr\|e_0\|_2}{1-r}.$$

Hence, we obtain $\lim_{k\to\infty} S_k < \infty$ and therefore

$$\lim_{i\to\infty} \frac{\|E_i^T \sigma_i\|_2}{\|\sigma_i\|_2} = 0.$$

$\square$

We would like to remark that so far we did not have to specify the direction σ_i, except $\sigma_i \neq 0$. Hence, the local convergence result in Theorem 8 and in particular the convergence of the approximation in the adjoint tangent direction in Lemma 11, i.e. the transposed Dennis-Moré property, do not depend on the choice of σ_i. As final step to prove superlinear convergence we have to show that the Dennis-Moré property (6) holds. For this purpose we assume, in addition to the result of the last lemma, that a specific property holds for the directions σ_i.

Assumption 12 (Residual property) *For the directions σ_i in the adjoint tangent rank-1 update (2), there exists a sequence $\{\lambda_i\} \subset \mathbb{R} \setminus \{0\}$ such that*

$$\lim_{i\to\infty} \frac{\|\lambda_i \sigma_i - (F'(x_*) - A_i)s_i\|_2}{\|s_i\|_2} = 0.$$

This is called 'residual property'.

As we will show in Section 2.3 this residual property has a strong connection to the direct tangent condition and the heredity property as discussed for the TR1 update in Section 2.1. In the comming proofs we use for technical reasons the following equivalent representation of the residual property: There exists a sequence $\{\lambda_i\} \subset \mathbb{R} \setminus \{0\}$ and a sequence $\{c_i\} \subset \mathbb{R}_+$ with $\lim_{i\to\infty} c_i = 0$ such that

$$\lambda_i \sigma_i = (F'(x_*) - A_i)s_i + r_i \quad \text{with} \quad \|r_i\|_2 \leq c_i \cdot \|s_i\|_2. \tag{7}$$

We find that as a consequence of the Cauchy-Schwarz inequality

$$\frac{\|Es\|}{\|s\|} \leq \frac{\|E^T E s\|}{\|Es\|}.$$

Now, we can prove the main result of this section:

Theorem 13 (Superlinear convergence) *Suppose the Assumptions 5 and 6 hold for the function F. Assume that the adjoint tangent rank-1 update (2) is applied in the quasi-Newton Algorithm 1 and σ_i fulfills the residual property. If $\{x_i\}$ converges linearly as in (5), then the Dennis-Moré property holds, i.e.*

$$\lim_{i\to\infty} \frac{\|(A_i - F'(x_*))s_i\|_2}{\|s_i\|_2} = 0.$$

Thus x_i converges q-superlinearly to x_*.

Proof: One has that

$$\begin{aligned}
\|\lambda_i\sigma_i\|_2^2 &= (s_i^T(F'(x_*) - A_i)^T + r_i^T)\lambda_i\sigma_i \\
&= s_i^T(F'(x_*) - A_i)^T\lambda_i\sigma_i + r_i^T\lambda_i\sigma_i \\
&\leq |\lambda_i|\|s_i\|_2\|(F'(x_*) - A_i)^T\sigma_i\|_2 + |\lambda_i|c_i\|s_i\|_2\|\sigma_i\|_2.
\end{aligned}$$

Division by $|\lambda_i|\|\sigma_i\|_2\|s_i\|_2$ yields

$$|\lambda_i|\frac{\|\sigma_i\|_2}{\|s_i\|_2} \leq \frac{\|(A_i - F'(x_*))^T\sigma_i\|_2}{\|\sigma_i\|_2} + c_i. \tag{8}$$

Furthermore we have

$$\frac{\|(A_i - F'(x_*))s_i\|_2}{\|s_i\|_2} \leq |\lambda_i|\frac{\|\sigma_i\|_2}{\|s_i\|_2} + c_i. \tag{9}$$

Inserting (8) into (9) finally yields

$$\frac{\|(A_i - F'(x_*))s_i\|_2}{\|s_i\|_2} \leq \frac{\|(A_i - F'(x_*))^T\sigma_i\|_2}{\|\sigma_i\|_2} + 2c_i.$$

With Lemma 11 and $\lim_{i\to\infty} c_i = 0$ this gives

$$\lim_{i\to\infty}\frac{\|(A_i - F'(x_*))s_i\|_2}{\|s_i\|_2} = 0.$$

□

2.3 Heredity

The above results are established using the least change property with respect to a fixed scale matrix norm of the ATR1 update. It is somewhat surprising that it also has the heredity property described in Lemma 15. We will show, that the residual property of the directions σ_i relates the adjoint tangent rank-1 update with the TR1 update. First of all we can show that the residual property in Assumption 12 is equivalent to an *approximate direct tangent condition*.

Lemma 14 (Approximate tangent condition) *Suppose the Assumption 5 holds for the function F. The adjoint tangent rank-1 update (2) is applied in the quasi-Newton Algorithm 1. Then σ_i satisfies the residual property, if and only if the resulting updated matrix A_{i+1} satisfies the approximate tangent condition*

$$\frac{\|(A_{i+1} - F'(x_*))s_i\|_2}{\|s_i\|_2} \leq L\|e_{i+1}\|_2 + c_i \tag{10}$$

with $\lim_{i\to\infty} c_i = 0$.

