Noname manuscript No.
(will be inserted by the editor)

On the convergence of the Broyden–like method and the Broyden–like matrices

Florian Mannel

Received: date / Accepted: date

Abstract We study the convergence properties of the matrices generated by the Broyden–like method for the solution of nonlinear systems of equations, with particular emphasis on Broyden's original method. We develop various sufficient conditions for the convergence of these matrices and use high-precision numerical experiments to demonstrate on several examples that these conditions are satisfied. We also show how the developed sufficient conditions are related to the rate of convergence of the iterates of the method. In particular, this work contains the following findings: In all numerical experiments the Broyden–like updates converge at least r-linearly, and if the Jacobian at the root is regular then the iterates appear to converge with an asymptotical q-order larger than one. Furthermore, the cluster points of the normalized steps span a one-dimensional linear space in all numerical experiments. This implies that the steps consistently violate uniform linear independence and indicates that the available convergence results for the Broyden–like matrices require assumptions that are unlikely to be satisfied. For the special case of the Broyden updates the numerical results suggest $2n$-step q-quadratical convergence under the standard assumptions for q-superlinear convergence of the iterates.

Keywords Broyden–like method · Broyden's method · quasi-Newton methods · convergence of Broyden–like matrices · rate of convergence of Broyden–like method · systems of nonlinear equations

Mathematics Subject Classification (2010) 49M15 · 65H10 · 90C53

1 Introduction

This work investigates the convergence of the matrices that the Broyden–like method generates, see Algorithm BL below. As part of this investigation we

F. Mannel
University of Graz
E-mail: florian.mannel@uni-graz.at

study the question whether the iterates of the Broyden–like method converge faster than q-superlinearly. Several researchers have pointed out that it is unknown if the Broyden–like matrices converge, for instance under the standard assumptions for q-superlinear convergence of the iterates, cf., e.g., the survey articles [11, Example 5.3], [26, p. 117], [16, p. 306] and [2, p. 940]. The main contributions of this work are

- to propose conditions that imply convergence of the Broyden–like matrices and hold in numerical experiments,
- to prove that these conditions are satisfied in some special cases,
- to relate these conditions to the rate of convergence of the iterates and to study the rate of convergence of the iterates in numerical experiments,
- to present numerical evidence that the Broyden–like matrices converge under the standard assumptions for q-superlinear convergence of the iterates.

Further contributions include the observations that the cluster points of the normalized Broyden–like steps consistently span a one-dimensional space and that the limit of the Broyden–like matrices never equals the true Jacobian. Each of these findings implies that the normalized Broyden–like steps consistently violate uniform linear independence. This suggests that the only previously available convergence result for the Broyden–like matrices requires assumptions that are frequently (possibly always) violated and thus underlines the significance of the new convergence conditions that we develop here.

The Broyden–like method, cf., e.g., [27], [32, Section 6] and [19, Algorithm 1], is a well-known tool for finding a solution of a smooth system of equations $F(u) = 0$, where F maps from $\mathbb{R}^n$ into $\mathbb{R}^n$. It reads as follows.

Algorithm BL: Broyden–like method

Input: $(u^0, B_0) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}$, B_0 invertible, $0 < \sigma_{\min} \le \sigma_{\max} < 2$
for $k = 0, 1, 2, \ldots$ **do**
  if $F(u^k) = 0$ **then** let $u^* := u^k$; STOP
  Solve $B_k s^k = -F(u^k)$ for s^k
  Let $u^{k+1} := u^k + s^k$ and $y^k := F(u^{k+1}) - F(u^k)$
  Choose $\sigma_k \in [\sigma_{\min}, \sigma_{\max}]$
  Let $B_{k+1} := B_k + \sigma_k (y^k - B_k s^k) \frac{(s^k)^T}{\|s^k\|^2}$
end
Output: u^*

Choosing $(\sigma_k) \equiv 1$ for the *updating sequence* (σ_k) recovers Broyden's method [5]. An appropriate choice of σ_k ensures that B_{k+1} is invertible if B_k is invertible. In fact, the Sherman–Morrison formula shows that all except at most one value of σ_k maintain invertibility of B_k. We emphasize that both Broyden's method and the more general Algorithm BL continue to attract the interest of researchers, cf. for instance the recent extensions to set-valued maps in, e.g., [3, 1] or the incorporation into infinite-dimensional semismooth quasi-Newton methods for PDE-constrained optimization in [28, 25].

This work is devoted to the convergence of (B_k). Following the approach of other studies on the convergence of quasi-Newton matrices such as [6] and [12],we develop conditions that ensure the convergence of (B_k) and verify in numerical experiments of high-precision that these conditions are satisfied. We also discuss the implications of these conditions for the rate of convergence of (u^k) and study this rate in the numerical experiments.

Let us briefly address the connection between convergence of (B_k) and (u^k). We restrict the discussion to the case that (u^k) converges to a root $\bar{u}$ of F at which $F'(\bar{u})$ is regular; the singular case is covered in [23]. If $F'(\bar{u})$ is invertible and (u^k) converges q-superlinearly to $\bar{u}$, then the Broyden–like updates $(B_{k+1} - B_k)$ and the iterates (u^k) satisfy

$$c \, \|B_{k+1} - B_k\| \leq \frac{\|u^{k+1} - \bar{u}\|}{\|u^k - \bar{u}\|} \leq C \, \|B_{k+1} - B_k\| \tag{1}$$

for all $k \geq 0$ with constants $c, C > 0$, cf. Lemma 3. Thus, by quantifying how fast $(\|B_{k+1} - B_k\|)$ converges to zero, we obtain additional information about the q-superlinear convergence of (u^k). For instance, if $(\|B_{k+1} - B_k\|)$ converges r-linearly to zero (which holds in all numerical experiments), then there is $c \in [0, 1)$ such that $\|u^{k+1} - \bar{u}\| \leq c^k \|u^k - \bar{u}\|$ for all k sufficiently large. This is slower than a q-order of convergence, i.e. $\|u^{k+1} - \bar{u}\| \leq C\|u^k - \bar{u}\|^\delta$ for some $C > 0$ and $\delta > 1$ and all k sufficiently large, but faster than

$$\sum_{k=0}^{\infty} \left(\frac{\|u^{k+1} - \bar{u}\|}{\|u^k - \bar{u}\|} \right)^2 < \infty,$$

a rate of convergence that has recently been proven for Algorithm BL and is evidently faster than q-superlinear convergence; cf. [24]. The numerical experiments for regular $F'(\bar{u})$ indicate quite convincingly that (u^k) always converges with a q-order $\delta > 1$; the value of δ is somewhat close to but smaller than $\frac{2n+1}{2n}$. This, in turn, yields that $(\|B_{k+1} - B_k\|)$ exhibits an r-order of convergence no smaller than δ and, in particular, converges r-superlinearly to zero. The numerical experiments with $n \leq 4$ indeed display this r-superlinear rate, but the experiments with larger n are inconclusive in this respect. However, since $\frac{2n+1}{2n}$ is rather close to 1 in the inconclusive experiments, we conjecture that $(\|B_{k+1} - B_k\|)$ does indeed tend to zero with r-order δ, but that experiments of even higher precision would be needed to confirm or reject this. Finally, since the concrete value of δ appears quite reliable in the experiments and since the experiments show that $(\|B_{k+1} - B_k\|)$ is not monotone, we infer from (1) that the r-order of convergence of $(\|B_{k+1} - B_k\|)$ is no larger than δ. In a certain sense this would be a rather complete answer to the question of how fast (u^k) and (B_k) converge in Algorithm BL. In passing let us emphasize again that the statements concerning convergence rates are based on numerical observations, not on theoretical results. Yet, such observations have not been available before and contribute to a better understanding of the convergence behavior of Algorithm BL. Furthermore, they may hopefully prove to be a valuable starting point for theoretical analysis in this direction.

Before discussing the novel conditions for convergence of (B_k) that we propose in this work, we now turn our attention to the conditions that are already available. To this end, we assume that (u^k) converges to a root $\bar{u}$ of F and distinguish the two cases that $F'(\bar{u})$ is either regular or singular. In the first case there is only one general result available that ensures convergence of (B_k): It is established in [27, Theorem 5.7] and in [20] that if the sequence of steps (s^k) is *uniformly linearly independent*, cf. [6, (AS.4)] for a definition, then (B_k) converges and $\lim_{k\to\infty} B_k = F'(\bar{u})$. However, conditions which imply uniform linear independence of (s^k) are unknown and we are not aware of a single example—be it theoretical or numerical—in which $\lim_{k\to\infty} B_k = F'(\bar{u})$ holds for $n > 1$ including the numerical examples contained in this work. Instead, available examples for Broyden's method such as [9, Example 5.3] and [10, Lemma 8.2.7] show that (u^k, B_k) can converge to $(\bar{u}, B)$ with $B \neq F'(\bar{u})$ in situations where the standard assumptions for q-superlinear convergence are satisfied. In addition, we have shown in [22, Corollary 1] that if one or more component functions of F are affine and B_0 agrees with $F'(u^0)$ on at least one of the corresponding rows, then (s^k) violates uniform linear independence. We conclude that while the available general convergence result may be helpful if the matrix update uses other directions than s^k, its applicability for the matrices generated by Algorithm BL seems quite limited if $F'(\bar{u})$ is regular. This statement also holds if $F'(\bar{u})$ is singular: The recent results in [23] show for Broyden's method that (B_k) converges, but that (s^k) violates uniform linear independence [23, Corollary 2]. Summarizing, there are no results available for general nonlinear F with regular $F'(\bar{u})$ that show convergence of (B_k) and whose assumptions are satisfied in numerical examples. The present work aims at closing this gap. We develop various sufficient conditions for the convergence of the Broyden–like matrices and we verify in numerical experiments with 1000 digits accuracy that several of these conditions are consistently satisfied. On a side note, all conditions allow for $\lim_{k\to\infty} B_k \neq F'(\bar{u})$.

Next we outline the conditions for the convergence of the Broyden–like matrices that are developed in this work. They are grouped into two sets. The first set evolves around the cluster points of the normalized steps

$$\hat{s}^k := \frac{s^k}{\|s^k\|}, \quad k \geq 0.$$

Our main result states that (B_k) converges if *all* cluster points of $(\hat{s}^k)$ are contained in a set of the form $\{\pm\bar{s}\}$ for some unit vector $\bar{s}$ and

$$\sum_k \min\left\{\left\|\hat{s}^k - \bar{s}\right\|, \left\|\hat{s}^k + \bar{s}\right\|\right\} < \infty \tag{2}$$

is satisfied (the result still holds under a somewhat weaker summability property); cf. section **??**. These assumptions may seem very restrictive, but they are, in fact, consistently satisfied in the numerical experiments. For the case that $F'(\bar{u})$ is singular with some additional structure, we will actually prove that (2) holds, complementing results from [23]. This case also serves as a motivation to derive the convergence conditions in this work without requiring invertibility of $F'(\bar{u})$ or superlinear convergence of (u^k) whenever possible (in the singular setting the convergence of (u^k) is only q-linear). Furthermore, we point out that if $(\hat{s}^k)$ satisfies (2), then it cannot be uniformly linearly independent according to [23, Corollary 1]. For the case that $F'(\bar{u})$ is regular, we were not able to prove that (2) is satisfied, except in the special case that F has only one nonlinear component function and the rows of B_0 that correspond to affine components of F match the corresponding rows of $F'(u^0)$: [22, Corollary 1] implies that (2) holds since for $k \geq 1$ all summands vanish.

The second set of sufficient conditions does not involve the cluster points of $(\hat{s}^k)$. Instead, we focus on the norm of the updates

$$\varepsilon_k := \|B_{k+1} - B_k\|, \quad k \geq 0.$$

The second set is divided into three blocks of conditions. In the first block we show, for instance, that the following condition ensures convergence of the Broyden–like matrices: (u^k) converges to some $\bar{u}$ and there are $M \in \mathbb{N} \cup \{0\}$ and $\gamma, C > 0$ such that

$$\left\|u^{k+1} - \bar{u}\right\| \leq C \left\|u^k - \bar{u}\right\| \left\|u^{k-M} - \bar{u}\right\|^{\gamma} \tag{3}$$

for all sufficiently large k. Under the standard conditions for q-superlinear convergence of (u^k) it seems natural to expect that (3) will hold for some M proportional to n, but already for Broyden's method—let alone the more general Broyden–like method—such results are almost non-existent in the literature (for $n > 1$), so a rigorous proof that (3) holds seems unavailable. The only result in this direction that we are aware of is [18, Theorem 4.1], but for it to apply a uniform linear independence–type assumption is required, which, in view of the violation of uniform linear independence that we observe in the numerical experiments, is why we have chosen to omit results based on [18, Theorem 4.1] in the present work. Let us, nonetheless, mention that if (3) holds and (u^k) converges at least r-superlinearly to $\bar{u}$, then (ε_k) converges at least r-superlinearly to zero, which by (1) implies that (u^k) converges quite a bit faster than q-superlinearly. In the numerical experiments we find consistently that (3) holds for $M = 0$ with a $\gamma > 1$ that depends on n and the choice of (σ_k) (but not on k if (σ_k) is constant), which is nothing else but the q-order of (u^k) that we discussed above. This would imply that (ε_k) converges at least with the respective r-order, which by (1) implies that (u^k) converges even faster than with q-order γ.[1] It is important to note that we actually prove convergence of (B_k) under a more general condition than (3), cf. Theorem 3. This enables us to prove that this conditions is satisfied for singular $F'(\bar{u})$, cf. Theorem 7 2), which is not true for (3) since in the singular case (u^k) does not converge q-superlinearly.

[1] Is this really true???

The second block uses that under well-known assumptions there holds $\sum_k \varepsilon_k^2 < \infty$. Therefore, if there are $N \in \mathbb{N}$ and $C > 0$ such that

$$\varepsilon_k \leq C \max\Big\{\varepsilon_{k-1}^2, \varepsilon_{k-2}^2, \ldots, \varepsilon_{k-N}^2\Big\} \tag{4}$$

is valid for all k sufficiently large, then $\sum_k \varepsilon_k < \infty$, so (B_k) converges. The requirement that we actually use is more general than (4), cf. Theorem 4. It is perhaps the most intriguing discovery of this work that for Broyden's method (4) seems to be satisfied with $N = 2n$ under the standard assumptions for q-superlinear convergence in a small vicinity of the root $\bar{u}$ (smaller than is needed for q-superlinear convergence). In fact, it seems that there we have

$$\varepsilon_{k+2n} \leq C\varepsilon_k^2$$

for some constant $C > 0$ and all sufficiently large k, which is to say that the updates $(\|B_{k+1} - B_k\|)$ are $2n$–step q-quadratically convergent. If this is at all true, then it appears likely that it is connected to Gay's famous theorem [13] on $2n$-step q-quadratic convergence of the iterates (u^k). In any case it is quite clear from the numerical experiments that the Broyden updates exhibit multi-step convergence of q-order greater than one, which implies in particular that they possess an r-order of convergence larger than one which matches the rate of convergence from the discussion above.

In the third block we directly involve Gay's theorem to derive a sufficient condition. A simplified version of this condition asserts convergence if there exist $M \in \mathbb{N}$ and $C > 0$ such that

$$\varepsilon_k \leq C \min\Big\{\varepsilon_{k-M-1}, \varepsilon_{k-M-2}, \ldots, \varepsilon_{k-M-2n}\Big\} \tag{5}$$

holds for all k large enough; cf. Theorem 6. We regard this as a monotonicity-type property and we point out that under the standard assumptions for q-superlinear convergence of (u^k), (ε_k) is a null sequence. Yet again, no result that implies (5) is available.

In summary, the theoretical results in combination with the numerical experiments presented in this work shed light on the convergence behavior of the Broyden–like matrices (B_k) and the iterates (u^k). In particular, they strongly suggest that these matrices converge at least r-linearly which would imply that (u^k) converges somewhat faster to $\bar{u}$ than q-superlinearly. The theoretical results contain several novel sufficient conditions for the convergence of (B_k) and some of these conditions are satisfied in *every single one* of the numerical experiments. Moreover, in the theoretical part we show that some of the novel sufficient conditions hold on certain singular and regular problems.

Let us now point out related literature. As mentioned above the approach to provide sufficient conditions for convergence and investigate in numerical experiments if these conditions are met is inspired by the existing literature on convergence of quasi-Newton matrices, where this is a common theme. For instance, this is done in [6] with uniform linear independence (ULI) as main assumption, in and in [12] with positive definiteness as main assumption. Powell in [30] shows that if ULI is satisfied, then the PSB matrices converge to the true Hessian, but in contrast to the SR1 results he applies algorithmic modifications to ensure that ULI holds. Strong convergence results are available for the DFP and BFGS matrices. In [14] it is shown that they converge under very general assumptions and in [33] the results of [14] are extended to the convex Broyden class excluding DFP. We stress that ULI is not used in [14,33]. More recent work like [4,15] is concerned with using quasi-Newton updates to invert matrices. However, the directions that are used for the matrix updates are *not* generated by a quasi-Newton method, so those works seem largely unrelated to the subject of this article.

Results on the (single-step) convergence of the iterates of quasi-Newton methods with q-order larger than 1 are extremely scarce. We are only aware of the very recent [31] that establishes such estimates for the convex Broyden class on self-concordant objectives.

This paper is organized as follows. In section 2 we establish notation and preparatory results. Section 3 develops the sufficient conditions for convergence of the Broyden–like matrices and section 4 shows that some of these conditions hold for problems with singular Jacobian. Section 5 is devoted to numerical experiments and section 6 provides a short summary of our findings.

2 Preliminaries

2.1 Notation

We use $\mathbb{N} := \{1, 2, 3, \ldots\}$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. We work in the Euclidean $\mathbb{R}^n$, $n \in \mathbb{N}$, whose norm we denote by $\|\cdot\|$. For matrices we exclusively use the spectral norm, which is also denoted by $\|\cdot\|$. We abbreviate $[n] := \{1, 2, \ldots, n\}$. For $A \in \mathbb{R}^{n\times n}$, A^j indicates the j-th row of A, regarded as a row vector. In contrast, A_k signifies an element of a sequence (A_k). The span of $C \subset \mathbb{R}^n$ is denoted by $\langle C\rangle$. If $C = \{\bar{s}\}$ for some $\bar{s} \in \mathbb{R}^n$, then we use $\langle \bar{s}\rangle$ instead of $\langle\{\bar{s}\}\rangle$. The number of elements of a finite set M is indicated by $|M|$.

2.2 The differentiability assumption

We will often require the following differentiability assumption of F.

Assumption 1 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be differentiable in a neighborhood of some $\bar{u}$ with $F(\bar{u}) = 0$ and let F' satisfy $\|F'(u) - F'(\bar{u})\| \le L\|u - \bar{u}\|^\alpha$ for all u in that neighborhood and constants $L, \alpha > 0$.*

2.3 Convergence of the Broyden–like method

To conveniently state convergence results for Algorithm BL let us introduce some notation. To this end, we remark that whenever an infinite sequence (u^k)

is generated by Algorithm BL, then $s^k \neq 0$ for all k is ensured. We use this tacitly from now on.