Proof: For any $\lambda_i \neq 0$ and any $r_i \in \mathbb{R}^n$ with $\lambda_i \sigma_i = (F'(x_*) - A_i)s_i + r_i$ we have from (2) that

$$\begin{aligned}
A_{i+1}s_i &= A_i s_i + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(F'(x_{i+1}) - A_i)s_i \\
&= A_i s_i + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(F'(x_*) - A_i)s_i + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(F'(x_{i+1}) - F'(x_*))s_i \\
&= A_i s_i + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(\lambda_i \sigma_i - r_i) + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(F'(x_{i+1}) - F'(x_*))s_i \\
&= A_i s_i + \lambda_i \sigma_i - \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i} r_i + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(F'(x_{i+1}) - F'(x_*))s_i \\
&= A_i s_i + (F'(x_*) - A_i)s_i + \left(I - \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}\right) r_i + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(F'(x_{i+1}) - F'(x_*))s_i \\
&= F'(x_*)s_i + \left(I - \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}\right) r_i + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(F'(x_{i+1}) - F'(x_*))s_i.
\end{aligned}$$

Subtracting $F'(x_*)s_i$, taking the norm, and dividing by $\|s_i\|_2$ gives

$$\frac{\|(A_{i+1} - F'(x_*))s_i\|_2}{\|s_i\|_2} \leq \frac{\|r_i\|_2}{\|s_i\|_2} + L\|e_{i+1}\|_2.$$

Thus, by denoting $c_i = \frac{\|r_i\|_2}{\|s_i\|_2}$, equation (10) with $\lim_{i\to\infty} c_i = 0$ holds, if and only if Assumption 12 holds. □

The following lemma shows, that the approximate direct tangent condition (10) holds in a generalized form for any previous step, too. Thus the adjoint tangent rank-1 update with the residual property yields *heredity* in this generalized fashion.

Lemma 15 (Heredity) *Suppose Assumption 5 holds for the function F. The adjoint tangent rank-1 update (2) is applied in the quasi-Newton Algorithm 1 and σ_i fulfills the residual property, then the estimate*

$$\frac{\|(A_i - F'(x_*))s_j\|_2}{\|s_j\|_2} \leq L \sum_{k=j+1}^{i} \|e_k\|_2 + c_j$$

is valid for all $j \in \mathbb{N}_0$ with $j < i$.

Proof: The assertion is proved by induction on i. For $i = 1$ and $j = 0$ we get the assertion from the approximate direct tangent condition. Now suppose the assertion holds for i. For $i + 1$ and and $j = i$ the assertion is proved again by the approximate direct tangent condition (10). If $j < i$ we obtain from (2)

$$A_{i+1} - F'(x_*) = \left(I - \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}\right)(A_i - F'(x_*)) + \frac{\sigma_i \sigma_i^T}{\sigma_i^T \sigma_i}(F'(x_{i+1}) - F'(x_*)).$$

Thus

$$\|(A_{i+1} - F'(x_*))s_j\|_2 \leq \|(A_i - F'(x_*))s_j\|_2 + \|(F'(x_{i+1}) - F'(x_*))s_j\|.$$

With the induction hypothesis this yields

$$\|(A_{i+1} - F'(x_*))s_j\|_2 \quad \leq \quad L \sum_{k=j+1}^{i} \|e_k\|_2\|s_j\|_2 + c_j\|s_j\|_2 + L\|e_{i+1}\|_2\|s_j\|_2$$

and thus finally

$$\frac{\|(A_{i+1} - F'(x_*))s_j\|_2}{\|s_j\|_2} \leq L \sum_{k=j+1}^{i+1} \|e_k\|_2 + c_j.$$

□

In this context we remark that the estimate only depends on the direction of the steps and is independent of their lengths, i.e.

$$\frac{\|E\lambda s\|}{\|\lambda s\|} = \frac{\|Es\|}{\|s\|} \quad \text{for} \quad 0 \neq s \in \mathbb{R}^n \quad \text{and} \quad 0 \neq \lambda \in \mathbb{R}.$$

The heredity in combination with the least change property (3) distinguishes the update formula (2) from other update formulas as, e.g., Broyden's or the SR1 method. Each of the latter update methods has only one of these two properties. Furthermore we can exploit the heredity to show even stronger convergence results. However, we have to assume that the current step s_i can be represented as a linear combination of previous steps in the following fashion.

Assumption 16 *Assume that the sequence $\{\tilde{s}_i\}$, defined by $\tilde{s}_i = s_i/\|s_i\|_2$ complies the following property: There exist $i_0 \in \mathbb{N}$, $c > 0$, and $k \in \mathbb{N}$, such that for all $i \geq i_0$*

$$\tilde{s}_i = \sum_{j=i-k}^{i-1} \lambda_j \tilde{s}_j \quad \textit{with} \quad |\lambda_j| \leq c. \tag{11}$$

Clearly the requirements on the steps in Assumption 16 are mathematically not desirable. They can be compared to the assumption of *uniformly linear independence* of the iteration steps. This property is in particular required to prove convergence of the SR1 method in [CGT91]. However we would like to point out, that Assumption 16 is a considerably weaker consequence of uniformly linear independence. Thus we do not require to be $k \geq n$. On the contrary linear dependent steps and a small number of k are favorable for the analysis as shown in the following lemma.