Definition 1 Let $F : \mathbb{R}^n \to \mathbb{R}^n$ and let (u^k), (s^k), (σ_k) and (B_k) be generated by Algorithm BL. For all $k \in \mathbb{N}_0$ denote

$$\hat{s}^k := \frac{s^k}{\|s^k\|} \qquad \text{and} \qquad \varepsilon_k := \|B_{k+1} - B_k\| = \sigma_k \frac{\left\|F(u^{k+1})\right\|}{\|s^k\|}.$$

If F is differentiable at some $\bar{u}$, then we set for $k \in \mathbb{N}_0$

$$E_k := B_k - F'(\bar{u}) \qquad \text{and} \qquad R^k_{\bar{u}} := \frac{F(u^{k+1}) - F(u^k) - F'(\bar{u})(s^k)}{\|s^k\|}.$$

To follow standard notation for quasi-Newton methods we have suppressed a superscript $\bar{u}$ in E_k. It will always be clear what $\bar{u}$ is, anyway.

Remark 1 We will often require $(\|R^k_{\bar{u}}\|)$ to converge sufficiently fast to zero. Therefore, note that due to Assumption 1 there holds

$$\left\|R^k_{\bar{u}}\right\| \le L \left\|u^k - \bar{u}\right\|^\alpha$$

for all sufficiently large k, provided that $u^k \to \bar{u}$ for $k \to \infty$. We conclude that $(\|R^k_{\bar{u}}\|)$ inherits its r-rate of convergence from (u^k). For instance, if (u^k) converges q-superlinearly to $\bar{u}$, then $(\|R^k_{\bar{u}}\|)$ converges r-superlinearly to zero.

Let us collect the convergence properties of Algorithm BL that we will use.

Theorem 1 *Let Assumption 1 hold.*

1) *If $F'(\bar{u})$ is invertible, then there is $\delta > 0$ such that for every (u^0, B_0) with $\|u^0 - \bar{u}\| \le \delta$ and $\|B_0 - F'(\bar{u})\| \le \delta$ Algorithm BL either terminates after finitely many iterations with output $u^* = \bar{u}$ or it generates a sequence (u^k) that converges q-linearly to $\bar{u}$; all B_k are invertible with $(\|B_k\|) < \infty$ and $(\|B_k^{-1}\|) < \infty$.*
2) *Let (u^k) be generated by Algorithm BL. If (u^k) satisfies*

$$\sum_{k=0}^{\infty} \left\|u^k - \bar{u}\right\|^\alpha < \infty, \tag{6}$$

then

$$\sum_{k=0}^{\infty} \left\|E_k \hat{s}^k\right\|^2 < \infty \tag{7}$$

and

$$\sum_{k=0}^{\infty} \|E_{k+1}E_{k+1}^T - E_k E_k^T\| < \infty$$

are satisfied. In particular, $(E_k E_k^T)$ converges, the singular values $(\Lambda_j(E_k))_k$, $j \in [n]$, converge and there holds $\sum_k \Lambda_1(E_k)^2 < \infty$ for the smallest singular value, $(\|E_k\|)$ converges and the row norms $(\|E_k^j\|)_k$, $j \in [n]$, converge.

3) Let (u^k) *be generated by Algorithm BL and suppose that it satisfies* (7). *If* $F'(\bar{u})$ *is invertible, then there holds*

$$\sum_{k=0}^{\infty}\left(\frac{\|u^{k+1}-\bar{u}\|}{\|u^k-\bar{u}\|}\right)^2<\infty.$$

In particular, (u^k) *converges q-superlinearly to* $\bar{u}$.
Moreover, there are constants $c, C>0$ *such that we have for all* $k\in\mathbb{N}_0$

$$c\|s^k\|\le\|F(u^k)\|\le C\|s^k\| \quad \text{and} \quad c\|u^k-\bar{u}\|\le\|F(u^k)\|\le C\|u^k-\bar{u}\|.$$

Proof The claims of 1) and 3) as well as (7) are established in [24]. We prove the remaining parts of 2). To this end, we denote

$$P_k:\mathbb{R}^n\to\mathbb{R}^n, \qquad P_k:=I-\sigma_k\hat{s}^k(\hat{s}^k)^T.$$

It is elementary to see that $E_{k+1}=E_kP_k+\sigma_kR^k_{\bar{u}}(\hat{s}^k)^T$, and this implies

$$E_{k+1}E_{k+1}^T=E_kP_kP_k^TE_k^T+\sigma_kE_kP_k\hat{s}^k(R^k_{\bar{u}})^T+\sigma_kR^k_{\bar{u}}(\hat{s}^k)^TP_k^TE_k^T+\sigma_k^2R^k_{\bar{u}}(R^k_{\bar{u}})^T.$$

Using $P_kP_k^T=I-\sigma_k(2-\sigma_k)\hat{s}^k(\hat{s}^k)^T$ and $P_k\hat{s}^k=(1-\sigma_k)\hat{s}^k$ we obtain

$$\left\|E_{k+1}E_{k+1}^T-E_kE_k^T\right\|\le\sigma_k(2-\sigma_k)\left\|E_k\hat{s}^k\right\|^2+2\sigma_k|1-\sigma_k|\left\|E_k\hat{s}^k\right\|\left\|R^k_{\bar{u}}\right\|+\sigma_k^2\left\|R^k_{\bar{u}}\right\|^2.$$

Since $t(2-t)\le 1$ for $t\in[0,2]$ and $2ab\le a^2+b^2$ for $a,b\in\mathbb{R}$, we find that

$$\left\|E_{k+1}E_{k+1}^T-E_kE_k^T\right\|\le 2\left\|E_k\hat{s}^k\right\|^2+5\left\|R^k_{\bar{u}}\right\|^2.$$

In view of (7) and $\sum_k\|R^k_{\bar{u}}\|<\infty$, the latter following from (6), this proves $\sum_k\|E_{k+1}E_{k+1}^T-E_kE_k^T\|<\infty$. We conclude that $\lim_{k\to\infty}E_kE_k^T$ exists. This, in turn, implies that the singular values of (E_k) converge for $k\to\infty$ since the eigenvalues of $(E_kE_k^T)$ converge. (For $A_k\to A$, (A_k) and A symmetric, convergence of the eigenvalues follows, e.g., from [17, Corollary 6.3.4].) Since the smallest eigenvalue of $E_k^TE_k$ is $\Lambda_1(E_k)^2$, we have

$$\Lambda_1(E_k)^2=\min_{\|v\|=1}v^TE_k^TE_kv\le\left\|E_k\hat{s}^k\right\|^2,$$

hence $\sum_k\Lambda_1(E_k)^2<\infty$ by (7). Since $\Lambda_n(E_k)=\|E_k\|$ for all k, $(\|E_k\|)$ converges, too. Since every component of the matrix $E_kE_k^T$ converges, we obtain the convergence of all $(E_k^j(E_k^i)^T)$ for $k\to\infty$, where $i,j\in[n]$. Taking $i=j$ shows that $(\|E_k^j\|^2)_k$ converges for all $j\in[n]$, which yields the claim on the row norms. Evidently, (7) implies $\lim_{k\to\infty}E_k\hat{s}^k=0$.

It remains to establish the existence of c, C in 3). Since it is well-known that $\|u^k-\bar{u}\|/\|s^k\|\to 1$ for $k\to\infty$ if (u^k) converges q-superlinearly to $\bar{u}$, cf. [8, Lemma 2.1] or [10, Lemma 8.2.3], we infer that it suffices to prove the existence of $c, C>0$ such that

$$c\|u^k-\bar{u}\|\le\|F(u^k)\|\le C\|u^k-\bar{u}\|$$

holds for all $k\in\mathbb{N}_0$. Taking $F(\bar{u})=0$ into account, the inequality to the right holds because of the differentiability of F at $\bar{u}$. The inequality to the left follows from the invertibility of $F'(\bar{u})$ by standard arguments. □

Remark 2 Since (6) holds in particular if (u^k) converges r-linearly to $\bar{u}$, we note that (6) is satisfied under the assumptions of part 1). We point out that (7) in part 2) follows from (6) without invertibility of $F'(\bar{u})$; we will use this when we treat singular problems in section 4. Regarding $(\|E_k^j\|)_k$ it can also be shown that for every $j \in [n]$

$$\sum_{k=0}^{\infty} \Big| \|E_{k+1}^j\| - \|E_k^j\| \Big| < \infty.$$

However, we do not need this stronger statement and therefore omit its proof. Lastly, note that (7) implies the Dennis–Moré condition $\lim_{k\to\infty} \|E_k \hat{s}^k\| = 0$.

2.4 Auxiliary results

In this section we collect results on the Broyden–like method that will be utilized in the convergence analysis in Section 3. In addition, we establish a connection between the convergence of the Broyden–like matrices and the superlinear convergence of the iterates, cf. Lemma 3

We start by providing conditions under which $(E_{k+1}\hat{s}^k)$ is summable.

Lemma 1 *Let Assumption 1 hold and let (u^k) be generated by Algorithm BL.*

1) Suppose that

$$\sum_{k=0}^{\infty} \|R_{\bar{u}}^k\| < \infty. \tag{8}$$

Then we have

$$\sum_k |1 - \sigma_k| \|E_k \hat{s}^k\| < \infty \qquad \Longrightarrow \qquad \sum_k \|E_{k+1}\hat{s}^k\| < \infty.$$

In particular, there holds $\sum_k \|E_{k+1}\hat{s}^k\| < \infty$ if $\sigma_k = 1$ for all k sufficiently large or if (7) and $\sum_k |1-\sigma_k|^2 < \infty$ are valid, provided that (8) is satisfied.

2) Suppose that the sequences

$$\Big(\|R_{\bar{u}}^k\|\Big) \qquad \text{and} \qquad \Big(|1 - \sigma_k| \|E_k \hat{s}^k\|\Big)$$

converge r-linearly (r-superlinearly/with r-order $p > 1$) to zero, then $(\|E_{k+1}\hat{s}^k\|)$ converges r-linearly (r-superlinearly/with r-order $p > 1$) to zero.

In particular, $(\|E_{k+1}\hat{s}^k\|)$ converges with the respective r-rate to zero if $(\|R_{\bar{u}}^k\|)$ does and $\sigma_k = 1$ for all k sufficiently large or if $(\|R_{\bar{u}}^k\|)$ and $(|1-\sigma_k|)$ converge with the respective r-rate to zero and $(\|E_k\|)$ is bounded.

Proof **Proof of 1):** From $E_{k+1} = E_k[I - \sigma_k \hat{s}^k (\hat{s}^k)^T] + \sigma_k R^k_{\bar{u}} (\hat{s}^k)^T$ we obtain

$$\left\|E_{k+1}\hat{s}^k\right\| \le |1-\sigma_k| \left\|E_k \hat{s}^k\right\| + \sigma_{\max} \left\|R^k_{\bar{u}}\right\|$$

for all $k \in \mathbb{N}_0$. Summation shows $\sum_k \|E_{k+1}\hat{s}^k\| < \infty$ if $\sum_k |1-\sigma_k| \|E_k \hat{s}^k\| < \infty$. If we use Young's inequality before taking the sum, then we find

$$\sum_{k=0}^{K} \left\|E_{k+1}\hat{s}^k\right\| \le \frac{1}{2}\sum_{k=0}^{K} |1-\sigma_k|^2 + \frac{1}{2}\sum_{k=0}^{K} \left\|E_k \hat{s}^k\right\|^2 + \sigma_{\max} \sum_{k=0}^{K} \left\|R^k_{\bar{u}}\right\|$$

for all $K \in \mathbb{N}_0$, which implies the last claim.
Proof of 2): Similar to the proof of 1). □

Remark 3 Boundedness of $(\|E_k\|)$ is addressed in part 2 of Theorem 1.

The next result connects $\|B_{k+1} - B_k\|$ to $\|E_k \hat{s}^k\|$.

Lemma 2 *Let (u^k) be generated by Algorithm BL. Then for all $k \in \mathbb{N}_0$*

$$\sigma_{\min} \left\|E_k \hat{s}^k\right\| - \sigma_{\max} \left\|R^k_{\bar{u}}\right\| \le \|B_{k+1} - B_k\| \le \sigma_{\max} \left\|E_k \hat{s}^k\right\| + \sigma_{\max} \left\|R^k_{\bar{u}}\right\|.$$

Proof We compute for all $k \in \mathbb{N}_0$

$$B_{k+1} - B_k = \sigma_k \left(y^k - B_k s^k\right) \frac{(s^k)^T}{\|s^k\|^2} = \sigma_k R^k_{\bar{u}} (\hat{s}^k)^T - \sigma_k E_k \hat{s}^k (\hat{s}^k)^T.$$

After some elementary considerations this yields the claim. □

In the sufficient conditions for the convergence of the Broyden–like matrices we will often assume that the Dennis–Moré condition holds. The previous lemma implies that this assumption is necessary.

Corollary 1 *Let Assumption 1 hold and let (u^k) be generated by Algorithm BL. Suppose that (u^k) converges to $\bar{u}$. If $\lim_{k\to\infty} B_k$ exists, then $\lim_{k\to\infty} E_k \hat{s}^k = 0$.*

Proof From $u^k \to \bar{u}$ and Assumption 1 we infer $\lim_{k\to\infty} R^k_{\bar{u}} = 0$. The convergence of (B_k) implies $\lim_{k\to\infty} \|B_{k+1} - B_k\| = 0$, so Lemma 2 yields the claim. □

We are particularly interested in rates of convergence of $(\|u^k - \bar{u}\|)$ that are faster than q-superlinear. The following result connects the convergence speed of $(\|B_{k+1} - B_k\|)$ to that of $(\|u^k - \bar{u}\|)$.

Lemma 3 *Let (u^k) be generated by Algorithm BL. Suppose that (u^k) converges q-superlinearly to some $\bar{u}$ at which F is differentiable and $F'(\bar{u})$ is invertible. Then there are constants $c, C > 0$ such that*

$$c\,\|B_{k+1} - B_k\| \le \frac{\left\|u^{k+1} - \bar{u}\right\|}{\|u^k - \bar{u}\|} \le C\,\|B_{k+1} - B_k\|$$

is satisfied for all $k \in \mathbb{N}_0$.[2]

[2] Konstanten spezifizieren!

Proof This is established in [24].

Remark 4 Under the assumptions of Lemma 3 we have, e.g., the implication

$$\sum_{k=0}^{\infty} \|B_{k+1} - B_k\| < \infty \qquad \Longrightarrow \qquad \sum_{k=0}^{\infty} \frac{\|u^{k+1} - \bar{u}\|}{\|u^k - \bar{u}\|} < \infty.$$

which shows that $\sum_k \|B_{k+1} - B_k\| < \infty$ yields a rate of convergence for (u^k) that is faster than q-superlinear, cf. also the discussion in the introduction.

3 Sufficient conditions for the convergence of the Broyden–like matrices

3.1 First set: Conditions based on the cluster points of normalized steps

The first set of sufficient conditions involves the subspace that is spanned by the cluster points of the normalized steps $\hat{s}^k$.

Definition 2 Let the sequence of steps (s^k) be generated by Algorithm BL. We denote by $\mathcal{S}$ the linear space

$$\mathcal{S} := \left\langle \left\{ s \in \mathbb{R}^n : \ s \text{ is a cluster point of } \ (\hat{s}^k) \right\} \right\rangle.$$

Remark 5 It is a surprising finding of this work that $\dim(\mathcal{S}) = 1$ holds in *all* numerical experiments. This indicates, in particular, that (s^k) frequently violates uniform linear independence, cf. the discussion in the introduction.

The following result establishes a connection between $\mathcal{S}$ and the cluster points of (E_k) in case $\dim(\mathcal{S}) = 1$. In addition, it will allow us to obtain the first convergence result for (B_k) based on $\mathcal{S}$.

Lemma 4 *Let (u^k) be generated by Algorithm BL. Let $\lim_{k\to\infty} E_k \hat{s}^k = 0$ and $\dim(\mathcal{S}) = 1$. Then there holds $\mathcal{S} \subset \ker(E)$ for every cluster point E of (E_k).*

Proof Let $K \subset \mathbb{N}$ be such that $\lim_{K\ni k\to\infty} E_k = E$. As $\lim_{k\to\infty} E_k \hat{s}^k = 0$ we obtain $\lim_{K\ni k\to\infty} E\hat{s}^k = 0$. By passing to a further subsequence we can assume without loss of generality that there is $\bar{s} \in \mathcal{S}$ with $\|\bar{s}\| = 1$ and $\lim_{K\ni k\to\infty} \hat{s}^k = \bar{s}$, which yields $E\bar{s} = 0$, hence $\bar{s} \in \ker(E)$. From $\dim(\mathcal{S}) = 1$ it follows that $\langle \bar{s} \rangle = \mathcal{S}$, thus we arrive at the claimed inclusion $\mathcal{S} \subset \ker(E)$. □

Remark 6 The fact that $\mathcal{S}$ is contained in $\ker(E)$ for *all* cluster points E of (E_k) says that all cluster points of (B_k) agree on $\mathcal{S}$ with $F'(\bar{u})$. While we have $\mathcal{S} = \ker(E) = 1$ in many of the numerical experiments of this paper, the inclusion $\mathcal{S} \subset \ker(E)$ can be strict: [22, Corollary 2] shows that $\dim(\mathcal{S}) = 1$, but $\dim(\ker(E)) \geq n - 1$ if F has at least $n - 1$ affine component functions and the corresponding rows of $F'(u^0)$ agree with those of B_0.

From the previous lemma we obtain the first sufficient condition for convergence of (B_k), albeit only for $n \in \{1, 2\}$. Let us first address the case $n = 1$. For $n = 1$ there holds $\dim(\mathcal{S}) = 1$ and the necessary condition $\lim_{k\to\infty} E_k \hat{s}^k = 0$ obviously implies $\lim_{k\to\infty} B_k = F'(\bar{u})$. In fact, for $n = 1$ it is possible to show much more and to extend the results to the case that F has only one nonlinear equation, cf. [22, Theorem 5]. For $n = 2$, we can prove the following.

Corollary 2 *Let $n = 2$ and let Assumption 1 hold. Suppose that (6) is satisfied and that $\dim(\mathcal{S}) = 1$. Then (B_k) converges.*

Proof Let $j \in [n]$. Let $\bar{s}$ with $\|\bar{s}\| = 1$ and $\mathcal{S} = \langle \bar{s} \rangle$ as well as $\tilde{s}$ with $\|\tilde{s}\| = 1$ and $\tilde{s}^T \bar{s} = 0$. Let E^j be a cluster point of (E_k^j) and set $w_j := E^j \tilde{s}$. From $E^j \bar{s} = 0$, which follows from (7), and $n = 2$ we infer that $E^j = w_j \tilde{s}^T$. As $(\|E_k^j\|)$ converges by Theorem 1 2) and $\tilde{s}$ is independent of the selected cluster point E^j, it follows that $\hat{E}^j \in \{\pm w_j \tilde{s}^T\}$ for *any* cluster point $\hat{E}$ of (E_k). Thus, (E_k^j) has at most two cluster points. Lemma 2 yields $\lim_{k\to\infty} \|E_{k+1}^j - E_k^j\| = \lim_{k\to\infty} \|B_{k+1}^j - B_k^j\| = 0$, hence (E_k^j) cannot have finitely many cluster points except if it converges, cf. [34, Lemma 10.11]. Thus, (B_k^j) converges. □

Remark 7 Using the technique of [22], cf. in particular [22, Theorem 4], the results for $n = 1$ and $n = 2$ can be extended to mappings F that have no more than two nonlinear component functions and $n - 2$ affine ones provided B_0 agrees with $F'(u^0)$ on the rows that correspond to affine components of F.

If $n > 2$, then $\dim(\mathcal{S}) = 1$ is not enough to ensure convergence of (B_k). Instead, it is necessary that $(\hat{s}^k)$ tends to $\pm\bar{s}$ sufficiently fast, where $\mathcal{S} = \langle \bar{s} \rangle$. Moreover, it seems also mandatory that σ_k tends to 1 fast enough, which is to say that Algorithm BL asymptotically turns into Broyden's method. Specifically, we have the following result and its corollary, the consequences of which for the rate of convergence of (u^k) are addressed in Remark 4.

Theorem 2 *Let Assumption 1 hold and let (u^k) be generated by Algorithm BL. Suppose that $(\|B_k\|)$ is bounded. Set $\zeta_k := \min\big\{\|\hat{s}^{k+1} - \hat{s}^k\|, \|\hat{s}^{k+1} + \hat{s}^k\|\big\}$, $k \in \mathbb{N}_0$.*

1) *If $\sum_k \|R_{\bar{u}}^k\| < \infty$, $\sum_k \|E_{k+1}\hat{s}^k\| < \infty$ and*

$$\sum_{k=0}^{\infty} \zeta_k < \infty \tag{9}$$

   are satisfied, then $\sum_k \|B_{k+1} - B_k\| < \infty$.

2) *If the sequences*

$$(\|R_{\bar{u}}^k\|), \qquad \Big(\|E_{k+1}\hat{s}^k\|\Big) \qquad \text{and} \qquad (\zeta_k)$$

   converge r-linearly (r-superlinearly/with r-order $p > 1$) to zero, then $(\|B_{k+1} - B_k\|)$ converges r-linearly (r-superlinearly/with r-order $p > 1$) to zero.

Proof By use of the triangle inequality we obtain for all $k \in \mathbb{N}$

$$\left\|E_k \hat{s}^k\right\| \leq \|E_k\| \min\left\{\left\|\hat{s}^k - \hat{s}^{k-1}\right\|, \left\|\hat{s}^k + \hat{s}^{k-1}\right\|\right\} + \left\|E_k \hat{s}^{k-1}\right\|.$$

The assumptions therefore imply $\sum_k \|E_k \hat{s}^k\| < \infty$, from which we the assertion follows by use of Lemma 2 and $\sum_k \|R_{\bar{u}}^k\| < \infty$. The proof of 2) is similar. □

Remark 8 Lemma 1 contains sufficient conditions for the summability, respectively, convergence with r-rate of $(\|E_{k+1}\hat{s}^k\|)$. In contrast, conditions that imply (9) or one of the stronger assumptions for (ζ_k) from 2) are unknown except for certain cases that we specify in Remark 10.

Condition (9) is less demanding than the condition (2) discussed in the introduction, but it still implies $\dim(\mathcal{S}) = 1$.

Corollary 3 *Let $(\hat{s}^k) \subset \mathbb{R}^n$ be a sequence with $\|\hat{s}^k\| = 1$ for all k. Then:*

1) (9) implies the existence of a vector $\bar{s}$ with $\|\bar{s}\| = 1$ and $\mathcal{S} = \langle \bar{s} \rangle$.
2) If there is a vector $\bar{s}$ with $\|\bar{s}\| = 1$ such that

$$\sum_{k=0}^{\infty} \min\left\{\left\|\hat{s}^k - \bar{s}\right\|, \left\|\hat{s}^k + \bar{s}\right\|\right\} < \infty \tag{10}$$

is satisfied, then (9) holds and $\mathcal{S} = \langle \bar{s} \rangle$.
3) If there is a vector $\bar{s}$ with $\|\bar{s}\| = 1$ such that $\bar{\zeta}_k := \min\{\|\hat{s}^k - \bar{s}\|, \|\hat{s}^k + \bar{s}\|\}$, $k \in \mathbb{N}_0$, converges r-linearly (r-superlinearly/with r-order $p > 1$) to zero, then so does (ζ_k) defined in Theorem 2 and there holds $\mathcal{S} = \langle \bar{s} \rangle$.

Proof **Proof of 1):** Let (9) be satisfied. We can inductively replace $\hat{s}^k$ by $-\hat{s}^k$ if necessary to obtain a sequence $(\tilde{s}^k)$ with $\tilde{s}^k \in \{\pm \hat{s}^k\}$ for all k and such that

$$\left\|\tilde{s}^{k+1} - \tilde{s}^k\right\| = \min\left\{\left\|\hat{s}^{k+1} - \hat{s}^k\right\|, \left\|\hat{s}^{k+1} + \hat{s}^k\right\|\right\}$$

for all k. The sequence $(\tilde{s}^k)$ thus satisfies $\sum_k \|\tilde{s}^{k+1} - \tilde{s}^k\| < \infty$ and is therefore convergent. Its limit, say $\bar{s}$, satisfies $\|\bar{s}\| = 1$, and by construction this implies that $(\hat{s}^k)$ can only have the cluster points $\pm\bar{s}$, so $\mathcal{S} = \langle \bar{s} \rangle$.
Proof of 2): (9) follows from (10) by the triangle inequality. Moreover, (10) implies that the cluster points of $(\hat{s}^k)$ belong to $\{\pm\bar{s}\}$, which yields $\mathcal{S} = \langle \bar{s} \rangle$.
Proof of 3): Similar to the proof of 2). □

Remark 9 We stress that (9) and (10) are satisfied in *all* numerical experiments conducted for this work. Still, let us mention that (10) can fail while (9) holds. This is the case, e.g., for $(\hat{s}^k) \subset \mathbb{R}^n$ with $\hat{s}^k := (\sqrt{1 - t_k^2}, t_k, 0, 0, \ldots, 0)^T$ and $\bar{s} := (1, 0, \ldots, 0)^T$, where $t_k := k^{-2}\sqrt{2k^2 - 1}$ for $k \in \mathbb{N}$.

Remark 10 Neither Theorem 2 nor Corollary 3 involve invertibility of $F'(\bar{u})$ or superlinear convergence of (u^k). In Theorem 7 we show that if $F'(\bar{u})$ is singular of a certain type and (u^0, B_0) is suitable, then (10) holds. This is a situation in which (u^k) converges q-linearly, but not faster. Furthermore, the results of [22, Lemma 3] show that if F has $n-1$ affine component functions and the corresponding $n-1$ rows of $F'(u^0)$ agree with those of B_0, then $(s^k)_{k\geq 1}$ is restricted to a one-dimensional subspace, so for $k \geq 1$ all summands in (9) and (10) vanish. In summary, the conditions developed in Theorem 2 and Corollary 3 are satisfied in all numerical experiments and can be rigorously proven in certain situations.

3.2 Second set: Three further sufficient conditions

In the remaining theoretical investigations we address sufficient conditions that are independent of $\mathcal{S}$, so they are complementary to those of section 3.1. To formulate these conditions conveniently, let us introduce the following quantity.

Definition 3 Let $(m_k) \subset \mathbb{N}_0$. We define $\mathfrak{C}\big((m_k)\big) \in \mathbb{N} \cup \{+\infty\}$ by

$$\mathfrak{C}\big((m_k)\big) := \sup_{M\in\mathbb{N}_0} \Big|\big\{k \in \mathbb{N}_0 : \;\; m_k = M\big\}\Big|.$$

Remark 11 $\mathfrak{C}\big((m_k)\big)$ provides an upper bound on how many times each natural number appears in the sequence (m_k). We will often require $\mathfrak{C}\big((m_k)\big) < \infty$.

All three sufficient conditions to come invoke the following lemma.

Lemma 5 *Let (u^k) be generated by Algorithm BL. Suppose that there exist sequences $(\alpha_k) \subset (0,\infty)$ and $(m_k) \subset \mathbb{N}_0$ with $\mathfrak{C}\big((m_k)\big) < \infty$ such that*

$$\varepsilon_k \leq \alpha_{m_k} \qquad \forall k \in \mathbb{N}_0.$$

Also assume that

$$\sum_{k=0}^{\infty} \alpha_k < \infty.$$

Then $\sum_k \varepsilon_k < \infty$.

Proof Summation shows for $K \in \mathbb{N}_0$ that

$$\sum_{k=0}^{K} \varepsilon_k \leq \sum_{k=0}^{K} \alpha_{m_k} \leq \mathfrak{C}\big((m_k)\big) \sum_{k=0}^{\infty} \alpha_k < \infty.$$

The right-hand side is independent of K, so the claim follows. □

3.2.1 Condition 1: A relationship between the updates and the steps/iterates

The following condition includes, in particular, condition (3) from the introduction. The *q-order/r-order of convergence* is defined in [29, Chapter 9].

Theorem 3 *Let (u^k) be generated by Algorithm BL. Suppose that there exist a constant $C > 0$ and sequences $(m_k) \subset \mathbb{N}_0$ and $(\gamma_k) \subset (0, \infty)$ such that $\mathfrak{C}\big((m_k)\big) < \infty$ and*

$$\varepsilon_k \leq C \, \|s^{m_k}\|^{\gamma_{m_k}} \tag{11}$$

are satisfied for all k sufficiently large.

1) If there holds

$$\sum_{k=0}^{\infty} \|s^k\|^{\gamma_k} < \infty, \tag{12}$$

then $\sum_k \varepsilon_k < \infty$.

2) If $\gamma_k \geq \gamma > 0$ for all k sufficiently large, then r-superlinear convergence [r-linear convergence with rate $\kappa \in (0,1)$] of (s^{m_k}) to zero implies r-superlinear convergence [r-linear convergence with rate $\kappa^\gamma \in (0,1)$] of (ε_k).

3) If (11) holds for $(m_k) \equiv k$ and $\gamma_k \geq \gamma > 0$ for all k sufficiently large, if (u^k) converges to some $\bar{u}$ and there is $\hat{C} > 0$ such that $\|s^k\| \leq \hat{C}\|u^k - \bar{u}\|$ holds for all k sufficiently large, and if F is differentiable at $\bar{u}$ with $F'(\bar{u})$ invertible, then

$$\limsup_{k\to\infty} \frac{\|u^{k+1} - \bar{u}\|}{\|u^k - \bar{u}\|^{1+\gamma}} \leq \frac{C\hat{C}^{1+\gamma} \, \|F'(\bar{u})^{-1}\|}{\sigma_{\min}}, \tag{13}$$

i.e., (u^k) converges with q-order at least $1 + \gamma$. Moreover, (ε_k) then converges with r-order at least $1 + \gamma$.

Proof **Proof of 1):** This follows from Lemma 5, applied with $\alpha_k := \|s^k\|^{\gamma_k}$.
Proof of 2): It is easy to check that if $(\|s^{m_k}\|)$ converges r-linearly with rate $\kappa \in (0,1)$ [r-superlinearly] to zero, then $(C\|s^{m_k}\|^\gamma)$, where $C, \gamma > 0$ are constants, has the same property with rate κ^γ.
Proof of 3): Since $s^k \to 0$ due to $u^k \to \bar{u}$, (11) implies

$$\sigma_{\min} \frac{\|F(u^{k+1}) - F(\bar{u})\|}{\hat{C}^{1+\gamma} \, \|u^k - \bar{u}\|^{1+\gamma}} \leq \sigma_k \frac{\|F(u^{k+1})\|}{\|s^k\|^{1+\gamma}} \leq C$$

for all k sufficiently large, where we used that $F(\bar{u}) = 0$ due to (11). It is elementary to infer from $u^k \to \bar{u}$ that $\limsup_{k\to\infty} \frac{\|u^{k+1} - \bar{u}\|}{\|F(u^{k+1}) - F(\bar{u})\|} \leq \|F'(\bar{u})^{-1}\|$, so (13) follows. The claim about the q-order is obviously true. By [22, Lemma 1] this implies that (ε_k) converges with r-order $1 + \gamma$. □

Remark 12 Using $\|s^k\| \le \|u^{k+1} - \bar{u}\| + \|u^k - \bar{u}\|$ we can easily formulate analogue statements with s^{m_k} replaced by $u^{m_k} - \bar{u}$. For instance, it follows readily that (12) holds if $\sum_k \|u^k - \bar{u}\|^\gamma < \infty$ is satisfied for some $\gamma > 0$, which is (6) for $\gamma = \alpha$. Also note that (s^{m_k}) converges r-linearly [r-superlinearly] to zero if $(u^{m_k} - \bar{u})$ does.

Remark 13 Conditions that ensure either (11) and (12) or (11) and the existence of γ are unknown. In Corollary 4 we prove for the Broyden updates and $(m_k) \equiv k$ that at least one of the $2n$ values $\varepsilon_k, \varepsilon_{k+1}, \dots, \varepsilon_{k+2n-1}$ satisfies (11) for $\gamma_k = \frac{1}{2n}$.

3.2.2 Condition 2: Multi-step q-order convergence of the Broyden–like updates

The next sufficient condition is motivated by the numerical observation that for Broyden's method the updates, while not monotone in their decline to zero, appear to converge with a $2d$–step q-order of convergence (in experiments with small n this order is at least 2, which means that the Broyden updates converge $2n$–step q-quadratically just like the iterates do by Gay's theorem), cf. the discussion in the introduction, specifically (4), and Remark 14. For the Broyden–like updates this does not seem to be true, but it is not difficult to include them in the following result anyway.

Theorem 4 *Let Assumption 1 hold and let (u^k) be generated by Algorithm BL. Let (7) and $\sum_k \|R^k_{\bar{u}}\|^2 < \infty$ be satisfied.*

1) If there exist constants $C, \delta > 0$ and a sequence $(m_k) \subset \mathbb{N}_0$ such that $\mathfrak{C}\big((m_k)\big) < \infty$ and

$$\varepsilon_k \le C \varepsilon_{m_k}^{1+\delta} \tag{14}$$

are satisfied for all sufficiently large k, then $\sum_k \varepsilon_k < \infty$.

2) If there exist constants $C, \delta > 0$, $M \in \mathbb{N}$, and a sequence $(m_k) \subset \mathbb{N}_0$ such that $\mathfrak{C}\big((m_k)\big) < \infty$ and

$$\varepsilon_k \le C \max\Big\{\varepsilon_{m_k}^{1+\delta}, \varepsilon_{m_k-1}^{1+\delta}, \dots, \varepsilon_{m_k-M}^{1+\delta}\Big\} \tag{15}$$

are satisfied for all sufficiently large k, then $\sum_k \varepsilon_k < \infty$.

3) If $N := \sup_{k \in \mathbb{N}_0} |m_k - k|$ is finite in 1), then (ε_k) converges with r-order at least $(1+\delta)^N$. If N is finite in 2), then (ε_k) converges with r-order at least $(1+\delta)^{N+M}$.

Proof **Proof of 1):** We observe first that by Lemma 2 and Young's inequality the assumptions imply $\sum_k \varepsilon_k^2 < \infty$. For $\delta = 1$ we thus obtain 1) from Lemma 5 with $\alpha_k := C\varepsilon_k^{1+\delta} = C\varepsilon_k^2$. Since $\varepsilon_k \to 0$ due to $\sum_k \varepsilon_k^2 < \infty$, this also implies that 1) is true for any $\delta > 1$. If (14) holds for some $\delta \in (0,1)$, then there are $\tilde{C} > 0$ and $(\tilde{m}_k) \subset \mathbb{N}_0$ with $\mathfrak{C}\big((\tilde{m}_k)\big) < \infty$ such that (14) holds for exponent 2, so the claim follows from the already proven case $\delta = 1$.

Proof of 2): We show that 2) follows from 1). Indeed, let for each sufficiently large k the number $\tilde{m}_k \in \{m_k, m_k - 1, \dots, m_k - M\}$ be an index that realizes the maximum in (15). This defines a sequence $(\tilde{m}_k)$ that satisfies $\mathfrak{C}\big((\tilde{m}_k)\big) < \infty$ and (14) (with m_k replaced by $\tilde{m}_k$).

Proof of 3): The claims follow from the elementary fact that if $(v_k) \subset [0, \infty)$ satisfies $v_k \leq C v_{k-L(k)}^{1+\delta}$ for constants $C, \delta > 0$ and a bounded sequence $L(k) \subset \mathbb{N}_0$, then (v^k) also converges with r-order $(1+\delta)^L$ for $L := \max_{k \in \mathbb{N}_0} L(k)$. □

Remark 14 In the numerical experiments with regular $F'(\bar{u})$ and $(\sigma_k) \equiv 1$ (Broyden's method) we observe that (14) holds for $m_k = k - 2d$. The same is true for $m_k = k - 2n$ if $\sigma_k \to 1$ or $\sigma_k = 1$ for all k sufficiently large. If F has only one nonlinear component function and $n-1$ affine ones and if B_0 agrees with $F'(u^0)$ on the rows that correspond to affine components of F and an additional assumption holds, then convergence of the Broyden updates with q-order $\frac{\sqrt{5}+1}{2}$ is established in [22, Theorem 5 1)]. Further theoretical results that confirm either (14) or (15) are not available for the time being.

3.2.3 Condition 3: Multi–step q-quadratic convergence of the iterates

In this section we derive a sufficient condition for convergence of the Broyden–matrices from the following generalization of Gay's result [13, Theorem 3.1] on $2n$-step q-quadratic convergence.

Theorem 5 *Let Assumption 1 hold with $\alpha = 1$ and let $F'(\bar{u})$ be invertible. Moreover, let $J \subset [n]$ be a set of indices (possibly empty) such that F_j is affine for $j \in J$ and such that $F_j'(u^0) = B_0^j$ for all $j \in J$, where B_0^j is the j-th row of B_0. Let $d := n - |J|$ (with $|J| := 0$ if $J = \emptyset$). Then there exist $\delta > 0$ and $C > 0$ such that for all (u^0, B_0) with $\|u^0 - \bar{u}\| \leq \delta$ and $\|B_0 - F'(\bar{u})\| \leq \delta$, Algorithm BL with $\sigma_{\min} = \sigma_{\max} = 1$ either terminates with output $u^* = \bar{u}$ or it generates a sequence (u^k) that satisfies*

$$\left\|u^{k+2d} - \bar{u}\right\| \leq C \left\|u^k - \bar{u}\right\|^2 \qquad \forall k \in \mathbb{N}. \tag{16}$$

If (u^k) satisfies (16), then it converges with r-order at least $2^{\frac{1}{2d}}$ and its q-order is no larger.

Proof (16) is established in [21, Theorem 3] and from this the maximal r-order $2^{\frac{1}{2d}}$ is elementary to conclude, cf. also [29, E 9.2-4.]. The claim on the q-order follows from the fact that the q-order is never larger than the r-order, cf. [29, 9.3.2.]. □

To derive a sufficient condition for convergence of the Broyden matrices we will use the following consequence of Theorem 5.