In the following proof, the number k is of particular relevance. It depends on the problem and describes, how far we have to go back in the iteration history to represent the current iterate as a linear combination of previous iterates which are *sufficiently* linear independent. Since $s_i \in \mathbb{R}^n$ it is reasonable to assume that $k \approx n$. This coincides with observations reported in [CGT91]. However as stated in [CGT91] there are cases for which fewer successive iteration steps may become linear dependent. While this contradicts the precondition of the convergence result in [CGT91], it is favorable for the analysis in Theorem 17 (iii), hence small values of k increase the lower bound of the r-order of convergence. For the definition of r-order of convergence we refer to [OR00]. An equivalent representation is given in [Pot89] as

$$r = \liminf_{i\to\infty} \left|\log \|x_i - x_*\|\right|^{1/i} = \liminf_{i\to\infty} \left|\frac{\log \|x_i - x_*\|}{\log \|x_0 - x_*\|}\right|^{1/i} \tag{12}$$

provided $x_i \neq x_*$. Hence r may be interpreted as the geometrical average of the sequence of values $\left|\frac{\log \|x_i - x_*\|}{\log \|x_{i-1} - x_*\|}\right|$.

Theorem 17 *Suppose the Assumptions 5 and 6 hold for the function F and $\{x_i\}$ converges linearly as in (5). The adjoint tangent rank-1 update (2) is applied in the quasi-Newton Algorithm 1 and σ_i satisfies the residual property (7) with $c_i \leq c_{max}\|e_i\|_2$ $(0 \leq c_{max} < \infty)$. Furthermore let the iteration steps comply Assumption 16. Then there exist constants $C_1, C_2 > 0$ independent of i and indices $i_1, i_2 \in \mathbb{N}$, such that*

(i) for all $i \geq i_1$

$$\frac{\|(A_i - F'(x_*))s_i\|_2}{\|s_i\|_2} \leq C_1 \|e_{i-k}\|_2, \tag{13}$$

(ii) for all $i \geq i_2$

$$\|e_{i+1}\|_2 \leq C_2 \, \|e_i\|_2 \, \|e_{i-k}\|_2, \tag{14}$$

and this yields

(iii) for the r-order of convergence $\mathcal{R} = \mathcal{R}(k)$ of the iteration

$$\mathcal{R}(k) \geq 1 + \eta_k \quad \textit{with} \quad \lim_{k\to\infty} \frac{\eta_k}{\frac{\log(k)}{k}} = 1. \tag{15}$$

Proof of Theorem 17 (i): Defining as before $E_i := A_i - F'(x_*)$, we have according to the proposition for $i \geq i_0$

$$E_i \tilde{s}_i = E_i \sum_{j=i-k}^{i-1} \lambda_j \tilde{s}_j = \sum_{j=i-k}^{i-1} \lambda_j E_i \tilde{s}_j.$$

This gives

$$\frac{\|E_i s_i\|_2}{\|s_i\|_2} = \|E_i \tilde{s}_i\|_2 \leq \sum_{j=i-k}^{i-1} |\lambda_j| \|E_i \tilde{s}_j\|_2 = \sum_{j=i-k}^{i-1} |\lambda_j| \frac{\|E_i s_j\|_2}{\|s_j\|_2}.$$

From Lemma 15 we get $\frac{\|E_i s_j\|_2}{\|s_j\|_2} \leq L \sum_{l=j+1}^{i} \|e_l\|_2 + c_j$ where $e_l := x_l - x_*$. Using $|\lambda_j| \leq c$ from Assumption 16 yields

$$\frac{\|E_i s_i\|_2}{\|s_i\|_2} \leq c \sum_{j=i-k}^{i-1} L \left(\sum_{l=j+1}^{i} \|e_l\|_2 \right) + c_j.$$

From equation (5) we have that $\|e_{i+1}\|_2 \leq r\|e_i\|_2$ with $r \in (0,1)$. Thus

$$\sum_{l=j+1}^{i} \|e_l\|_2 \leq \|e_{j+1}\|_2 \sum_{l=0}^{i-j-1} r^l = \|e_{j+1}\|_2 \frac{1-r^{i-j}}{1-r} \leq \frac{1}{1-r}\|e_{j+1}\|_2.$$

Accordingly we have from the assumptions that $c_j \leq c_{max}\|e_j\|_2$ with $0 \leq c_{max} < \infty$. This yields

$$\sum_{j=i-k}^{i-1} c_j \leq c_{max} \sum_{j=i-k}^{i-1} \|e_j\|_2 \leq \frac{c_{max}}{1-r}\|e_{i-k}\|_2.$$

This leads to

$$\begin{aligned}
\frac{\|E_i s_i\|_2}{\|s_i\|_2} &\leq c \sum_{j=i-k}^{i-1} \frac{L}{1-r}\|e_{j+1}\|_2 + c_j \;=\; \frac{cL}{1-r} \sum_{j=i-k}^{i-1} \|e_{j+1}\|_2 + c \sum_{j=i-k}^{i-1} c_j \\
&\leq \frac{cL}{(1-r)^2}\|e_{i-k+1}\|_2 + \frac{c\, c_{max}}{1-r}\|e_{i-k}\|_2 \\
&\leq \underbrace{\left(\frac{cLr}{(1-r)^2} + \frac{c\, c_{max}}{1-r} \right)}_{C_1} \|e_{i-k}\|_2.
\end{aligned}$$

□

To prove Theorem 17 (ii), we first state the following result [DS96, Lemma 4.1.15/16].