Corollary 4 *Under the assumptions of Theorem 5 there is a constant $\hat{C} > 0$ such that for each $k \in \mathbb{N}$ there holds at least one of the $2d$ inequalities*

$$\varepsilon_{k+j} \leq \hat{C} \|s^k\|^{\frac{1}{2d}}, \qquad j = 0, \dots, 2d-1.$$

In particular, (u^k) converges q-superlinearly, then (ε_k) contains an r-superlinearly convergent subsequence (ε_{k_j}) with $k_{j+1} - k_j \leq 4d$ for all $j \in \mathbb{N}_0$.

Proof From part 3) of Theorem 1 we obtain constants $c, C > 0$ such that

$$c\|s^k\| \le \left\|F(u^k)\right\| \le C\|s^k\| \;\text{ and }\; c\left\|u^k - \bar{u}\right\| \le \left\|F(u^k)\right\| \le C\left\|u^k - \bar{u}\right\| \tag{17}$$

hold for all $k \in \mathbb{N}_0$. Hence, it suffices to show that at least one of the inequalities

$$\frac{\left\|F(u^{k+1+j})\right\|}{\|F(u^{k+j})\|} \le \bar{C}\left\|F(u^k)\right\|^{\frac{1}{2d}}, \quad j = 0, \ldots, 2d-1$$

must hold for each $k \in \mathbb{N}$ and some constant $\bar{C} > 0$. Suppose to the contrary that for each $\bar{C} > 0$ there is a $k \in \mathbb{N}$ such that none of the $2d$ inequalities is satisfied. Then for any $\bar{C}$ and the associated k we have

$$\prod_{j=0}^{2d-1} \frac{\left\|F(u^{k+1+j})\right\|}{\|F(u^{k+j})\|} > \bar{C}^{2d}\left\|F(u^k)\right\|, \quad \text{hence} \quad \left\|F(u^{k+2d})\right\| > \bar{C}^{2d}\left\|F(u^k)\right\|^2.$$

In view of (17) we can therefore find for any $\tilde{C} > 0$ a $k \in \mathbb{N}$ such that

$$\left\|u^{k+2d} - \bar{u}\right\| > \tilde{C}\left\|u^k - \bar{u}\right\|^2$$

is satisfied, which contradicts Theorem 5.

The additional claim follows from $\|s^k\| \le 2\|u^k - \bar{u}\|$ and the fact that $\|u^k - \bar{u}\|^{\frac{1}{2d}}$ is still q-superlinearly convergent. □

Based on Corollary 4 we can derive the following sufficient condition.

Theorem 6 *Let the assumptions of Theorem 5 hold and suppose that there exist a constant $C > 0$ and a sequence $(m_k) \subset \mathbb{N}_0$ such that $\mathfrak{C}\big((m_k)\big) < \infty$ and*

$$\varepsilon_k \le C \min\Big\{\varepsilon_{m_k-1}, \varepsilon_{m_k-2}, \ldots, \varepsilon_{m_k-2d}\Big\} \tag{18}$$

are satisfied for all sufficiently large k. Then $\sum_k \varepsilon_k < \infty$.

If $N := \sup_{k\in\mathbb{N}_0} |m_k - k| < \infty$, then (ε_k) has r-order at least $(\frac{2d+1}{2d})^{N+2d}$.

Proof Using Corollary 4 we obtain for all sufficiently large $k \in \mathbb{N}$, say $k \ge K$,

$$\varepsilon_k \le C \min\big\{\varepsilon_{m_k-1}, \varepsilon_{m_k-2}, \ldots, \varepsilon_{m_k-2d}\big\} \le C\|s^{m_k-2d}\|^{\frac{1}{2d}},$$

so the first claim follows from Theorem 3 1). The proof of the additional part follows from the same argument as the proof of Theorem 4 3). □

Remark 15 For divergence of (B_k) both (15) and (18) must be violated.

4 Application to a class of singular problems

In this section we show that Theorem 2 with the stronger summability property (10) and Theorem 3 can be applied to certain singular problems to obtain convergence of (B_k). The results of this section extend findings from [23], where only Broyden's method is addressed. We stress that the following convergence analysis of (B_k) builds on the convergence analysis of (u^k) presented in [7]. In particular, the assumptions on the singularity of the problem coincide. They read as follows.

Assumption 2 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be differentiable in a neighborhood of $\bar{u}$ and twice differentiable at $\bar{u}$, where $\bar{u}$ satisfies $F(\bar{u}) = 0$. Moreover, suppose that the following conditions are satisfied:*

- *There is $\phi \in \mathbb{R}^n$ with $\|\phi\| = 1$ such that $N := \ker(F'(\bar{u})) = \langle\phi\rangle$, where $\langle\phi\rangle$ denotes the linear hull of ϕ.*
- *There holds $P_N(F''(\bar{u})(\phi, \phi)) \neq 0$, where $P_N : \mathbb{R}^n \to \mathbb{R}^n$ denotes the orthogonal projection onto N, i.e., $P_N(v) = v^T\phi\phi$ for all $v \in \mathbb{R}^n$.*

Remark 16 Assumption 2 implies Assumption 1 with $\alpha = 1$.

We will use the range space $X := \operatorname{img}(F'(\bar{u}))$ and the orthogonal projection $P_X : \mathbb{R}^n \to \mathbb{R}^n$ onto X. Also, for $\bar{u} \in \mathbb{R}^n$ and $(\rho, \theta) \in (0, \infty) \times (0, \infty)$ let

$$W_{\bar{u}}(\rho, \theta) := \Big\{u \in \mathbb{R}^n : \ \|P_X(u - \bar{u})\| \leq \theta \|P_N(u - \bar{u})\|\Big\} \cap \mathbb{B}_\rho(\bar{u}),$$

where $\mathbb{B}_\rho(\bar{u})$ is the open ball of radius ρ centered at $\bar{u}$.

The following results are obtained in [7].

Lemma 6 *Let Assumption 2 hold and let $\mu_X, \mu_N > 0$. Then there exist $\rho > 0$ and $\theta > 0$ such that for all pairs $(u^0, B_0) \in W_{\bar{u}}(\rho, \theta) \times \mathbb{R}^{n\times n}$ that satisfy*

$$\left\|\left(B_0 - F'(u^0)\right) P_X(v)\right\| \leq \mu_X \rho \qquad \text{and} \qquad \left\|\left(B_0 - F'(u^0)\right) P_N(v)\right\| \leq \mu_N \rho^2$$

for all $v \in \mathbb{R}^n$ with $\|v\| = 1$, Algorithm BL with $(\sigma_k) \equiv 1$ (Broyden's method) generates a sequence (u^k) with

$$\lim_{k\to\infty} \frac{\left\|u^{k+1} - \bar{u}\right\|}{\|u^k - \bar{u}\|} = \frac{\sqrt{5} - 1}{2} \qquad \text{and} \qquad \lim_{k\to\infty} \frac{\left\|P_X(u^k - \bar{u})\right\|}{\left\|P_N(u^k - \bar{u})\right\|^2} = 0.$$

Moreover, there holds for all $k \geq 1$

$$P_N(u^{k+1} - \bar{u}) = \lambda_k P_N(u^k - \bar{u}), \tag{19}$$

where $(\lambda_k)_{k\geq1}$ satisfies $(\lambda_k)_{k\geq1} \subset \left(\frac{3}{8}, \frac{4}{5}\right)$ and $\lim_{k\to\infty} \lambda_k = \frac{\sqrt{5}-1}{2}$.

Proof All claims are established in [7]: (1.13) in [7] gives the left limit and from (2.13) together with $\theta_n \to 0$ we deduce the right limit. (The assumption $\mathbb{R}^n = X \oplus N$ imposed as part of (1.4) in [7] is superfluous and the index "$n+1$" that appears in (1.14) of [7] is a misprint that should read "n".) The claim (19) is contained in [7, (2.14)]. □

The following result applies in particular to the setting of Lemma 6.

Theorem 7 *Let Assumption 2 hold. Let (u^k) be generated by Algorithm BL with (σ_k) satisfying $\sum_k |1-\sigma_k|^2 < \infty$. Suppose that (u^k) satisfies*

$$\frac{\|u^{k+1}-\bar{u}\|}{\|u^k-\bar{u}\|} \le \kappa \qquad \text{and} \qquad \lim_{k\to\infty} \frac{\|P_X(u^k-\bar{u})\|}{\|P_N(u^k-\bar{u})\|^2} = 0 \tag{20}$$

for some $\kappa \in (0,1)$ and all k sufficiently large. Then:

1) The assumptions of Theorem 2 are fulfilled and for $k \to \infty$ there holds $\bar{\zeta}_k := \min\{\|\hat{s}^k-\phi\|, \|\hat{s}^k+\phi\|\} = o(\|u^k-\bar{u}\|)$. In particular, $(\bar{\zeta}_k)$ converges at least r-superlinearly to zero, (10) is satisfied for $\bar{s} := \phi$, and $\mathcal{S} = N$.
2) The assumptions of Theorem 3 are valid for $m_k := k$ and $\gamma_k := 1$, $k \in \mathbb{N}_0$.
3) We have $\sum_k \varepsilon_k < \infty$ and (ε_k) converges at least r-linearly with rate κ.

Proof **Proof of 1):** The claim $\bar{\zeta}_k = o(\|u^k-\bar{u}\|)$ is [23, Theorem 2, part 1)] (we remark that although [23] addresses Broyden's method, the result also holds for the Broyden–like method). It implies r-superlinear convergence of $(\bar{\zeta}_k)$ since (u^k) converges q-linearly to $\bar{u}$ by (20). As a simple consequence we now obtain that (10) holds for $\bar{s} = \phi$. In turn, this yields (9) and $\mathcal{S} = \langle\phi\rangle = N$ by part 2) of Corollary 3. Lastly, we verify the remaining assumptions of Theorem 2.

- The boundedness of $(\|B_k\|)$ follows from part 2) of Theorem 1, which is applicable because (u^k) converges q-linearly to $\bar{u}$ by (20).
- Inequality (8) also follows from the linear convergence of (u^k).
- The boundedness $\sum_k \|E_{k+1}\hat{s}^k\| < \infty$ is implied by Lemma 1.

Proof of 2): We verify the assumptions of Theorem 3. Evidently, the choice of (m_k) implies that $\mathfrak{C}((m_k)) = 1 < \infty$. Since (u^k) converges q-linearly, (12) holds. Estimate (11) follows from [23, Theorem 2, part 2)].
Proof of 3): Both Theorem 2 and Theorem 3 imply $\sum_k \varepsilon_k < \infty$. The r-linear convergence of (ε_k) follows from the second part of Theorem 3 as $(\gamma_k) \equiv 1$. □

The majority of convergence results for (B_k) from [23] can be extended to Algorithm BL provided $\sum_k (1-\sigma_k)^2 < \infty$. Exemplarily, let us demonstrate this for [23, Corollary 3], which states q-linear convergence of (ε_k).

Theorem 8 *Let Assumption 2 hold. Let (u^k) be generated by Algorithm BL with (σ_k) satisfying $\sum_k |1-\sigma_k|^2 < \infty$. Suppose that (u^k) satisfies (19) as well as*

$$\lim_{k\to\infty} \frac{\|u^{k+1}-\bar{u}\|}{\|u^k-\bar{u}\|} = \kappa \qquad \text{and} \qquad \lim_{k\to\infty} \frac{\|P_X(u^k-\bar{u})\|}{\|P_N(u^k-\bar{u})\|^2} = 0$$

for some $\kappa \in (0,1)$. Then we have

$$\lim_{k\to\infty} \frac{\varepsilon_{k+1}}{\varepsilon_k} = \kappa.$$

Proof This can be argued almost identically as in [23, Corollary 3]. □

5 Numerical experiments

We study the validity of the developed sufficient conditions on numerical examples. In section 5.1 we discuss the design of the experiments, while section 5.2 contains the examples and results.

5.1 Design of the experiments

5.1.1 Implementation and accuracy

The experiments are carried out in MATLAB 2017A using the *variable precision arithmetic (vpa)* with a precision of 1000 digits. The termination criterion $F(u^k) = 0$ in Algorithm BL is replaced by $\|F(u^k)\| \leq 10^{-320}$. By $\bar{k} \in \mathbb{N}_0$ we denote the final value of k in Algorithm BL.

5.1.2 Known solution and random initialization

All examples have $\bar{u} = 0$ as a solution and the experiments are set up in such a way that convergence to $\bar{u}$ takes place in all runs. Except for the last example, $F'(\bar{u})$ is invertible. The initial point u^0 is always generated using `rand`. It has random entries in $[-\alpha, \alpha]$, where $\alpha \in [10^{-10}, 0.1]$ will be specified for each example. For B_0 we choose $B_0 = F'(u^0) + \hat{\alpha}\|F'(u^0)\|R$, where $R \in \mathbb{R}^{n\times n}$ is a random matrix with entries in $[-1, 1]$ and $\hat{\alpha} \in \{0\} \cup [10^{-10}, 0.1]$ is example-dependent.

5.1.3 Quantities of interest

For (u^k), (s^k) and (B_k) from Algorithm BL we define

$$F_k := F(u^k), \qquad \delta_k := \frac{\ln\bigl(\|F_k\|\bigr)}{\ln\bigl(\|s^{k-1}\|\bigr)}, \qquad \epsilon_k := \|B_k - B_{k-1}\|, \quad \rho_\epsilon^k := \sqrt[k+1]{\epsilon_k},$$

$$\beta_k := \frac{\epsilon_k}{\epsilon_{k-2d}^2} \qquad \text{and} \qquad R_k := \frac{\ln\bigl(\epsilon_k\bigr)}{\ln\bigl(\epsilon_{k-2d}\bigr)}, \tag{21}$$

where d plays the same role as in Theorem 5. We regard $d = n$ as the standard choice and mention only if $d \neq n$ is selected. Also, we set

$$\zeta_k := \min\Bigl\{\|\hat{s}^k - \hat{s}^{k-1}\|, \|\hat{s}^k + \hat{s}^{k-1}\|\Bigr\} \qquad \text{and} \qquad \rho_\zeta^k := \sqrt[k+1]{\zeta_k}$$

as well as

$$\bar{\zeta}_k := \min\Bigl\{\|\hat{s}^k - \hat{s}^{\bar{k}}\|, \|\hat{s}^k + \hat{s}^{\bar{k}}\|\Bigr\} \quad \text{and} \quad \bar{\rho}_\zeta^k := \sqrt[k+1]{\bar{\zeta}_k}. \tag{22}$$

Whenever any of these quantities is undefined we set it to -1; e.g., $\delta_0 := -1$.

Let us point out some aspects that we want to study. To this end, we first remark that (u^k) converges at least q-linearly in all experiments, so we tacitly assume in the following discussion that (6) is satisfied.

- We want to assess if $\|E_k\| \to 0$ for $k \to \infty$. Since we find in *all* experiments that $\|E_k\| \not\to 0$, we always have $\lim_{k\to\infty} B_k \neq F'(\bar{u})$. This implies that (s^k) is *never* uniformly linearly independent, cf. the discussion in the introduction.
- It is easy to see that if $\delta_k \geq \delta$ for some $\delta > 1$ and all k sufficiently large, then (11) is satisfied for $m_k = k$ and $\gamma_k = \delta - 1$. In particular, the last part of Theorem 3 can be applied if $F'(\bar{u})$ is invertible, which implies that $(\|F_k\|)$, $(\|s^k\|)$ and $(\|u^k - \bar{u}\|)$ converge to zero with q-order at least δ and (ϵ_k) goes to zero with r-order at least δ.
- Theorem 4 1) implies that (B_k) converges if (β_k) is bounded.
- From (14) it follows that an estimate of the form $\epsilon_k \leq C\epsilon_{k-2d}^{1+\delta}$ for a constant $\delta > 0$ and all k sufficiently large is sufficient for convergence of (B_k). Such an estimate implies $R_k \geq 1 + \delta$ for all k sufficiently large and we are therefore interested to see whether R_k stays safely above 1 for large k.
- We use $\bar{\zeta}_k$ as approximation of $\min\{\|\hat{s}^k - \bar{s}\|, \|\hat{s}^k + \bar{s}\|\}$ which appears in (10). Observe that $\bar{\zeta}_{\bar{k}} = 0$ by definition.
- We include the three smallest singular values Λ_1^k, Λ_2^k and Λ_3^k in the results. From Theorem 1 we know that (Λ_1^k) converges to zero and that each (Λ_j^k), $j \in [n]$, converges. Furthermore, if F has d affine component functions and the corresponding d rows of B_0 match those of $F'(u^0)$, then d singular values remain exactly zero throughout the entire algorithm, cf. [22, Lemma 3]. It is now interesting to observe that in *all* numerical experiments the number of singular values converging to zero is *exactly* 1, respectively, d. Since $n > 1$ and $d < n$ this shows again that (B_k) does not converge to $F'(\bar{u})$.

5.1.4 Single run and cumulative run

For each example we perform at least one *single run* and one *cumulative run*. In single runs we display the quantities of interest during the course of the algorithm. In cumulative runs we perform $m := 2000$ single runs with the initial data varying according to section 5.1.2. With the cumulative run we want to gauge the *worst-case behavior* of Algorithm BL with respect to the quantities of interest. To explain this in more detail, consider ϵ_k and ζ_k. We recall from $\epsilon_k = \|B_k - B_{k-1}\|$, respectively, Lemma 3, that (B_k) converges if $\sum_k \epsilon_k < \infty$, respectively, if $\sum_k \zeta_k < \infty$. In single runs we provide $\rho_\epsilon^k = \sqrt[k+1]{\epsilon_k}$ and $\rho_\zeta^k = \sqrt[k+1]{\zeta_k}$ for $k \in \mathbb{N}_0$ to assess if the series $\sum_k \epsilon_k$ and $\sum_k \zeta_k$ converge. In addition, we compute

$$\rho_\epsilon^j := \max_{k_0(j)\leq k\leq \bar{k}(j)} \rho_\epsilon^k \qquad \text{and} \qquad \rho_\zeta^j := \max_{k_0(j)\leq k\leq \bar{k}(j)} \rho_\zeta^k$$

in each of the m single runs of the cumulative run, where $j \in [m]$ indicates the respective single run and we use here and in the remainder of this work the value $k_0(j) := \lfloor 0.75\bar{k}(j)\rfloor$. For the cumulative run we display

$$\rho_\epsilon := \max_{j\in[m]} \rho_\epsilon^j \qquad \text{and} \qquad \rho_\zeta := \max_{j\in[m]} \rho_\zeta^j,$$

and we conclude that (B_k) converges in all of the m single runs if ρ_ϵ, respectively, ρ_ζ are somewhat smaller than 1. In the same manner as ρ_ζ we define $\bar\rho_\zeta$ from $\bar\zeta_k$. Similar considerations apply to the following quantities that we use to display the results of cumulative runs. As before we let $j \in [m]$ indicate the respective single run. Also, we set $K(j) := \{k_0(j), k_0(j)+1, \ldots, \bar{k}(j)\}$.

$$\|F\| := \max_{j\in[m]} \|F(u^{\bar{k}(j)})\| \qquad \text{and} \qquad \|E\| := \min_{j\in[m]} \min_{k\in[\bar{k}]} \|E_k\|$$

as well as

$$\delta := \min_{j\in[m]} \min_{k\in K(j)} \delta_k^j, \qquad \beta := \max_{j\in[m]} \max_{k\in K(j)} \beta_k^j \qquad \text{and} \qquad R := \min_{j\in[m]} \min_{k\in K(j)} R_k^j,$$

and for the singular values

$$\Lambda_1 := \max_{j\in[m]} \Lambda_1^{\bar{k}(j)}, \quad \Lambda_2^- := \min_{j\in[m]} \Lambda_2^{\bar{k}(j)}, \quad \Lambda_2^+ := \max_{j\in[m]} \Lambda_2^{\bar{k}(j)}, \quad \Lambda_3 := \min_{j\in[m]} \Lambda_3^{\bar{k}(j)}.$$

We observe that these definitions allow for $\Lambda_1 > \Lambda_2^-$ and $\Lambda_2^+ > \Lambda_3$.

5.2 Numerical examples

5.2.1 Example 1

We first consider an example with only one nonlinear component function. Let

$$F : \mathbb{R}^3 \to \mathbb{R}^3, \qquad F(u) = \begin{pmatrix} u_1 + u_2 + u_3 \\ u_2 - 2(1+u_3)^2 + 2 \\ u_1 - 5u_3 \end{pmatrix}.$$

We fix $\alpha = 0.1$ and $\hat\alpha = 0$ in this example, so $B_0 = F'(u^0)$. Our focus is on the effect of the choice of (σ_k). We use $(\sigma_k) \equiv 1$ (Broyden's method), $(\sigma_k) \equiv 0.9$, $(\sigma_k) \equiv 1.1$, $(\sigma_k) \equiv 0.1$, $\sigma_k = 0.9$ for $0 \le k \le 4$ and $\sigma_k = 1$ else, $(\sigma_k) \equiv 1 - \frac{1}{(2+k)^2}$, $(\sigma_k) \equiv 1 - \frac{1}{(2+k)^4}$. The numerical outcome of a single run for Broyden's method is displayed in Table 1 and for $(\sigma_k) \equiv 0.9$ in Table 2. The results of cumulative runs are given in Table 3. From [22, Theorem 5 1)] we know that if $\sigma_k = 1$ for all k sufficiently large, then (u^k) converges with q-order $\frac{\sqrt{5}+1}{2} \approx 1.618$, which implies $\delta_k \approx 1.62$ for Broyden's method and also for the variant that uses $\sigma_k = 0.9$ only for the first five updates. [22, Theorem 5 1)] further asserts that (ϵ_k) converges faster than 2–step q-quadratically if σ_k is eventually 1. For the associated two choices of (σ_k) we thus compute β_k and R_k using $d = 1$ in (21), while for the other choices of (σ_k) we use $d = n = 3$. From [22, Lemma 3] we know that $\Lambda_1^k = \Lambda_2^k = 0$ for all $k \ge 0$ and that, up to its sign, $\hat{s}^k$ is constant for all $k \ge 1$, which implies that (9) and (10) are satisfied. For the same reason ζ_k, $\bar\zeta_k$, ρ_ζ^k and $\bar\rho_\zeta^k$ will be zero for $k \ge 1$ up to machine precision, which is why we suppress the latter two in the results for this example. The numerical outcomes of the single runs in Table 1 and Table 2 and of the cumulative runs in Table 3 confirm the theoretical results.

Table 1 Example 1: Results for one run with $B_0 = F'(u^0)$ and $(\sigma_k) \equiv 1$

k	$\|\|F_k\|\|$	$\|\|E_k\|\|$	δ_k	ϵ_k	ρ_ϵ^k	β_k	R_k	ζ_k	$\bar{\zeta}_k$	Λ_1^k	Λ_2^k
0	0.61	0.38	-1	-1	-1	-1	-1	-1	0.64	0.0	0.0
1	0.017	0.31	2.06	0.12	0.35	-1	-1	0.64	3.8e-862	6.4e-506	6.4e-506
2	7.1e-4	0.3	1.7	0.051	0.37	-1	-1	1.1e-1007	3.8e-862	3.7e-506	6.8e-506
3	2.4e-7	0.3	2.03	4.3e-4	0.14	0.028	3.71	1.4e-1005	3.8e-862	3.7e-506	8.1e-507
4	3.5e-12	0.3	1.71	1.8e-5	0.11	7.0e-3	3.67	3.3e-1004	3.8e-862	4.8e-506	3.7e-506
5	1.7e-20	0.3	1.71	6.2e-9	0.043	0.033	2.44	1.3e-1000	3.8e-862	4.3e-506	0.0
6	1.2e-33	0.3	1.66	8.8e-14	0.014	2.7e-4	2.75	4.6e-996	3.8e-862	5.2e-506	3.7e-506
7	4.0e-55	0.3	1.65	4.3e-22	2.1e-3	1.1e-5	2.6	2.4e-987	3.8e-862	4.7e-506	3.7e-506
8	9.3e-90	0.3	1.63	3.0e-35	1.5e-4	3.8e-9	2.64	2.2e-974	3.8e-862	2.3e-506	3.7e-506
9	7.4e-146	0.3	1.63	1.0e-56	2.5e-6	5.5e-14	2.62	8.4e-953	3.8e-862	1.9e-506	3.7e-506
10	1.4e-236	0.3	1.62	2.4e-91	5.8e-9	2.7e-22	2.62	8.4e-953	3.8e-862	3.7e-506	2.9e-506
11	2.1e-383	0.3	1.62	1.9e-147	5.9e-13	1.9e-35	2.62	3.8e-862	0.0	5.7e-506	3.7e-506

They indicate convergence of $\sum_k \epsilon_k$, of $\sum_k \zeta_k$ and of $\sum_k \bar{\zeta}_k$ in each of the 2000 runs due to $\rho_\epsilon, \rho_\zeta, \bar{\rho}_\zeta \ll 1$. Since the latter two values are provably zero in exact arithmetic, but are rather far away from this for $(\sigma_k) \equiv 0.1$, we will not use the choice $(\sigma_k) \equiv 0.1$ in subsequent examples. Due to $n = 3$ there holds $\|E_k\| = \Lambda_3^k$ for all $k \geq 0$, so we suppress Λ_3^k, respectively, Λ_3 in the tables. We notice that δ_k and δ are significantly smaller if $\sigma_k \neq 1$ for sufficiently large k. Also, 2/6–step convergence of (ϵ_k) with some q-order larger than one seems to hold. Table 3 may be interpreted as showing the existence of a lower bound for δ for all choices of (σ_k), but taking into account the development of (δ_k) for $(\sigma_k) \equiv 0.9$ in Table 2 we are hesitant. Rather, it seems clear that (ϵ_k) converges q-linearly with q-factor 0.1. However, the existence of a lower bound $\delta > 1$ implies convergence of (ϵ_k) with r-order δ, which is faster than r-superlinear convergence, so in particular ρ_ϵ would have to converge to zero in this case. For this reason we lean towards the conjecture that the existence of a lower bound $\delta > 1$ may hold if $\sigma_k \to 1$ (fast enough), but likely not if σ_k is bounded away from 1. Observing in Table 2 that β_k seems to grow by a factor of 10 in each iteration, we suspect similarly that $R_k \to 1$, so we deem the values for R in Table 3 untrustworthy and guess that $R > 1$ only holds if $\sigma_k \to 1$ (fast enough). Since the theory of [22] has shown that the example at hand behaves essentially like a one-dimensional example in which just the second equation appears, we conclude that this is the best possible situation and it is to be expected that δ and R also do not exist in other examples if $\sigma_k \not\to 1$.

5.2.2 Example 2

The next example contains two nonlinear equations:

$$F : \mathbb{R}^4 \to \mathbb{R}^4, \qquad F(u) = \begin{pmatrix} 25\sin(u_1) + 10\cos(u_2) + 10u_3^3 - 0.1u_4^2 - 10 \\ u_1 + u_3 \\ (1 + u_1)u_2(u_3 - 1) \\ u_3 - u_4 \end{pmatrix}.$$

We fix $\alpha = 0.1$ and $\hat{\alpha} = 0$ in this example and study again the influence of (σ_k). From [22] we obtain that $(s^k)_{k\geq 1}$ is confined to an affine space of dimension

Table 2 Example 1: Results for one run with $B_0 = F'(u^0)$ and $(\sigma_k) \equiv 0.9$

k	$\|\|F_k\|\|$	$\|\|E_k\|\|$	δ_k	ϵ_k	ρ_ϵ^k	β_k	R_k	ζ_k	$\bar{\zeta}_k$	Λ_1^k	Λ_2^k
0	0.53	0.33	-1	-1	-1	-1	-1	-1	1.2	0.0	0.0
1	0.013	0.2	1.77	0.14	0.37	-1	-1	1.2	4.4e-984	6.0e-506	0.0
2	2.1e-5	0.2	2.36	1.8e-3	0.12	-1	-1	4.5e-1006	4.4e-984	0.0	8.2e-506
3	2.2e-9	0.2	1.81	1.2e-4	0.1	-1	-1	2.8e-1005	4.4e-984	2.6e-506	6.4e-506
4	2.2e-14	0.2	1.56	1.2e-5	0.1	-1	-1	8.9e-1004	4.4e-984	6.9e-506	2.6e-506
5	2.2e-20	0.2	1.43	1.2e-6	0.1	-1	-1	8.7e-1003	4.4e-984	0.0	5.0e-506
6	2.2e-27	0.2	1.35	1.2e-7	0.1	-1	-1	1.5e-1002	4.4e-984	2.6e-506	6.6e-507
7	2.2e-35	0.2	1.3	1.2e-8	0.1	6.1e-7	9.22	3.9e-1001	4.4e-984	0.0	6.2e-506
8	2.3e-44	0.2	1.26	1.2e-9	0.1	3.5e-4	3.26	4.3e-1000	4.4e-984	3.4e-506	0.0
9	2.3e-54	0.2	1.23	1.2e-10	0.1	8.0e-3	2.53	7.1e-999	4.4e-984	2.6e-506	3.0e-506
10	2.3e-65	0.2	1.2	1.2e-11	0.1	0.087	2.21	6.5e-998	4.4e-984	5.0e-506	2.6e-506
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
20	2.4e-230	0.2	1.1	1.2e-21	0.1	8.7e8	1.4	6.2e-988	4.4e-984	7.0e-506	3.7e-506
21	2.5e-252	0.2	1.1	1.2e-22	0.1	8.7e9	1.38	9.9e-987	4.4e-984	8.1e-506	0.0
22	2.5e-275	0.2	1.09	1.2e-23	0.1	8.7e10	1.35	5.1e-986	4.4e-984	2.6e-506	1.7e-506
23	2.5e-299	0.2	1.09	1.2e-24	0.1	8.7e11	1.33	6.1e-985	4.9e-984	5.2e-506	0.0
24	2.5e-324	0.2	1.08	1.2e-25	0.1	8.7e12	1.32	4.9e-984	0.0	0.0	6.2e-506

Table 3 Example 1: Results for cumulative runs with $B_0 = F'(u^0)$ and varying (σ_k)

(σ_k)	$\|\|F\|\|$	$\|\|E\|\|$	δ	ρ_ϵ	β	R	ρ_ζ	$\bar{\rho}_\zeta$	Λ_1	Λ_2^-	Λ_2^+
1	8e-321	2e-4	1.62	3e-4	2e-7	2.59	4e-68	3e-74	1e-505	0	1e-505
0.9	1e-320	2e-5	1.08	0.13	8e13	1.3	4e-37	2e-38	1e-505	0	1e-505
1.1	1e-320	2e-4	1.08	0.13	7e13	1.3	4e-37	2e-38	2e-505	0	2e-505
0.1	1e-320	2e-4	1.02	0.86	4e6	1.02	3e-11	2e-11	7e-506	0	9e-506
5×0.9	8e-321	2e-4	1.57	9e-3	8e-6	2.25	5e-63	3e-68	2e-505	0	1e-505
$1-(k+2)^{-2}$	1e-320	2e-4	1.12	0.03	7e17	1.61	7e-47	4e-49	1e-505	0	1e-505
$1-(k+2)^{-4}$	9e-321	2e-4	1.17	4e-3	4	1.98	8e-57	2e-60	1e-505	0	1e-505

2 and that $(\Lambda_1^k) \equiv (\Lambda_2^k) \equiv 0$. Because of the restriction to two dimensions we conjecture 4–step q-quadratic convergence of (ϵ_k) if $\sigma_k = 1$ for all large k; accordingly, we compute β_k and R_k using $d = 2$ in (21) for such choices of (σ_k), while for all other choices we use $d = n = 4$. The results in Table 5 and Table 6 show that multi–step convergence of (ϵ_k) with q-order larger than one is only ensured if $\sigma_k \to 1$. We notice that (B_k) converges in all runs, for instance because the worst-case rates δ are always larger than one, cf. Theorem 3 with $m_k = k$. Furthermore, let us point out that for Broyden's method Corollary 4 asserts that for large k every set of the form $\{\delta_k, \delta_{k+1}, \delta_{k+2}, \delta_{k+3}\}$ contains at least one number close to or larger than 1.25; this behavior is apparent in Table 5. Again (10) is satisfied in all runs for this example.

5.2.3 Example 3

Next we choose a fully nonlinear F. Let

$$F : \mathbb{R}^3 \to \mathbb{R}^3, \qquad F(u) = \begin{pmatrix} (1+u_1)^2(1+u_2) + (1+u_2)^2 + u_3 - 2 \\ e^{u_1} + (1+u_2)^3 + u_3^2 - 2 \\ e^{u_3^2} + (1+u_2)^2 - 2 \end{pmatrix}.$$

We fix $\hat{\alpha} = 0$ and investigate various choices for (σ_k), each one for $\alpha \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-8}\}$. The results are displayed in Tables 7–9. The convergence of (B_k) in all runs and the validity of (10) are obvious based on the

Table 4 Example 2: Results for one run with $B_0 = F'(u^0)$ and $(\sigma_k) \equiv 1$

k	$\|\|F_k\|\|$	$\|\|E_k\|\|$	δ_k	ϵ_k	ρ_ϵ^k	β_k	R_k	ζ_k	ρ_ζ^k	$\bar{\zeta}_k$	$\bar{\rho}_\zeta^k$	Λ_1^k	Λ_2^k	Λ_3^k
0	1.5	0.85	-1	-1	-1	-1	-1	-1	-1	1.1	1.1	1.6e-506	2.0e-505	0.094
1	0.037	0.72	1.72	0.25	0.5	-1	-1	1.1	1.0	0.06	0.24	9.2e-506	1.2e-505	0.079
2	6.1e-3	0.26	1.1	0.62	0.85	-1	-1	0.22	0.6	0.28	0.65	4.7e-506	9.7e-506	0.065
3	2.7e-5	0.26	1.53	0.026	0.4	-1	-1	0.45	0.82	0.17	0.64	9.4e-506	4.0e-506	0.059
4	2.6e-6	0.24	1.21	0.1	0.64	-1	-1	0.18	0.71	8.7e-3	0.39	4.7e-507	6.1e-506	0.057
5	4.0e-8	0.24	1.24	0.037	0.58	0.58	2.4	8.2e-3	0.45	5.4e-4	0.29	1.6e-505	2.9e-506	0.057
6	2.9e-11	0.24	1.36	1.7e-3	0.4	4.4e-3	13.3	1.1e-3	0.38	5.2e-4	0.34	1.5e-505	3.9e-506	0.057
7	2.8e-15	0.24	1.34	2.2e-4	0.35	0.32	2.31	5.9e-4	0.39	6.6e-5	0.3	5.2e-506	5.1e-506	0.057
8	1.5e-19	0.24	1.26	1.2e-4	0.37	0.011	3.99	6.6e-5	0.34	8.7e-8	0.16	1.5e-505	5.5e-506	0.057
9	8.8e-25	0.24	1.25	1.3e-5	0.33	9.8e-3	3.4	8.7e-8	0.2	7.5e-11	0.097	1.6e-506	4.5e-506	0.057
10	6.9e-33	0.24	1.32	1.8e-8	0.2	6.5e-3	2.79	7.2e-11	0.12	3.0e-12	0.09	1.2e-505	2.4e-506	0.057
11	4.5e-44	0.24	1.33	1.5e-11	0.13	3.1e-4	2.96	3.0e-12	0.11	2.2e-14	0.073	9.3e-507	1.3e-505	0.057
12	1.2e-56	0.24	1.28	6.2e-13	0.12	4.2e-5	3.12	2.2e-14	0.089	4.0e-21	0.027	4.5e-506	6.1e-506	0.057
13	2.4e-71	0.24	1.26	4.4e-15	0.094	2.4e-5	2.95	4.0e-21	0.035	3.6e-30	7.9e-3	1.0e-505	1.7e-506	0.057
14	8.5e-93	0.24	1.3	8.2e-22	0.039	2.6e-6	2.72	3.6e-30	0.011	5.3e-36	4.4e-3	1.9e-505	2.0e-506	0.057
15	2.8e-123	0.24	1.33	7.5e-31	0.013	3.4e-9	2.78	5.3e-36	6.2e-3	1.1e-41	2.8e-3	1.2e-505	3.5e-506	0.057
16	1.3e-159	0.24	1.29	1.1e-36	7.7e-3	2.8e-12	2.95	1.1e-41	3.9e-3	2.8e-57	4.7e-4	3.7e-506	1.5e-505	0.057
17	1.4e-201	0.24	1.26	2.3e-42	4.9e-3	1.2e-13	2.9	2.8e-57	7.2e-4	4.3e-82	3.0e-5	1.3e-505	4.8e-506	0.057
18	3.4e-259	0.24	1.28	5.8e-58	9.7e-4	8.5e-16	2.71	4.3e-82	5.2e-5	8.2e-103	4.2e-6	2.2e-506	1.2e-505	0.057
19	1.3e-341	0.24	1.32	8.8e-83	7.9e-5	1.6e-22	2.72	8.2e-103	7.9e-6	0.0	0.0	8.0e-506	4.4e-506	0.057