Lemma 18 *If the Assumption 5 holds for the function F, one has for any $u, v \in \mathbb{R}^n$, that*

$$\|F(v) - F(u) - F'(x_*)(v-u)\| \leq \frac{L}{2} \left(\|v - x_*\| + \|u - x_*\| \right) \|v - u\|.$$

If additionally the Assumption 6 holds for the function F, then there exist $\epsilon > 0$, $0 < \alpha < \beta$, such that for all $u, v \in \mathbb{R}^n$ with $\max\{\|v-x\|, \|u-x\|\} \leq \epsilon$

$$\alpha\|v-u\| \leq \|F(v) - F(u)\| \leq \beta\|v-u\|.$$

Furthermore, one can show the following result:

Lemma 19 *Suppose the Assumptions 5 and 6 hold, $\{x_i\}$ converges linearly as in (5), and the adjoint tangent rank-1 update (2) is applied in the quasi-Newton Algorithm 1. Then there exists an $i_0 \in \mathbb{N}$, $\alpha > 0$ such that for all $i \geq i_0$ and $e_i := x_i - x_*$*

$$\|e_{i+1}\| \leq \frac{\|(A_i - F'(x_*))s_i\|}{\|s_i\|} \frac{\|e_i\| + \|e_{i+1}\|}{\alpha} + \frac{L}{2\alpha} \left(\|e_i\| + \|e_{i+1}\|\right)^2. \qquad (16)$$

Proof:

$$\begin{aligned} 0 &= A_i s_i + F(x_i) \\ &= (A_i - F'(x_*))s_i + F(x_i) + F'(x_*)s_i \\ -F(x_{i+1}) &= (A_i - F'(x_*))s_i + [-F(x_{i+1}) + F(x_i) + F'(x_*)s_i] \end{aligned}$$

Using Lemma 18 and $\|s_i\| \leq \|e_i\| + \|e_{i+1}\|$ yields that there exists an $i_0 \in \mathbb{N}$ such that for all $i \geq i_0$

$$\begin{aligned} \|F(x_{i+1})\| &\leq \|(A_i - F'(x_*))s_i\| + ||-F(x_{i+1}) + F(x_i) + F'(x_*)s_i|| \\ &\leq \frac{\|(A_i - F'(x_*))s_i\|}{\|s_i\|}\|s_i\| + \frac{L}{2}(\|e_i\| + \|e_{i+1}\|)\|s_i\| \\ \alpha\|e_{i+1}\| &\leq \frac{\|(A_i - F'(x_*))s_i\|}{\|s_i\|}(\|e_i\| + \|e_{i+1}\|) + \frac{L}{2}\left(\|e_i\| + \|e_{i+1}\|\right)^2. \end{aligned}$$

□

Now Theorem 17 (ii) can be proved using the result of the previous lemma. This is followed by the proof of Theorem 17 (iii), which goes back to a more general convergence result in [OR00].

Proof of Theorem 17 (ii) and (iii): Applying the result of Theorem 17 (i) to equation (16) yields for sufficiently large indices i

$$\begin{aligned} \|e_{i+1}\|_2 &\leq C\|e_{i-k}\|_2 \frac{1+r}{\alpha}\|e_i\|_2 + \frac{(1+r)^2}{2\alpha}\|e_i\|_2^2 \\ &= \frac{1+r}{\alpha}\left(C\|e_{i-k}\|_2 + \frac{1+r}{2}\|e_i\|_2\right)\|e_i\|_2. \end{aligned}$$

Since $k \geq 1$ we have $\|e_i\|_2 \leq r^k\|e_{i-k}\|_2$ and thus

$$\|e_{i+1}\|_2 \leq \underbrace{\frac{1+r}{\alpha}\left(C + \frac{1+r}{2}r^k\right)}_{C_2}\|e_{i-k}\|_2 \; \|e_i\|_2.$$

This proves Theorem 17 (ii).

From [OR00, 9.2.9] we have that if a sequence $\{\varepsilon_i\} \subset (0, \infty)$ satisfies for sufficiently large indices i, that there are a constant $C < \infty$ and $k \in \mathbb{N}_0$ such that $\varepsilon_{i+1} \leq C\, \varepsilon_i\, \varepsilon_{i-k}$, then the r-order of convergence of $\{\varepsilon_i\}$ is $\mathcal{R} = \mathcal{R}(k) \geq \tau_k$, where τ_k is the unique positive root of $\tau^{k+1} - \tau^k - 1 = 0$. Moreover, $\tau_k \in (1, 2)$, $\tau_k > \tau_{k+1}$, and $\lim_{k\to\infty} \tau_k = 1$. Substituting $\eta_k := \tau_k - 1$ yields

$$(1+\eta_k)^k\eta_k = 1 \quad \text{with} \quad 1 > \eta_k > 0 \qquad (17)$$

and we prove, that

$$\lim_{k\to\infty} \frac{\eta_k}{\frac{\log(k)}{k}} = 1. \tag{18}$$

Since η_k decreases monotonically and $\eta_k \to 0$ for $k \to \infty$, we have that

$$\lim_{k\to\infty} \frac{\eta_k}{\frac{\log(k)}{k}} = \lim_{\eta_k\to 0} \frac{k\ \eta_k}{\log(k)}.$$

From equation (17), we get that $k = \frac{-\log(\eta_k)}{\log(1+\eta_k)}$. Thus using Bernoulli-l'Hospitales rule yields

$$\lim_{\eta_k\to 0} \frac{k\ \eta_k}{\log(k)} = \lim_{\eta_k\to 0} \frac{-\log(\eta_k)}{\log(k)} \cdot \underbrace{\lim_{\eta_k\to 0} \frac{\eta_k}{\log(1+\eta_k)}}_{=1} = \lim_{\eta_k\to 0} \frac{-\log(\eta_k)}{\log(k)}.$$