Table 5 Example 2: Results for one run with $B_0 = F'(u^0)$ and $(\sigma_k) \equiv 0.9$

k	$\|\|F_k\|\|$	$\|\|E_k\|\|$	δ_k	ϵ_k	ρ_ϵ^k	β_k	R_k	ζ_k	ρ_ζ^k	$\bar{\zeta}_k$	$\bar{\rho}_\zeta^k$	Λ_1^k	Λ_2^k	Λ_3^k
0	1.2	0.49	-1	-1	-1	-1	-1	-1	-1	0.93	0.93	4.7e-506i	4.3e-507i	0.068
1	0.012	0.42	1.74	0.14	0.37	-1	-1	0.97	0.98	0.14	0.38	5.5e-506	5.0e-506	0.063
2	9.1e-4	0.17	1.16	0.34	0.7	-1	-1	0.68	0.88	0.81	0.93	6.2e-506	5.0e-506	0.051
3	4.7e-6	0.16	1.29	0.057	0.49	-1	-1	0.9	0.97	0.097	0.56	2.3e-506	8.8e-506	0.044
4	3.2e-7	0.12	1.17	0.1	0.63	-1	-1	0.11	0.65	0.017	0.44	2.6e-506	3.2e-506	0.04
5	9.8e-11	0.12	1.44	8.0e-4	0.3	-1	-1	1.4	1.1	1.4	1.1	4.7e-506	1.3e-506	0.04
6	6.4e-13	0.085	1.09	0.081	0.7	-1	-1	1.4	1.0	1.4e-4	0.28	7.8e-507	4.7e-507	0.024
7	1.2e-15	0.085	1.18	4.6e-3	0.51	-1	-1	3.3e-3	0.49	3.2e-3	0.49	1.6e-507	4.2e-506	0.024
8	2.1e-19	0.085	1.22	4.3e-4	0.42	-1	-1	2.1e-4	0.39	3.4e-3	0.53	6.7e-507	4.8e-506	0.024
9	3.6e-24	0.085	1.23	4.1e-5	0.36	2.1e-3	5.1	1.5e-4	0.41	3.6e-3	0.57	4.9e-507	3.7e-506	0.024
10	4.2e-30	0.085	1.23	2.8e-6	0.31	2.5e-5	11.8	1.6e-3	0.56	5.1e-3	0.62	7.9e-507	3.4e-506	0.024
11	2.3e-35	0.085	1.16	1.3e-5	0.39	4.1e-3	3.92	5.2e-3	0.65	1.1e-4	0.47	7.9e-506	6.0e-507	0.024
12	4.0e-40	0.085	1.12	4.3e-5	0.46	4.3e-3	4.38	1.2e-4	0.5	3.4e-6	0.38	2.1e-506	8.5e-507	0.024
13	5.5e-46	0.085	1.14	3.3e-6	0.41	5.2	1.77	1.0e-6	0.37	4.4e-6	0.41	5.5e-507	7.0e-506	0.024
14	7.3e-53	0.085	1.14	3.3e-7	0.37	4.9e-5	5.96	1.2e-7	0.35	4.6e-6	0.44	6.4e-507	6.0e-506	0.024
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
29	8.7e-216	0.085	1.07	2.0e-15	0.32	5.7e5	1.44	3.8e-15	0.33	7.4e-13	0.39	4.1e-506	2.2e-507	0.024
30	6.0e-232	0.085	1.07	1.7e-16	0.31	4.8e6	1.4	1.4e-13	0.38	8.8e-13	0.41	8.8e-507	5.3e-507	0.024
31	2.9e-247	0.085	1.06	1.2e-15	0.34	4.4e9	1.21	8.9e-13	0.42	1.3e-14	0.37	4.5e-506	2.0e-507	0.024
32	8.9e-262	0.085	1.06	7.5e-15	0.37	8.2e8	1.23	1.3e-14	0.38	3.6e-17	0.32	4.5e-507	9.3e-507	0.024
33	2.3e-277	0.085	1.06	6.4e-16	0.36	1.1e6	1.43	3.4e-17	0.33	2.4e-18	0.3	4.7e-507	3.6e-506	0.024
34	6.1e-294	0.085	1.06	6.4e-17	0.34	1.5e7	1.39	1.0e-18	0.31	1.3e-18	0.31	4.8e-506	4.4e-507	0.024
35	1.6e-311	0.085	1.06	6.4e-18	0.33	1.6e8	1.35	3.2e-19	0.31	1.0e-18	0.32	1.2e-506	4.9e-506	0.024
36	4.1e-330	0.085	1.06	6.3e-19	0.32	1.6e9	1.33	1.0e-18	0.33	0.0	0.0	9.1e-506	7.8e-507	0.024

Table 6 Example 2: Results for cumulative runs with $B_0 = F'(u^0)$ and varying (σ_k)

(σ_k)	$\|\|F\|\|$	$\|\|E\|\|$	δ	ρ_ϵ	β	R	ρ_ζ	$\bar{\rho}_\zeta$	Λ_1	Λ_2^-	Λ_2^+	Λ_3
1	9e-321	3e-3	1.2	0.08	3e-3	2.27	0.04	0.03	3e-505	0	2e-505	2e-4
0.9	1e-320	3e-3	1.04	0.5	2e23	0.68	0.56	0.55	2e-505	0	3e-505	3e-4
1.1	9e-321	3e-3	1.04	0.57	2e25	0.69	0.67	0.55	3e-505	0	3e-505	3e-4
5×0.9	1e-320	3e-3	1.19	0.1	0.05	2.09	0.07	0.03	3e-505	0	3e-505	3e-4
$1-(k+2)^{-2}$	1e-320	3e-3	1.07	0.23	1e16	1.19	0.26	0.21	2e-505	0	3e-505	3e-4
$1-(k+2)^{-4}$	1e-320	3e-3	1.11	0.09	2e14	1.4	0.11	0.10	3e-505	0	3e-505	1e-4

indicators δ, ρ_ϵ, ρ_ζ and $\bar{\rho}_\zeta$ if $(\sigma_k) \equiv 1$. For $(\sigma_k) \equiv 0.1$ we find, judging from ρ_ϵ and δ, that (B_k) converges for $\alpha \leq 10^{-3}$. The values of ρ_ϵ may be interpreted as showing convergence of (B_k) also for $\alpha \in \{10^{-1}, 10^{-2}\}$, in particular taking into account that the convergence for $(\sigma_k) \equiv 1$ appears to be rather slow in comparsion to other choices of (σ_k), which follows from the smaller values of δ and also from the comparably large value of Λ_1. The conjectured 6–step

Table 7 Example 3: Results for one run with $B_0 = F'(u^0)$ and $(\sigma_k) \equiv 1$ for $\alpha = 10^{-1}$

k	$\|\|F_k\|\|$	$\|\|E_k\|\|$	δ_k	ϵ_k	ρ_ϵ^k	β_k	R_k	ζ_k	ρ_ζ^k	$\bar{\zeta}_k$	$\bar{\rho}_\zeta^k$	Λ_1^k	Λ_2^k
0	0.26	0.26	-1	-1	-1	-1	-1	-1	-1	1.2	1.2	0.051	0.18
1	0.012	0.23	2.01	0.11	0.33	-1	-1	1.3	1.1	0.57	0.75	0.042	0.12
2	9.9e-4	0.18	1.31	0.19	0.58	-1	-1	0.84	0.94	0.92	0.97	3.5e-3	0.044
3	7.0e-5	0.16	1.35	0.084	0.54	-1	-1	1.1	1.0	1.1	1.0	8.3e-4	0.034
4	1.6e-6	0.15	1.27	0.057	0.56	-1	-1	0.58	0.9	0.54	0.88	4.7e-6	0.013
5	3.0e-8	0.15	1.27	0.025	0.54	-1	-1	0.51	0.89	0.036	0.57	5.7e-7	0.011
6	1.6e-9	0.15	1.21	0.031	0.61	-1	-1	0.048	0.65	0.012	0.53	6.2e-8	9.5e-3
7	5.6e-12	0.14	1.31	2.2e-3	0.47	0.2	2.72	0.012	0.58	3.7e-4	0.37	2.8e-9	9.4e-3
8	5.2e-15	0.14	1.29	5.8e-4	0.44	0.015	4.53	4.1e-4	0.42	3.7e-5	0.32	1.0e-11	9.4e-3
9	1.7e-19	0.14	1.33	2.1e-5	0.34	3.0e-3	4.34	5.3e-3	0.59	5.4e-3	0.59	9.3e-15	9.4e-3
10	6.9e-23	0.14	1.19	2.5e-4	0.47	0.077	2.9	5.4e-3	0.62	7.4e-5	0.42	3.0e-19	9.4e-3
11	2.9e-26	0.14	1.16	2.6e-4	0.5	0.43	2.23	6.7e-5	0.45	6.7e-6	0.37	1.2e-22	9.4e-3
12	1.5e-31	0.14	1.22	3.2e-6	0.38	3.3e-3	3.64	6.6e-6	0.4	4.0e-8	0.27	5.1e-26	9.4e-3
13	7.4e-38	0.14	1.21	3.1e-7	0.34	0.062	2.46	4.0e-8	0.3	1.3e-11	0.17	2.6e-31	9.4e-3
14	2.3e-46	0.14	1.24	1.9e-9	0.26	5.7e-3	2.69	8.4e-11	0.21	7.0e-11	0.21	1.3e-37	9.4e-3
15	1.4e-57	0.14	1.25	4.0e-12	0.19	9.3e-3	2.43	7.0e-11	0.23	1.4e-10	0.24	4.0e-46	9.4e-3
16	7.7e-69	0.14	1.2	3.3e-12	0.21	5.3e-5	3.19	1.4e-10	0.26	1.4e-13	0.18	2.6e-57	9.4e-3
17	8.3e-80	0.14	1.16	6.7e-12	0.24	1.0e-4	3.11	1.4e-13	0.19	1.7e-17	0.12	1.4e-68	9.4e-3
18	8.9e-94	0.14	1.18	6.7e-15	0.18	6.6e-4	2.58	1.7e-17	0.13	6.2e-23	0.068	1.5e-79	9.4e-3
19	1.2e-111	0.14	1.19	8.1e-19	0.12	8.2e-6	2.78	6.2e-23	0.078	1.6e-30	0.032	1.6e-93	9.4e-3
20	5.5e-135	0.14	1.21	3.0e-24	0.076	8.1e-7	2.7	1.7e-30	0.038	2.4e-33	0.028	2.1e-111	9.4e-3
21	6.9e-166	0.14	1.23	7.8e-32	0.039	4.9e-9	2.73	5.7e-33	0.034	8.1e-33	0.035	9.8e-135	9.4e-3
22	3.0e-199	0.14	1.2	2.7e-34	0.035	2.4e-11	2.93	8.1e-33	0.04	6.9e-39	0.022	1.2e-165	9.4e-3
23	1.9e-232	0.14	1.17	3.8e-34	0.041	8.6e-12	2.99	6.9e-39	0.026	2.4e-49	9.4e-3	5.3e-199	9.4e-3
24	9.8e-272	0.14	1.17	3.3e-40	0.026	7.3e-12	2.79	2.4e-49	0.011	3.9e-64	2.9e-3	3.3e-232	9.4e-3
25	1.8e-321	0.14	1.18	1.1e-50	0.012	1.7e-14	2.76	3.9e-64	3.6e-3	0.0	0.0	1.7e-271	9.4e-3

Table 8 Example 3: Results for one run with $B_0 = F'(u^0)$ and $(\sigma_k) \equiv 1$ for $\alpha = 10^{-3}$

k	$\|\|F_k\|\|$	$\|\|E_k\|\|$	δ_k	ϵ_k	ρ_ϵ^k	β_k	R_k	ζ_k	ρ_ζ^k	$\bar{\zeta}_k$	$\bar{\rho}_\zeta^k$	Λ_1^k	Λ_2^k
0	1.2e-3	3.4e-3	-1	-1	-1	-1	-1	-1	-1	1.0	1.0	5.8e-5	8.1e-4
1	1.6e-6	2.1e-3	1.95	1.5e-3	0.039	-1	-1	0.28	0.53	0.83	0.91	4.9e-5	7.7e-4
2	2.2e-9	1.4e-3	1.47	1.7e-3	0.12	-1	-1	0.12	0.5	0.95	0.98	2.6e-7	4.4e-4
3	1.3e-13	1.4e-3	1.47	8.0e-5	0.095	-1	-1	0.94	0.98	8.0e-3	0.3	2.9e-10	4.4e-4
4	2.1e-16	1.3e-3	1.26	5.5e-4	0.22	-1	-1	0.011	0.4	0.019	0.45	1.2e-13	2.6e-4
5	2.4e-21	1.3e-3	1.36	3.9e-6	0.13	-1	-1	3.4e-3	0.39	0.015	0.5	1.9e-16	2.6e-4
6	8.6e-27	1.3e-3	1.29	1.2e-6	0.14	-1	-1	0.015	0.55	1.4e-5	0.2	2.2e-21	2.6e-4
7	1.4e-31	1.3e-3	1.21	5.2e-6	0.22	2.2	1.88	1.4e-5	0.25	1.0e-7	0.13	8.0e-27	2.6e-4
8	1.9e-39	1.3e-3	1.27	4.7e-9	0.12	1.6e-3	3.01	5.2e-7	0.2	6.2e-7	0.2	1.3e-31	2.6e-4
9	1.0e-48	1.3e-3	1.26	1.8e-10	0.11	0.027	2.38	6.2e-7	0.24	4.4e-11	0.092	1.8e-39	2.6e-4
10	6.6e-58	1.3e-3	1.2	2.1e-10	0.13	7.1e-4	2.97	4.3e-11	0.11	1.4e-13	0.068	9.6e-49	2.6e-4
11	3.0e-71	1.3e-3	1.24	1.5e-14	0.07	9.8e-4	2.56	1.0e-12	0.1	8.6e-13	0.099	6.1e-58	2.6e-4
12	3.1e-86	1.3e-3	1.22	3.4e-16	0.065	2.5e-4	2.61	8.6e-13	0.12	4.4e-19	0.039	2.7e-71	2.6e-4
13	2.7e-101	1.3e-3	1.18	2.9e-16	0.078	1.1e-5	2.94	4.4e-19	0.049	3.9e-24	0.021	2.9e-86	2.6e-4
14	1.2e-122	1.3e-3	1.22	1.5e-22	0.035	6.8e-6	2.62	4.1e-24	0.028	2.1e-25	0.023	2.5e-101	2.6e-4
15	5.3e-149	1.3e-3	1.22	1.4e-27	0.021	4.5e-8	2.75	2.1e-25	0.029	1.2e-33	8.8e-3	1.2e-122	2.6e-4
16	1.2e-176	1.3e-3	1.19	7.2e-29	0.022	1.6e-9	2.91	1.2e-33	0.012	4.7e-41	4.2e-3	4.9e-149	2.6e-4
17	1.4e-212	1.3e-3	1.21	4.1e-37	9.5e-3	1.8e-9	2.63	4.6e-41	5.7e-3	8.0e-43	4.6e-3	1.1e-176	2.6e-4
18	6.8e-256	1.3e-3	1.21	1.6e-44	5.0e-3	1.3e-13	2.83	8.0e-43	6.1e-3	1.8e-55	1.3e-3	1.3e-212	2.6e-4
19	5.6e-301	1.3e-3	1.18	2.7e-46	5.3e-3	3.1e-15	2.93	1.8e-55	1.8e-3	8.0e-72	2.8e-4	6.3e-256	2.6e-4
20	1.0e-358	1.3e-3	1.19	6.2e-59	1.7e-3	2.7e-15	2.67	8.0e-72	4.1e-4	0.0	0.0	5.2e-301	2.6e-4

convergence of (ϵ_k) with q-order larger than one seems to hold for $\alpha \leq 10^{-3}$ if $(\sigma_k) \equiv 1$, but not for $\alpha \in \{10^{-1}, 10^{-2}\}$ and also not for other choices of (σ_k). A closer inspection of $\alpha \in \{10^{-1}, 10^{-2}\}$ for Broyden's method reveals that out of the 2000 runs only 2, respectively, 1 fail to exhibit $R \geq 2$. By Corollary 4 we should see in Broyden's method for large k that every set $\{\delta_k, \delta_{k+1}, \ldots, \delta_{k+5}\}$ contains at least one number close to or larger than $7/6 \approx 1.17$; Table 7 and 8 confirm this. Moreover, since $\delta \geq 1.12$ for $(\sigma_k) \equiv 1$ in Table 9 we conclude that (δ_k) always stays safely away from 1, which implies convergence of (B_k) via Theorem 3.

Table 9 Example 3: Results for cumulative runs for $B_0 = F'(u^0)$, $\alpha = 10^{-j}$ and $(\sigma_k) \equiv \sigma$, respectively, $\sigma_k = 0.9$ for $k \leq 9$ and $\sigma_k = 1$ else (represented by $\sigma = X$), and $\sigma_k = 1 - (k+2)^{-2}$ (represented by $\sigma = Y$)

(j,σ)	$\|\|F\|\|$	$\|\|E\|\|$	δ	ρ_ϵ	β	R	ρ_ζ	$\bar\rho_\zeta$	Λ_1	Λ_2^-	Λ_2^+
$(1,1)$	8e-321	1e-2	1.12	0.3	982	1.72	0.27	0.19	1e-257	1e-7	0.27
$(2,1)$	1e-320	1e-3	1.13	0.15	644	1.74	0.14	0.11	1e-259	2e-7	0.02
$(3,1)$	9e-321	7e-5	1.14	0.06	8e-3	2.17	0.08	0.06	2e-257	2e-8	2e-3
$(8,1)$	9e-321	6e-10	1.13	0.02	3e-5	2.31	6e-3	0.03	5e-251	4e-14	2e-8
$(1,X)$	9e-321	8e-3	1.09	0.42	4e5	1.55	0.39	0.29	8e-262	9e-8	0.26
$(2,X)$	9e-321	8e-4	1.08	0.42	3e11	1.29	0.41	0.34	1e-263	1e-8	2e-2
$(3,X)$	9e-321	8e-5	1.09	0.26	1e18	1.46	0.4	0.32	6e-241	1e-8	2e-3
$(1,Y)$	1e-320	1e-2	1.06	0.38	2e27	0.82	0.45	0.44	2e-71	2e-7	0.25
$(2,Y)$	1e-320	9e-4	1.06	0.37	1e32	0.71	0.51	0.42	2e-68	4e-8	2e-2
$(3,Y)$	1e-320	9e-5	1.05	0.33	7e27	0.77	0.49	0.48	6e-66	6e-9	2e-3
$(1,0.9)$	1e-320	8e-3	1.03	0.61	1e23	0.65	0.76	0.73	5e-38	3e-9	0.26
$(2,0.9)$	1e-320	8e-4	1.03	0.63	2e20	0.66	0.8	0.79	2e-37	3e-8	2e-2
$(3,0.9)$	1e-320	8e-5	1.03	0.61	2e22	0.63	0.82	0.82	1e-34	2e-10	2e-3
$(1,1.1)$	1e-320	5e-3	1.03	0.66	1e22	0.58	0.75	0.75	1e-39	4e-7	0.22
$(2,1.1)$	1e-320	1e-3	1.03	0.63	1e23	0.64	0.82	0.81	2e-37	2e-8	2e-2
$(3,1.1)$	1e-320	9e-5	1.03	0.59	3e22	0.61	0.89	0.88	1e-35	3e-9	2e-3
$(1,0.5)$	1e-320	1e-2	1.01	0.91	2e18	0.35	1.0	1.01	9e-18	2e-7	0.2
$(2,0.5)$	1e-320	1e-3	1.01	0.88	5e18	0.37	1.01	1.0	2e-16	3e-8	2e-2
$(3,0.5)$	1e-320	1e-4	1.02	0.84	1e16	0.49	1.01	1.01	3e-15	2e-9	2e-3

5.2.4 Example 4

We consider another mapping without affine components:

$$F:\mathbb{R}^7 \to \mathbb{R}^7, \qquad F(u) = \begin{pmatrix} u_1 + u_2 + (1+u_3)^2 + u_4 + u_5 + u_6 - u_7^3 \\ u_2 - 2(1+u_3)^2 + 3u_5 - \sin(u_7) + 2 \\ u_1 - u_3^2 + u_5 u_6 u_7 \\ 0.5\ln(1+u_2^2) - 2e^{u_3} + 0.1u_7^{10} + 2 \\ \sin(u_1 + u_3 - 10u_2) - u_4^5 - u_6 \\ u_1^2 + u_3^2 + u_5^2 + (1+u_7)^2 - 1 \\ u_6 - u_7 - u_7^6 \end{pmatrix}.$$

We use $\alpha \in \{10^{-2}, 10^{-4}\}$, $\hat\alpha \in \{0, 10^{-2}, 10^{-4}, 10^{-10}\}$ and $(\sigma_k) \equiv 1$ as well as $(\sigma_k) \equiv 0.9$. The results in Table 10–12 show convergence of (B_k) in all runs, since $\delta > 1$ and since ρ_ϵ is safely smaller than one. In contrast to previous examples the indicators R, ρ_ζ and $\bar\rho_\zeta$ are inconclusive in the case $(\sigma_k) \equiv 0.9$ and it is unclear if the summability property (9) holds. We mention that for $\alpha = 10^{-5}$ and $\alpha = 10^{-6}$ there were only two runs out of 4000 with $R < 2$. We confirm for $(\sigma_k) \equiv 1$ in Table 10 that every $2n = 14$ steps there is at least one for which δ_k is close to or larger than $15/14 \approx 1.07$ (in fact, all are) and infer from $\delta \geq 1.05$ that (B_k) converges, cf. Theorem 3.

Table 10 Example 4: Results for one run with $B_0 = F'(u^0)$, $\alpha = 0.01$ and $(\sigma_k) \equiv 1$

k	$\|\|F_k\|\|$	$\|\|E_k\|\|$	δ_k	ϵ_k	ρ_ϵ^k	β_k	R_k	ζ_k	ρ_ζ^k	$\bar{\zeta}_k$	$\bar{\rho}_\zeta^k$	Λ_1^k	Λ_2^k	Λ_3^k
0	0.092	0.043	-1	-1	-1	-1	-1	-1	-1	1.0	1.0	5.0e-507	9.7e-6	2.7e-5
1	2.6e-4	0.039	2.05	0.015	0.12	-1	-1	0.89	0.94	0.83	0.91	3.7e-13	7.1e-6	2.7e-5
2	1.0e-6	0.039	1.52	8.8e-3	0.21	-1	-1	0.7	0.89	0.65	0.87	1.6e-14	3.3e-6	2.6e-5
3	9.1e-9	0.038	1.31	0.013	0.34	-1	-1	1.0	1.0	1.0	1.0	1.6e-16	3.2e-6	2.5e-5
4	5.2e-11	0.037	1.25	9.4e-3	0.39	-1	-1	0.25	0.75	0.87	0.97	2.3e-19	1.8e-6	2.5e-5
5	5.5e-14	0.037	1.26	1.7e-3	0.34	-1	-1	0.38	0.85	1.2	1.0	3.3e-21	1.8e-6	2.5e-5
6	1.2e-16	0.037	1.17	4.3e-3	0.46	-1	-1	0.85	0.98	0.97	0.99	6.1e-25	1.7e-6	2.5e-5
7	3.6e-19	0.036	1.11	0.015	0.59	-1	-1	0.97	1.0	1.4e-3	0.44	4.9e-28	1.2e-6	2.5e-5
8	7.7e-22	0.036	1.1	0.01	0.6	-1	-1	0.032	0.68	0.031	0.68	1.1e-29	7.1e-7	2.5e-5
9	3.3e-26	0.036	1.17	2.1e-4	0.43	-1	-1	5.8e-3	0.6	0.025	0.69	2.3e-32	7.1e-7	2.5e-5
10	2.6e-31	0.036	1.17	3.8e-5	0.4	-1	-1	1.4e-3	0.55	0.024	0.71	1.0e-36	7.1e-7	2.5e-5
11	4.7e-37	0.036	1.16	9.1e-6	0.38	-1	-1	7.6e-3	0.67	0.016	0.71	7.9e-42	7.1e-7	2.5e-5
12	5.0e-42	0.036	1.12	5.1e-5	0.47	-1	-1	0.016	0.73	5.5e-5	0.47	1.5e-47	7.1e-7	2.5e-5
13	1.1e-46	0.036	1.09	1.1e-4	0.52	-1	-1	2.2e-4	0.55	1.6e-4	0.54	1.6e-52	7.1e-7	2.5e-5
14	3.3e-53	0.036	1.13	1.5e-6	0.41	-1	-1	1.7e-5	0.48	1.5e-4	0.56	3.5e-57	7.1e-7	2.5e-5
15	7.9e-61	0.036	1.13	1.2e-7	0.37	5.4e-4	3.79	1.3e-4	0.57	2.0e-5	0.51	1.0e-63	7.1e-7	2.5e-5
16	1.4e-67	0.036	1.1	8.5e-7	0.44	0.011	2.96	2.0e-5	0.53	9.2e-7	0.44	2.5e-71	7.1e-7	2.5e-5
17	3.9e-75	0.036	1.1	1.4e-7	0.42	8.2e-4	3.64	7.9e-6	0.52	7.0e-6	0.52	4.3e-78	7.1e-7	2.5e-5
18	4.2e-83	0.036	1.1	5.3e-8	0.41	6.0e-4	3.59	6.7e-6	0.53	3.1e-7	0.45	1.2e-85	7.1e-7	2.5e-5
19	3.9e-91	0.036	1.09	4.5e-8	0.43	0.016	2.65	3.1e-7	0.47	3.4e-10	0.34	1.3e-93	7.1e-7	2.5e-5
20	1.7e-100	0.036	1.1	2.1e-9	0.39	1.1e-4	3.67	2.6e-9	0.39	2.9e-9	0.39	1.2e-101	7.1e-7	2.5e-5
21	6.0e-112	0.036	1.11	1.7e-11	0.32	7.3e-8	5.93	6.4e-11	0.34	2.8e-9	0.41	5.3e-111	7.1e-7	2.5e-5
22	5.2e-125	0.036	1.11	4.3e-13	0.29	4.0e-9	6.23	2.8e-9	0.42	8.8e-11	0.37	1.9e-122	7.1e-7	2.5e-5
23	2.0e-136	0.036	1.09	1.8e-11	0.36	4.1e-4	2.92	8.8e-11	0.38	1.4e-14	0.26	1.6e-135	7.1e-7	2.5e-5
24	2.4e-149	0.036	1.09	5.9e-13	0.32	4.0e-4	2.77	1.7e-13	0.31	1.9e-13	0.31	6.2e-147	7.1e-7	2.5e-5
25	5.7e-165	0.036	1.1	1.2e-15	0.27	1.4e-5	2.96	6.3e-13	0.34	4.4e-13	0.33	7.5e-160	7.1e-7	2.5e-5
26	5.0e-180	0.036	1.09	4.2e-15	0.29	1.6e-6	3.35	4.4e-13	0.35	3.5e-16	0.27	1.8e-175	7.1e-7	2.5e-5
27	3.0e-195	0.036	1.08	3.0e-15	0.3	2.5e-7	3.66	3.5e-16	0.28	1.3e-19	0.21	1.6e-190	7.1e-7	2.5e-5
28	1.4e-213	0.036	1.09	2.3e-18	0.25	1.1e-6	3.02	6.5e-19	0.24	5.3e-19	0.23	9.4e-206	7.1e-7	2.5e-5
29	1.3e-234	0.036	1.1	4.4e-21	0.21	3.3e-7	2.94	5.1e-19	0.25	2.1e-20	0.22	4.5e-224	7.1e-7	2.5e-5
30	9.0e-256	0.036	1.09	3.4e-21	0.22	4.7e-9	3.37	1.9e-20	0.23	2.6e-21	0.22	4.0e-245	7.1e-7	2.5e-5
31	2.3e-278	0.036	1.09	1.2e-22	0.21	6.6e-9	3.19	3.2e-21	0.23	6.9e-22	0.22	2.8e-266	7.1e-7	2.5e-5
32	1.0e-301	0.036	1.08	2.2e-23	0.21	7.7e-9	3.12	6.9e-22	0.23	9.8e-25	0.19	7.2e-289	7.1e-7	2.5e-5
33	9.9e-326	0.036	1.08	4.7e-24	0.21	2.3e-9	3.18	9.8e-25	0.2	0.0	0.0	3.2e-312	7.1e-7	2.5e-5

Table 11 Example 4: Results for one run with $B_0 = F'(u^0)$, $\alpha = 0.01$ and $(\sigma_k) \equiv 0.9$

k	$\|\|F_k\|\|$	$\|\|E_k\|\|$	δ_k	ϵ_k	ρ_ϵ^k	β_k	R_k	ζ_k	ρ_ζ^k	$\bar{\zeta}_k$	$\bar{\rho}_\zeta^k$	Λ_1^k	Λ_2^k	Λ_3^k
0	0.092	0.043	-1	-1	-1	-1	-1	-1	-1	1.1	1.1	5.0e-507	9.7e-6	2.7e-5
1	2.6e-4	0.039	2.05	0.013	0.12	-1	-1	0.89	0.94	1.2	1.1	3.1e-13	7.2e-6	2.7e-5
2	1.1e-6	0.039	1.51	8.4e-3	0.2	-1	-1	0.66	0.87	0.86	0.95	1.3e-13	2.7e-6	2.5e-5
3	8.7e-9	0.038	1.31	0.011	0.32	-1	-1	1.1	1.0	1.4	1.1	1.5e-14	2.7e-6	2.5e-5
4	5.0e-11	0.037	1.24	8.5e-3	0.39	-1	-1	0.36	0.81	1.2	1.0	3.2e-15	1.6e-6	2.5e-5
5	6.2e-14	0.037	1.26	1.8e-3	0.35	-1	-1	0.53	0.9	1.2	1.0	3.3e-16	1.5e-6	2.5e-5
6	1.5e-16	0.037	1.17	4.5e-3	0.46	-1	-1	0.29	0.84	1.0	1.0	3.8e-17	1.4e-6	2.5e-5
7	3.2e-19	0.037	1.13	6.6e-3	0.53	-1	-1	1.4	1.0	0.6	0.94	4.0e-18	1.4e-6	2.5e-5
8	1.3e-21	0.036	1.09	0.016	0.63	-1	-1	0.22	0.85	0.38	0.9	2.6e-18	2.5e-7	2.5e-5
9	5.5e-25	0.036	1.12	2.0e-3	0.54	-1	-1	5.6e-3	0.6	0.39	0.91	4.3e-19	1.6e-7	2.5e-5
10	2.3e-29	0.036	1.15	1.9e-4	0.46	-1	-1	0.024	0.71	0.41	0.92	4.3e-20	1.5e-7	2.5e-5
11	6.8e-35	0.036	1.16	1.3e-5	0.39	-1	-1	0.44	0.93	0.026	0.74	4.3e-21	1.5e-7	2.5e-5
12	7.5e-39	0.036	1.09	5.5e-4	0.56	-1	-1	0.16	0.87	0.14	0.86	4.8e-22	1.4e-7	2.5e-5
13	2.2e-43	0.036	1.1	1.4e-4	0.53	-1	-1	0.076	0.83	0.21	0.9	4.8e-23	1.4e-7	2.5e-5
14	4.9e-48	0.036	1.09	1.1e-4	0.54	-1	-1	0.062	0.83	0.28	0.92	4.8e-24	1.4e-7	2.5e-5
15	7.4e-53	0.036	1.09	7.2e-5	0.55	0.41	2.21	0.46	0.95	0.19	0.9	4.8e-25	1.4e-7	2.5e-5
16	7.8e-57	0.036	1.06	5.2e-4	0.64	7.3	1.58	0.35	0.94	0.17	0.9	5.4e-26	1.2e-7	2.5e-5
17	5.0e-61	0.036	1.06	3.1e-4	0.64	2.7	1.78	0.015	0.79	0.15	0.9	5.6e-27	1.2e-7	2.5e-5
18	1.6e-66	0.036	1.08	1.6e-5	0.56	0.22	2.32	3.1e-3	0.74	0.15	0.9	5.6e-28	1.2e-7	2.5e-5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
44	1.4e-257	0.036	1.04	2.8e-10	0.61	1.1e4	1.41	1.4e-5	0.78	8.9e-6	0.77	5.7e-54	1.2e-7	2.5e-5
45	3.8e-266	0.036	1.03	1.3e-8	0.67	9.7e4	1.22	8.3e-6	0.78	5.7e-7	0.73	5.7e-55	1.2e-7	2.5e-5
46	6.8e-275	0.036	1.03	9.0e-9	0.67	5.2e6	1.09	2.6e-8	0.69	5.9e-7	0.74	5.7e-56	1.2e-7	2.5e-5
47	1.2e-284	0.036	1.03	8.8e-10	0.65	8.8e4	1.29	6.4e-9	0.67	6.0e-7	0.74	5.7e-57	1.2e-7	2.5e-5
48	2.0e-295	0.036	1.04	8.2e-11	0.62	588.0	1.57	4.4e-8	0.71	6.4e-7	0.75	5.7e-58	1.2e-7	2.5e-5
49	1.3e-306	0.036	1.04	3.2e-11	0.62	7.3e4	1.37	8.1e-7	0.76	1.6e-7	0.73	5.7e-59	1.2e-7	2.5e-5
50	1.9e-316	0.036	1.03	7.5e-10	0.66	2.0e7	1.11	1.6e-7	0.74	1.4e-9	0.67	5.7e-60	1.2e-7	2.5e-5
51	3.0e-327	0.036	1.03	7.8e-11	0.64	1.7e5	1.32	1.4e-9	0.68	0.0	0.0	5.7e-61	1.2e-7	2.5e-5

Table 12 Example 4: Results for cumulative runs with $B_0 = F'(u^0) + \hat{\alpha}\|F'(u^0)\|R$ for $\alpha = 10^{-j}$, $\hat{\alpha} = 10^{-l}$ and $(\sigma_k) \equiv \sigma$

(j, l, σ)	$\|\|F\|\|$	$\|\|E\|\|$	δ	ρ_ϵ	β	R	ρ_ζ	$\bar{\rho}_\zeta$	Λ_1	Λ_2^-	Λ_2^+	Λ_3
$(2, 0, 1)$	1e-320	6e-3	1.05	0.53	2e4	1.54	0.71	0.53	5e-299	1e-13	6e-5	6e-8
$(2, 10, 1)$	1e-320	6e-3	1.06	0.48	3e7	1.39	0.7	0.5	2e-297	1e-12	6e-5	6e-8
$(2, 4, 1)$	1e-320	5e-3	1.05	0.58	63	1.76	0.78	0.7	2e-292	1e-11	6e-4	2e-5
$(2, 2, 1)$	1e-320	8e-2	1.05	0.64	1e4	1.3	0.73	0.67	9e-291	3e-10	2e-2	8e-5
$(4, 0, 1)$	1e-320	4e-5	1.07	0.27	663	1.76	0.33	0.28	7e-309	4e-20	6e-9	5e-12
$(4, 10, 1)$	1e-320	4e-5	1.07	0.27	20	1.9	0.42	0.32	1e-296	5e-15	5e-9	2e-10
$(4, 4, 1)$	1e-320	1e-3	1.05	0.47	60	1.8	0.7	0.68	7e-288	3e-12	2e-4	1e-6
$(4, 2, 1)$	9e-321	1e-1	1.05	0.59	5e5	1.37	0.68	0.68	4e-289	1e-11	0.021	6e-5
$(5, 0, 1)$	1e-320	4e-6	1.07	0.21	183	1.95	0.37	0.28	5e-311	1e-22	5e-11	2e-14
$(6, 0, 1)$	1e-320	3e-7	1.08	0.13	0.03	1.97	0.3	0.23	5e-311	2e-23	6e-13	3e-15
$(2, 0, 0.9)$	1e-320	6e-3	1.02	0.8	1e18	0.55	1.01	0.98	3e-55	1e-13	6e-5	3e-8
$(2, 10, 0.9)$	1e-320	6e-3	1.02	0.78	3e16	0.54	1.01	0.98	4e-50	2e-13	6e-5	3e-8
$(2, 4, 0.9)$	1e-320	5e-3	1.02	0.83	7e15	0.44	1.01	1.01	1e-49	1e-12	7e-4	1e-5
$(2, 2, 0.9)$	1e-320	9e-2	1.02	0.87	2e15	0.53	1.01	1.01	2e-54	4e-12	3e-2	3e-4
$(4, 0, 0.9)$	1e-320	4e-5	1.02	0.7	1e17	0.58	1.01	1.01	7e-53	3e-20	6e-9	3e-12
$(4, 10, 0.9)$	1e-320	4e-5	1.02	0.69	1e18	0.56	1.01	1.01	6e-40	3e-16	6e-9	2e-10
$(4, 4, 0.9)$	1e-320	8e-4	1.02	0.79	1e15	0.63	1.01	1.01	4e-45	2e-13	2e-4	2e-6
$(4, 2, 0.9)$	1e-320	9e-2	1.02	0.85	5e15	0.47	1.01	1.0	1e-54	1e-11	2e-2	1e-4

5.2.5 Example 5

We consider $F : \mathbb{R}^{10} \to \mathbb{R}^{10}$ given by

$$F(u) = \begin{pmatrix} u_1 + u_2 + (1+u_3)^2 + u_4 + u_5 + u_6 - u_7^3 - 2u_8 + \sin(u_{10}) \\ u_2 - 2(1+u_3)^2 + 3u_5 - \sin(u_7) - u_{10} + 2 \\ u_1 - u_3^2 + u_5 u_6 u_7 - (1+u_8)(u_9 - 1) - 1 \\ 0.5\ln(1+u_1^2) - 2e^{u_3} + 0.1u_7^{10} + 0.3u_9^4 + 2 \\ \sin(u_1 + u_3 - 10u_2) - u_4^5 - u_6 - u_8 \\ u_1^2 + u_3^2 + (1+u_5)^2 + (1+u_7)^2 + \sin(u_9) - 1 \\ u_6 - u_7 + u_9^2 \\ u_1 + 0.5\ln(1+u_9^2) - 2e^{u_{10}} + 2 \\ u_2 + 0.5\ln(1+u_8^2) - e^{u_{10}} + 1 \\ (1+u_3)^2 + u_8^2 + u_9^2 + u_{10} - 1 \end{pmatrix}.$$

We choose $\alpha = 0.01$ and $\hat{\alpha} = 10^{-j}$ with $j \in \{2, 3, 4\}$ as well as $(\sigma_k) \equiv 1$. The results are comprised in Table 13–14 and indicate that (B_k) converges in all runs. At least every 20-th value of (δ_k) should be close to or larger than $21/20 = 1.05$ and, as in previous experiments, the worst-case value $\delta = 1.03$ in Table 14 suggests a rather uniform behavior of (δ_k) implying convergence of (B_k).