Furthermore $\log(k) = \log(-\log(\eta_k)) - \log(\log(1+\eta_k))$ and

$$\frac{\log(k)}{-\log(\eta_k)} = \frac{\log(\log(1+\eta_k)) - \log(-\log(\eta_k))}{\log(\eta_k)} = \frac{\log(\log(1+\eta_k))}{\log(\eta_k)} - \frac{\log(-\log(\eta_k))}{\log(\eta_k)}.$$

Applying again Bernoulli-l'Hospitales rule gives

$$\lim_{\eta_k\to 0} \frac{\log(\log(1+\eta_k))}{\log(\eta_k)} = \lim_{\eta_k\to 0} \frac{\eta_k}{\log(1+\eta_k)(1+\eta_k)} = 1$$

and

$$\lim_{\eta_k\to 0} \frac{\log(-\log(\eta_k))}{\log(\eta_k)} = \lim_{\eta_k\to 0} \frac{1}{\log(\eta_k)} = 0,$$

which finally yields $\lim_{\eta_k\to 0} \frac{-\log(\eta_k)}{\log(k)} = 1$ and proves (18). This concludes the proof. □

From Theorem 17 (ii), we get in particular $(k+1)$-step quadratic convergence of the iteration. However Theorem 17 (iii) yields a somewhat faster r-order of convergence then just $(k+1)$-step quadratic. The ratio of the corresponding efficiencies in the sense of Ostrowski [Ost66] is

$$\frac{\log(1+\eta_k)}{\log\left(1+\frac{1}{k+1}\right)} \approx \log(k).$$

If we assume that a full Newton step is $q \cdot n$ times as expensive as our quasi-Newton step, the latter efficiency is $q \cdot \log(k)$ times better than that of Newtons method. For some values of k, the quantity η_k is displayed it Table 1.

Though Broyden's update does not share heredity, we may note that it solves affine problems within $2n$ iteration steps. This finite termination results, first formulated by Burmeister (cf. [Sch79, B. 5.5.1.]) and published by Gay [Gay79], suggests, that on nonlinear but smooth problems the rate of convergence is at least $2n$-step quadratic, which corresponds to an r-order of $1 + \frac{1}{2n} + O(\frac{1}{n^2})$ by binomial expansion.

Table 1: Comparison of values for k and η_k

k	1	2	3	4	10	100	1000	10^5
η_k	0.618	0.466	0.380	0.325	0.184	0.0343	0.00526	9.28e-5
$\log(k)/k$	0	0.347	0.366	0.347	0.230	0.0461	0.00690	1.15e-4
$\frac{\eta_k}{\log(k)/k}$	–	1.34	1.04	0.94	0.80	0.75	0.76	0.81

Comparing the computational effort of the ATR1 updates to Broyden's update, we may state, that in terms of function evaluations an ATR1 update is at least twice as expensive as Broyden's update since it requires the evaluation of the adjoint vector $\sigma_i^T F'(x_{i+1})$. However concerning the linear algebra, the update of an existing factorization dominates. Hence the computational effort of an ATR1 update is approximately equal to Broyden's update.

2.4 Variants of ATR1 updates

As stated in the previous section, the choice of σ_i in formula (2) is crucial for the properties of the ATR1 update method. Especially the approximate direct tangent condition is important. Thus σ_i should be a sufficiently good approximation to $(F'(x_*) - A_i)s_i = F(x_i) + F'(x_*)s_i$. However to apply the method, it is essential that σ_i can be computed with a reasonable computational effort. In Section 2.2 we proposed three approaches for σ_i. All of them comply the residual property in Assumption 12 as shown in the following lemma.

Lemma 20 *Suppose Assumption 5 holds for the function F and the adjoint tangent rank-1 update (2) is applied in the quasi-Newton Algorithm 1. If σ_i is defined by*

(A), then

$$\|\sigma_i - (F'(x_*) - A_i)s_i\|_2 \leq L\|e_{i+1}\|_2\|s_i\|_2,$$

(B), then

$$\|\sigma_i - (F'(x_*) - A_i)s_i\|_2 \leq \frac{L}{2}(\|e_i\|_2 + \|e_{i+1}\|_2)\|s_i\|_2,$$

(C), then

$$\|\alpha_i\sigma_i - (F'(x_*) - A_i)s_i\|_2 \leq \frac{L}{2}(\|e_i\|_2 + \|e_{i+1}\|_2)\|s_i\|_2 + (1-\alpha_i)\|A_i\|_2\|s_i\|_2.$$

Proof: The inequality for (A) follows by the Lipschitz continuity of F' in x_*. For (B) we get the inequality by Lemma 18. For (C) the inequality follows by (B) and the identity $A_i s_i = -F(x_i)$. □

Thus adjoint tangent rank-1 updates using (A), (B), and (C) converge locally q-superlinear. Additionally we can apply Theorem 17 to Approach (A) and (B).

Furthermore for (A) we have that it complies Definition 2 of the TR1 update. Thus, beside the adjoint tangent condition, it fulfills the direct tangent condition $A_{i+1}s_i = F'(x_{i+1})s_i$, too. With this we can prove, that in Lemma 15 even

$$\frac{\|(A_i - F'(x_*))s_j\|_2}{\|s_j\|_2} \leq L \sum_{k=j+1}^{i} \|e_k\|_2 \qquad (j < i)$$

holds. Moreover it is possible to prove even k-step quadratic convergence for (A). However it requires the additional computation of the directional derivative $F'(x_{i+1})s_i$. That can be performed by the forward mode of Automatic Differentiation but causes additional computational effort. Approach (B) neither fulfills a direct tangent condition nor a secant condition as for example Broyden's method. Therefore it avoids the computation of the directional derivative. The same is valid for (C).

Additionally, an important feature of the methods (A), (B), and (C) is, that they are invariant with respect to regular linear transformations of the state space of x. This means, for any regular $T \in \mathbb{R}^{n\times n}$ and

- $\tilde{F} : \mathbb{R}^n \to \mathbb{R}^n$ with $\tilde{F}(\tilde{x}) = F(T^{-1}\tilde{x})$ and
- $\tilde{x}_i = Tx_i$, $\tilde{A}_i = A_iT^{-1}$

applied in Algorithm 1 with the update formula (2), we get that

$$\tilde{A}_{i+1} = A_{i+1}T^{-1} \quad \text{and} \quad \tilde{x}_{i+1} = Tx_{i+1}.$$

This property is shared also with Broyden's bad update formula. In contrast Broyden's (good) update is only invariant with respect to regular linear transformations of the range of F. However an often applied norm-dependent line search destroys this invariance of Broyden's method. The invariance in the domain of F is not influenced by such a line search.

Furthermore the adjoint tangent condition of approach (B) yields, as long as A_i is non-singular, that

$$-F(x_i)^T F'(x_i)A_i^{-1}F(x_i) = -F(x_i)^T A_iA_i^{-1}F(x_i) = -\|F(x_i)\|_2^2.$$

Thus the quasi-Newton direction $s_i = -A_i^{-1}F(x_i)$ is a decent direction in the l_2 norm and the decent is the same as for the Newton step $\bar{s} = -F'(x_i)^{-1}F(x_i)$. This fact can be exploited for a globalization strategy based on a line search approach. A forthcoming paper will focus on this global convergence analysis.

3 Implementation and numerical results

The ATR1 update methods are applied to nonlinear equation problems of the Moré test set [MGH81] and two particular test functions. The methods are compared to Newton's and other quasi-Newton methods as freezing the Jacobian and using Broyden's (good) update method. Since the focus of this paper is the local convergence behavior we do not involve a line search or trust region method. Thus the numerical tests of the ATR1 update are performed only for method (A) and (B).

3.1 Algorithmic implementation

We use the following standard algorithm to solve the nonlinear equation problem $F(x) = 0$:

1. Initialize $x := x_0$, $A := F'(x_0)$
2. Compute a factorization $PLU = A$
3. While $\max\{\|F(x)\|, \|s\|\} > \varepsilon$:
   3.1. Solve $PLUs = -F(x)$
   3.2 Compute $x := x + s$
   3.3. Update PLU
4. Set $x_* := x$

Here the update procedure used in step 3.3. is determined by the specific choice of the rank-1 formula. For the comparison to other methods, this step is modified accordingly. To compute the PLU factorization the LAPACK routine `DGETRF` is used. The initialization of the iterations and Newton's method requires the Jacobian of F. Additionally the ATR1 update formulas are based on the terms $F'(x)s$ and $\sigma^T F'(x)$. To evaluate these derivative information we use the AD tool ADOL-C [GJU96] for the differentiation of C/C++ codes.

The overall computing effort depends heavily on the factorization used. In this study a LU factorization with partial pivoting is used. This factorization is updated according to an algorithm by Bennett [Ben65] without readjustment of the pivoting. This algorithm is implemented under consideration of efficient memory access as described in [SGB05]. The update algorithm by Kielbasinsky/Schwetlick [KS88] with readjusting the pivoting was tested, too. However for the test problems considered here, the algorithm had no relevant influence on the iteration counts but the run time increases by a factor of about 2. As alternative, one may use the update of a QR factorization, which has very good stability properties but results in considerably higher computational costs for the solution of the linear systems as well as for the update.

3.2 Numerical tests

To illustrate the quality of the approximation we applied the ATR1 update methods to various test problems and compared the number of iterations needed to reach convergence with a reasonable tolerance. Additionally we state the run time required for the whole iteration process. For that purpose we compiled the program using *gcc 3.3.2.* and executed it on a PC with AMD Athlon(tm) XP 1.6 GHz (256 KB cache) processor. The results for the higher dimensional nonlinear equation problems of the Moré test set are displayed in Table 2. The numbers in the first column refer to the number of the test problem in [MGH81]. If not otherwise stated, these tests are performed for the dimension $n = 1000$ using the initial iterates as proposed in the test set. The iteration is performed up to a tolerance of $\varepsilon = 10^{-14}$ in the residual $\|F(x_i)\|_\infty$ and in the step size $\|s_i\|_\infty$.

Table 2: Iteration counts (a) and run times in seconds (b) of Moré test set

Test problem		Newton	Method A	Method B	Broyden	frozen J.
(21)	(a)	2	3	3	5	4
	(b)	1.54	0.682	0.677	0.748	0.624
(22)	(a)	47	47	47	67	-
	(b)	25.3	3.39	3.16	3.59	-
(26)[1]	(a)	7	18	19	22	-
	(b)	4.34	1.56	1.63	1.57	-
(27)[2]	(a)	349	349	350	-	-
	(b)	3.85e-2	3.34e-2	3.33e-2	-	-
(28)	(a)	3	5	5	5	8
	(b)	2.13	0.830	0.810	0.764	0.741
(29)	(a)	3	5	5	5	8
	(b)	122	34.8	34.6	31.8	32.5
(30)	(a)	5	14	14	17	34
	(b)	3.17	1.29	1.31	1.30	1.43
(31)	(a)	6	21	20	31	104
	(b)	4.10	1.77	1.73	2.02	3.28

To compare the behavior of the methods for different dimensions n, we use additionally the test problem

$$F : \mathbb{R}^n \to \mathbb{R}^n \qquad F(x) = (f_i(x))_{i=1,\ldots,n}\,,$$
$$f_i(x) = \xi_i + \sum_{j=1, j\neq i}^{n} \xi_j^2 \quad \text{and} \quad \xi_i = \frac{x_i - (i-1)}{i}. \tag{19}$$

It has a solution $F(x_*) = 0$ for $x_* = (0, 1, 2, \ldots, n-1)^T$ and $F'(x_*) = I$. As initial iterate we take $x_0 = 0$. The iteration is computed up to a tolerance of $\varepsilon = 10^{-12}$ in the residual $\|F(x_i)\|_\infty$ and in the step size $\|s_i\|_\infty$. The results are given in Table 3.

For a numerical estimation of the r-order of convergence, we monitore the descent of the error in the last five iterations and define

$$R_5 = \left(\frac{\log \|e_{i_{end}}\|_2}{\log \|e_{i_{end}-5}\|_2}\right)^{\frac{1}{5}}.$$

as an estimation of the r-order of convergence. Numerical results of R_5 for different dimensions n are displayed in Table 4.

A particular field of application for the solution of nonlinear equations are implicit methods for ODEs and DAEs. Therefore we consider also the solution of the first implicit Euler step with varying integration step sizes h for

[1] Initial iterate is chosen with $x_0 = \frac{1}{2}\hat{x}_0$ with $\hat{x}_0$, proposed in the test set. Otherwise no convergence was achieved for dimension $n = 1000$.

[2] The dimension is chosen with $n = 20$ to avoid floating point overflow.

Table 3: Iteration counts (a) and run times in seconds (b) of test function (19)

Dimension		Newton	Method A	Method B	Broyden
10	(a)	8	17	17	26
	(b)	3.87e-4	8.21e-4	7.78e-4	4.61e-4
100	(a)	12	20	22	36
	(b)	4.24e-2	9.27e-3	9.82e-3	9.46e-3
500	(a)	14	23	23	43
	(b)	1.48	0.43	0.40	0.53
1000	(a)	15	24	24	51
	(b)	8.66	1.87	1.91	2.88
2000	(a)	16	24	25	59
	(b)	57.70	9.06	9.32	15.48

Table 4: Estimated r-order of convergence (R_5) for test function (19)

n	2	4	10	100	1000
Method (A)	1.802	1.418	1.143	1.114	1.121
Method (B)	1.826	1.385	1.134	1.114	1.136
Broyden	1.395	1.159	1.071	1.027	1.049

the Robertson initial value problem (IVP). The problem describes three chemical reactions with three components. For further details to this IVP and its integration we refer to [Rob66] and [HW91]. Hence, the dimension of the problem is only three, the run times for the solution are negligible. However the iteration counts, displayed in Table 5, nicely illustrate the performance of the approximation.

Table 5: Iteration counts for first implicit Euler step of Robertson problem

Step size h	Newton	Method A	Method B	Broyden	frozen J.
10^{-4}	3	3	3	3	4
10^{-3}	5	5	5	7	-
0.01	8	8	9	15	-
0.1	12	13	13	47	-
1	15	27	19	-	-
10	19	21	92[3]	-	-

Inspecting the results for the Moré test set given in Table 2, one can see, that as expected Newton's method requires least iterations in all tests. The ATR1 methods mostly need some more iterations than Newton's method. Thereby

[3]The large iteration count may be due to the distance between initial iterate and solution.

Method (A) and (B) require roughly the same numbers of iteration to converge. However, the ATR1 methods need significantly less iteration than Broyden's method. Convergence of the frozen Jacobian method is rather poor. The run time for Newton's method is significantly higher than for the quasi-Newton methods. Since the computation of the ATR1 update is more expensive than Broyden's update the difference in the run time is not that much for these test problems.

The results in Table 3 for the test problem (19) verify the results of the iteration counts of the Moré test set. Thereby the iteration counts increase with the dimension n. Here Newton's method and the ATR1 methods perform much better than Broyden's method. This may be caused by the fact, that the scaling of the components of x gets worse if the dimension is increased. The run time of Newton's method increases significantly faster than that of the quasi-Newton methods if the value of n grows. Due to the significantly fewer iterations of the ATR1 methods compared to Broyden's method, the run times of the ATR1 methods are less than the run times of Broyden's method if the dimension is not too small. For this test problem the frozen Jacobian method did not converge.

The estimated r-orders of convergence displayed in Table 4 confirme for $n = 2$ to $n = 100$ that the rate of convergence of the quasi-Newton methods decreases if the number of dimensions increases. The fact, that R_5 for $n = 1000$ is larger than for $n = 100$ may be due to the choice of F. Furthermore we see, that, except for $n = 10$, the values of R_5 of the ATR1 update methods are larger than $1 + \eta_n$, which is predicted as lower bound in Theorem 17, provided that $k = n$. Moreover for all dimensions n, the rates of convergence of the ATR1 updates are significantly larger than the ones of Broyden's update. The estimated rates for Broyden's update for $n = 2$ to $n = 10$ are slightly larger than $1 + \frac{1}{2n}$. This confimes the predicted r-order of Broyden's update in [Gay79].

For the integration of the Robertson problem we can state that the iteration counts of the ATR1 methods are mostly in the scope of that of Newton's method. In contrast to this, convergence of Broyden's method is rather poor and for larger integration step sizes it even does not converge. The frozen Jacobian method is not appropriate for this problem.

4 Conclusion

We analyze a new class of quasi-Newton methods for the solution of nonlinear equations. Here the TR1 update is related with the class of adjoint tangent rank-1 updates. We give sufficient conditions for local linear and q-superlinear convergence of these methods. Additionally we show a heredity property. With this and reasonable assumptions on the iteration steps we can even prove $(k+1)$-step quadratic convergence and estimate the r-order of convergence.

Three specific variants of update formulas are proposed. All of them are in particular invariant with respect to the scaling of the domain of the nonlinear function. Numerical results verify the convergence properties of the ATR1 methods. Thereby the ATR1 methods converge significantly faster than Broyden's method. However the computation of the ATR1 update is slightly more expensive than that of Broyden's update. Nevertheless the run times of the ATR1 methods are mostly less than the run times of Broyden's method, especially for problems which are of higher dimension and badly scaled in the domain.

The considered class of adjoint tangent rank-1 updates combines for the first time the least change property with heredity. This yields favorable properties for local convergence. Future work will focus on combining the proposed update methods with line search and trust region algorithms to ensure global convergence.

References

[BCH+05] H. M. Bücker, G. F. Corliss, P. Hovland, U. Naumann, and B. Norris, editors. *Automatic Differentiation: Applications, Theory, and Implementations*, volume 50 of *Lecture Notes in Computational Science and Engineering*. Springer, New York, NY, 2005.

[BDM73] C. G. Broyden, J. E. Jr. Dennis, and J. J. Moré. On the local and superlinear convergence of quasi-Newton methods. *J.I.M.A.*, 12:223–246, 1973.

[Ben65] J.M. Bennett. Triangular factors of modified matrices. *Numerische Mathematik*, 7:217–221, 1965.

[Bro65] C. G. Broyden. A class of methods for solving nonlinear simultaneous equations. *Math. Comp.*, 19:577–593, 1965.

[CGT91] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. Convergence of quasi-newton matrices generated by the symmetric rank one update. *Math. Programming*, 50:177–195, 1991.

[DM74] J. E. Jr. Dennis and J. J. Moré. A characterization of superlinear convergence and its application to quasi-Newton methods. *Math. Comp.*, 28:549–560, 1974.

[DS96] J. E. Jr. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, 1996.

[Gay79] D. M. Gay. Some convergence properties of broyden's method. *SIAM J. Numer. Anal.*, 16:623–630, 1979.

[GJU96] A. Griewank, D. Juedes, and J. Utke. ADOL-C: A package for automatic differentiation of algorithms written in C/C++. *TOMS*, 22:131–167, 1996.

[Gri00] A. Griewank. *Evaluating derivatives: principles and techniques of algorithmic differentiation*. SIAM, 2000.

[GS78] D. M. Gay and R. B. Schnabel. Solving systems of nonlinear equations by Broyden's method with projected updates. *In Nonlinear Programming 3, O. Mangasarian, R. Meyer and S. Robinson, eds., Academic Press, NY*, pages 245–281, 1978.

[GW02] A. Griewank and A. Walther. On constrained optimization by adjoint based quasi-Newton methods. *Opt. Meth. and Soft.*, 17:869–889, 2002.

[Hab04] E. Haber. Quasi-newton methods for large scale electromagnetic inverse problems. *Inverse Problems*, 21:305–317, 2004.

[HW91] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II*. Springer-Verlag, 1991.

[KS88] A. Kielbasinski and H. Schwetlick. *Numerische lineare Algebra*. VEB Deutscher Verlag der Wissenschaften, 1988.

[MGH81] J. J. Moré, B. S. Garbow, and K. E. Hillstrom. Testing unconstrained optimization software. *TOMS*, 7:17–41, 1981.

[NW99] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer-Verlag, 1999.

[OR00] J. M. Ortega and W. C. Reinboldt. Iterative solution of nonlinear equations in several variables. *Academic Press*, 2000.

[Ost66] A. Ostrowski. *Solution of Equations and Systems of Equations*. Academic Press, New York, 1966.

[Pot89] F. A. Potra. On q-order and r-order of convergence. *Journal of Optimization Theory and Applications*, 63(3):415–431, 1989.

[Rob66] H. H. Robertson. The solution of a set of reaction rate equations. *In J. Walsh, ed.: Numerical Analysis, an Introduction, Academic Press*, pages 178–182, 1966.

[Sch79] H. Schwetlick. *Numerische Lösung nichtlinearer Gleichungen*. VEB Deutscher Verlag der Wissenschaften, 1979.

[SGB05] P. Stange, A. Griewank, and M. Bollhöfer. On the efficient update of rectangular LU factorizations subject to low rank modifications. TU Berlin, Preprint 2005/27, 2005.

[SWG05] S. Schlenkrich, A. Walther, and A. Griewank. AD-based quasi-Newton methods for the integration of stiff ODEs. In Bücker et al. [BCH+05], pages 89–98.