5.2.6 Example 6: Degenerate Jacobian

Finally, let us consider [23, Example 2], i.e.

$$F : \mathbb{R}^3 \to \mathbb{R}^3, \qquad F(u) = \begin{pmatrix} u_1^2 + u_2 + u_3 \\ u_2 - 2u_3^3 \\ 5u_3 + u_3^2 \end{pmatrix}.$$

Table 13 Example 5: Results for one run with $B_0 = F'(u^0)$ and $\alpha = 10^{-2}$

k	$\|\|F_k\|\|$	$\|\|E_k\|\|$	δ_k	ϵ_k	ρ_ϵ^k	β_k	R_k	ζ_k	ρ_ζ^k	$\bar{\zeta}_k$	$\bar{\rho}_\zeta^k$	Λ_1^k	Λ_2^k	Λ_3^k
0	0.089	0.046	-1	-1	-1	-1	-1	-1	-1	0.89	0.89	6.5e-508	2.0e-6	2.6e-5
1	3.4e-4	0.043	2.16	0.014	0.12	-1	-1	0.89	0.94	0.027	0.17	3.6e-9	2.4e-5	4.6e-5
2	3.9e-5	0.043	1.89	8.4e-3	0.2	-1	-1	0.041	0.35	0.047	0.36	4.3e-8	2.5e-5	2.7e-4
3	6.0e-7	0.043	1.9	1.1e-3	0.18	-1	-1	0.057	0.49	0.023	0.39	3.7e-8	2.4e-5	5.4e-5
4	2.5e-8	0.043	1.65	9.9e-4	0.25	-1	-1	0.023	0.47	0.013	0.42	2.9e-9	2.2e-5	3.2e-5
5	3.3e-10	0.043	1.56	4.0e-4	0.27	-1	-1	0.018	0.51	0.024	0.54	9.6e-11	2.2e-5	3.2e-5
6	2.8e-12	0.043	1.43	3.3e-4	0.32	-1	-1	8.0e-3	0.5	0.028	0.6	9.9e-13	2.2e-5	3.2e-5
7	8.2e-15	0.043	1.39	1.1e-4	0.32	-1	-1	0.085	0.73	0.095	0.75	8.1e-15	2.2e-5	3.2e-5
8	5.8e-17	0.043	1.21	1.4e-3	0.48	-1	-1	0.61	0.95	0.66	0.95	4.5e-18	2.2e-5	3.2e-5
9	8.0e-19	0.043	1.12	0.01	0.63	-1	-1	0.65	0.96	0.021	0.68	8.6e-21	2.2e-5	3.1e-5
10	2.1e-19	0.043	1.12	8.6e-3	0.65	-1	-1	4.5e-3	0.61	0.018	0.69	1.6e-21	2.2e-5	2.8e-5
11	2.6e-22	0.043	1.25	5.1e-5	0.44	-1	-1	8.7e-4	0.56	0.018	0.72	5.7e-22	2.2e-5	2.8e-5
12	5.3e-26	0.043	1.25	8.6e-6	0.41	-1	-1	0.024	0.75	0.02	0.74	7.0e-25	2.2e-5	2.8e-5
13	2.4e-28	0.043	1.15	2.4e-4	0.55	-1	-1	0.027	0.77	0.021	0.76	1.1e-28	2.2e-5	2.8e-5
14	1.7e-30	0.043	1.13	2.9e-4	0.58	-1	-1	3.9e-3	0.69	0.023	0.78	6.6e-31	2.2e-5	2.8e-5
15	1.3e-33	0.043	1.16	3.1e-5	0.52	-1	-1	6.8e-3	0.73	0.026	0.8	4.6e-33	2.2e-5	2.8e-5
16	1.7e-36	0.043	1.13	5.8e-5	0.56	-1	-1	0.054	0.84	0.028	0.81	3.2e-36	2.2e-5	2.8e-5
17	1.3e-38	0.043	1.09	7.2e-4	0.67	-1	-1	0.029	0.82	1.4e-3	0.69	2.4e-39	2.2e-5	2.8e-5
18	1.0e-40	0.043	1.09	3.8e-4	0.66	-1	-1	1.8e-3	0.72	3.5e-4	0.66	3.1e-41	2.2e-5	2.8e-5
19	4.4e-44	0.043	1.12	2.3e-5	0.59	-1	-1	3.9e-4	0.68	4.5e-5	0.61	2.3e-43	2.2e-5	2.8e-5
20	4.4e-48	0.043	1.13	5.2e-6	0.56	-1	-1	1.4e-5	0.59	3.1e-5	0.61	1.0e-46	2.2e-5	2.8e-5
21	1.5e-53	0.043	1.15	1.8e-7	0.49	9.7e-4	3.62	2.6e-6	0.56	3.3e-5	0.63	1.0e-50	2.2e-5	2.8e-5
22	1.0e-59	0.043	1.14	3.4e-8	0.47	4.9e-4	3.59	2.4e-5	0.63	9.5e-6	0.6	3.5e-56	2.2e-5	2.8e-5
23	6.0e-65	0.043	1.11	3.1e-7	0.54	0.25	2.2	3.7e-5	0.65	4.6e-5	0.66	2.3e-62	2.2e-5	2.8e-5
24	5.5e-70	0.043	1.1	4.8e-7	0.56	0.49	2.1	3.0e-6	0.6	4.3e-5	0.67	1.4e-67	2.2e-5	2.8e-5
25	4.1e-76	0.043	1.11	3.9e-8	0.52	0.24	2.18	1.4e-5	0.65	2.9e-5	0.67	1.3e-72	2.2e-5	2.8e-5
26	1.4e-81	0.043	1.09	1.8e-7	0.56	1.7	1.94	5.7e-6	0.64	2.3e-5	0.67	9.4e-79	2.2e-5	2.8e-5
27	2.1e-87	0.043	1.09	7.5e-8	0.56	5.7	1.81	2.4e-5	0.68	9.4e-7	0.61	3.3e-84	2.2e-5	2.8e-5
28	1.3e-92	0.043	1.08	3.2e-7	0.6	0.16	2.28	5.7e-7	0.61	3.7e-7	0.6	4.7e-90	2.2e-5	2.8e-5
29	1.8e-99	0.043	1.09	7.5e-9	0.54	6.9e-5	4.1	2.1e-8	0.55	4.0e-7	0.61	2.9e-95	2.2e-5	2.8e-5
30	9.4e-108	0.043	1.1	2.7e-10	0.49	3.7e-6	4.63	1.5e-8	0.56	3.8e-7	0.62	4.1e-102	2.2e-5	2.8e-5
31	3.5e-116	0.043	1.09	1.9e-10	0.5	0.075	2.26	4.8e-7	0.63	1.0e-7	0.6	2.1e-110	2.2e-5	2.8e-5
32	4.2e-123	0.043	1.07	6.3e-9	0.56	86.0	1.62	4.8e-7	0.64	3.8e-7	0.64	7.9e-119	2.2e-5	2.8e-5
33	5.1e-130	0.043	1.07	6.4e-9	0.57	0.11	2.27	2.1e-8	0.59	3.6e-7	0.65	9.5e-126	2.2e-5	2.8e-5
34	2.7e-138	0.043	1.07	2.8e-10	0.53	3.2e-3	2.71	8.6e-8	0.63	2.8e-7	0.65	1.2e-132	2.2e-5	2.8e-5
35	5.9e-146	0.043	1.07	1.1e-9	0.56	1.2	1.98	3.6e-7	0.66	8.6e-8	0.64	6.2e-141	2.2e-5	2.8e-5
36	5.4e-153	0.043	1.06	4.8e-9	0.6	1.4	1.96	9.5e-8	0.65	8.3e-9	0.6	1.3e-148	2.2e-5	2.8e-5
37	1.3e-160	0.043	1.06	1.2e-9	0.58	2.4e-3	2.83	2.4e-9	0.59	5.9e-9	0.61	1.2e-155	2.2e-5	2.8e-5
38	7.7e-170	0.043	1.07	3.1e-11	0.54	2.1e-4	3.08	5.9e-9	0.62	1.8e-11	0.53	2.9e-163	2.2e-5	2.8e-5
39	1.1e-178	0.043	1.06	7.8e-11	0.56	0.14	2.18	1.8e-11	0.54	4.8e-15	0.44	1.7e-172	2.2e-5	2.8e-5
40	5.2e-190	0.043	1.07	2.4e-13	0.49	8.9e-3	2.39	4.8e-15	0.45	5.9e-18	0.38	2.6e-181	2.2e-5	2.8e-5
41	6.3e-205	0.043	1.09	6.3e-17	0.41	1.9e-3	2.4	9.7e-17	0.42	9.1e-17	0.42	1.2e-192	2.2e-5	2.8e-5
42	1.5e-221	0.043	1.09	1.3e-18	0.38	1.1e-3	2.4	1.9e-16	0.43	1.0e-16	0.42	1.4e-207	2.2e-5	2.8e-5
43	7.5e-238	0.043	1.08	2.6e-18	0.4	2.6e-5	2.71	9.4e-17	0.43	9.0e-18	0.41	3.5e-224	2.2e-5	2.8e-5
44	1.8e-254	0.043	1.08	1.2e-18	0.4	5.3e-6	2.84	1.8e-17	0.42	8.7e-18	0.42	1.7e-240	2.2e-5	2.8e-5
45	7.8e-272	0.043	1.07	2.3e-19	0.39	1.5e-4	2.51	3.1e-16	0.46	3.2e-16	0.46	4.0e-257	2.2e-5	2.8e-5
46	6.1e-288	0.043	1.06	4.1e-18	0.43	1.2e-4	2.58	3.5e-16	0.47	3.3e-17	0.45	1.8e-274	2.2e-5	2.8e-5
47	5.4e-304	0.043	1.06	4.6e-18	0.44	8.3e-4	2.43	3.3e-17	0.45	2.2e-20	0.39	1.4e-290	2.2e-5	2.8e-5
48	4.4e-321	0.043	1.06	4.3e-19	0.42	4.2e-6	2.83	2.2e-20	0.4	0.0	0.0	1.2e-306	2.2e-5	2.8e-5

Table 14 Example 5: Results for cumulative runs with $B_0 = F'(u^0)$, $\alpha = 10^{-j}$, and $(\sigma_k) \equiv 1$ (represented by $\sigma = 1$), respectively, $\sigma_k = 0.9$ for $k \leq 4$ and $\sigma_k = 1$ else (represented by $\sigma = Y$)

(j, σ)	$\|\|F\|\|$	$\|\|E\|\|$	δ	ρ_ϵ	β	R	ρ_ζ	$\bar{\rho}_\zeta$	Λ_1	Λ_2^-	Λ_2^+	Λ_3
$(2, 1)$	1e-320	9e-3	1.03	0.73	1e9	0.99	0.96	0.81	5e-300	2e-10	1e-4	2e-7
$(2, Y)$	1e-320	9e-3	1.03	0.72	2e8	1.0	0.89	0.83	9e-300	3e-12	1e-4	2e-7
$(5, 1)$	1e-320	9e-6	1.05	0.39	2366	1.59	0.7	0.69	1e-301	1e-18	9e-11	2e-13
$(5, Y)$	1e-320	9e-6	1.05	0.46	2288	1.52	0.67	0.66	1e-294	2e-16	8e-11	3e-13
$(8, 1)$	9e-321	9e-9	1.06	0.25	1777	1.6	0.61	0.6	2e-300	4e-23	9e-17	2e-19
$(8, Y)$	1e-320	9e-9	1.06	0.25	643	1.69	0.63	0.61	1e-301	2e-22	1e-16	3e-19

Table 15 Example 6: Results for one run with $B_0 = F'(u^0) + \hat{\alpha}\|F'(u^0)\|R$ for $\hat{\alpha} = 0.1$, $\alpha = 0.1$, and $(\sigma_k) \equiv 0.9$

k	$\|\|F_k\|\|$	$\|\|E_k\|\|$	δ_k	ϵ_k	ρ_ϵ^k	β_k	R_k	ζ_k	ρ_ζ^k	$\bar{\zeta}_k$	$\bar{\rho}_\zeta^k$	Λ_1^k	Λ_2^k
0	0.38	0.91	-1	-1	-1	-1	-1	-1	-1	0.83	0.83	0.22	0.42
1	0.023	0.95	2.05	0.13	0.36	-1	-1	0.59	0.77	0.33	0.57	0.25	0.44
2	0.023	0.85	1.39	0.31	0.68	-1	-1	0.39	0.73	0.072	0.42	0.14	0.36
3	0.015	0.77	1.37	0.29	0.73	-1	-1	0.13	0.6	0.17	0.64	0.059	0.36
4	4.4e-3	0.75	1.59	0.12	0.65	-1	-1	0.032	0.5	0.19	0.72	0.044	0.36
5	4.1e-4	0.74	1.67	0.039	0.58	-1	-1	0.11	0.7	0.076	0.65	0.039	0.36
6	1.6e-4	0.74	1.46	0.058	0.67	-1	-1	0.089	0.71	0.013	0.54	0.025	0.35
7	1.2e-4	0.74	1.45	0.054	0.69	3.2	1.43	2.0e-3	0.46	0.015	0.59	0.01	0.35
8	3.6e-5	0.74	1.82	9.0e-3	0.59	0.093	4.03	7.9e-3	0.58	6.9e-3	0.58	5.8e-3	0.35
9	2.4e-5	0.74	1.83	7.3e-3	0.61	0.088	3.95	7.8e-3	0.62	9.0e-4	0.5	3.7e-3	0.35
10	1.3e-5	0.74	1.78	6.5e-3	0.63	0.45	2.37	3.8e-3	0.6	4.7e-3	0.61	2.5e-3	0.35
11	3.4e-6	0.74	1.77	3.7e-3	0.63	2.4	1.73	5.0e-4	0.53	4.2e-3	0.63	1.9e-3	0.35
12	4.4e-7	0.74	1.81	1.3e-3	0.6	0.39	2.33	3.1e-3	0.64	1.0e-3	0.59	1.3e-3	0.35
13	2.8e-7	0.74	1.7	1.7e-3	0.64	0.61	2.17	1.5e-3	0.63	4.9e-4	0.58	7.1e-4	0.35
14	1.5e-7	0.74	1.76	1.0e-3	0.63	12.0	1.46	7.1e-6	0.45	5.0e-4	0.6	3.7e-4	0.35
15	5.0e-8	0.74	1.91	3.0e-4	0.6	5.6	1.65	3.1e-4	0.6	1.8e-4	0.58	2.3e-4	0.35
16	3.6e-8	0.74	1.88	3.0e-4	0.62	7.0	1.61	2.6e-4	0.62	7.4e-5	0.57	1.5e-4	0.35
17	1.6e-8	0.74	1.85	2.3e-4	0.63	17.0	1.49	9.6e-5	0.6	1.7e-4	0.62	1.0e-4	0.35
18	3.8e-9	0.74	1.85	1.2e-4	0.62	70.0	1.36	5.5e-5	0.6	1.1e-4	0.62	7.2e-5	0.35
19	7.0e-10	0.74	1.86	5.3e-5	0.61	17.0	1.55	1.0e-4	0.63	1.1e-5	0.57	4.7e-5	0.35
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
120	1.5e-49	0.74	1.98	5.9e-25	0.63	7.1e21	1.05	7.8e-26	0.62	1.2e-25	0.62	4.5e-25	0.35
121	6.7e-50	0.74	1.98	4.0e-25	0.63	1.0e22	1.05	1.1e-25	0.62	7.7e-27	0.61	2.9e-25	0.35
122	2.9e-50	0.74	1.98	2.7e-25	0.63	1.6e22	1.05	5.1e-26	0.62	4.3e-26	0.62	1.8e-25	0.35
123	1.1e-50	0.74	1.98	1.7e-25	0.63	2.8e22	1.05	7.2e-27	0.62	3.6e-26	0.62	1.2e-25	0.35
124	3.8e-51	0.74	1.98	1.0e-25	0.63	4.6e22	1.05	2.7e-26	0.62	8.6e-27	0.62	7.7e-26	0.35
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
460	3.2e-183	0.74	1.99	8.9e-92	0.63	4.9e88	1.01	3.8e-95	0.62	3.0e-94	0.63	6.7e-92	0.35
461	1.3e-183	0.74	1.99	5.7e-92	0.63	7.7e88	1.01	2.2e-94	0.63	8.4e-95	0.63	4.3e-92	0.35
462	5.3e-184	0.74	1.99	3.6e-92	0.63	1.2e89	1.01	1.4e-94	0.63	5.9e-95	0.63	2.7e-92	0.35
463	2.1e-184	0.74	1.99	2.3e-92	0.63	1.9e89	1.01	1.7e-95	0.62	7.6e-95	0.63	1.7e-92	0.35
464	8.7e-185	0.74	1.99	1.5e-92	0.63	3.0e89	1.01	4.5e-95	0.63	3.1e-95	0.63	1.1e-92	0.35
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
806	3.1e-319	0.74	2.0	8.8e-160	0.64	5.0e156	1.01	3.8e-164	0.63	3.3e-164	0.63	6.6e-160	0.35
807	1.3e-319	0.74	2.0	5.6e-160	0.64	7.9e156	1.01	7.8e-165	0.63	2.6e-164	0.63	4.2e-160	0.35
808	5.1e-320	0.74	2.0	3.5e-160	0.64	1.2e157	1.01	2.0e-164	0.63	5.3e-165	0.63	2.7e-160	0.35
809	2.1e-320	0.74	2.0	2.3e-160	0.64	1.9e157	1.01	1.2e-164	0.63	6.3e-165	0.63	1.7e-160	0.35
810	8.3e-321	0.74	2.0	1.4e-160	0.64	3.1e157	1.01	3.3e-166	0.63	6.7e-165	0.63	1.1e-160	0.35

Note that $F'(0)$ is not invertible, so this example does not satisfy the standard assumptions for q-superlinear convergence of the iterates. We choose $\alpha = 0.1$, $\hat{\alpha} \in \{0, 0.1\}$ as well as $(\sigma_k) \equiv 1$ and $(\sigma_k) \equiv 1 - (k+2)^{-2}$. Based on Theorem 7 1), which applies because Assumption 2 is readily seen to hold with $\phi = (1, 0, 0)^T$, we expect $\bar{s} = \phi$ in this example, so we replace $\hat{s}^{\bar{k}}$ in the definition (22) of $\bar{\zeta}_k$ by ϕ. The results are displayed in Table 15 and Table 16, where we suppress (Λ_3^k) and Λ_3 since they agree with $(\|E_k\|)$ and $\|E\|$, respectively. From Theorem 7 we obtain the convergence of (B_k), which is numerically confirmed in Table 16. Under the assumptions of Lemma 6 the iterates are q-linear convergent with exact asymptotic convergence factor $\frac{\sqrt{5}-1}{2}$, so

6 Summary

We have investigated under which conditions the matrices of the Broyden–like method converge, with particular emphasis on Broyden's method. Our findings suggest that the Broyden–like matrices (B_k) converge frequently (possi-

Table 16 Example 6: Results for cumulative runs with $B_0 = F'(u^0) + \hat{\alpha}\|F'(u^0)\|R$, where X represents the choice $\sigma_k = 1 - (k+2)^{-2}$ for all $k \geq 0$

$(\alpha, \hat{\alpha}, \sigma)$	$\|\|F\|\|$	$\|\|E\|\|$	δ	ρ_ϵ	β	R	ρ_ζ	$\bar{\rho}_\zeta$	Λ_1	Λ_2^-	Λ_2^+	Λ_3
$(1,1,1)$	1e-320	0.23	1.95	0.73	4e160	0.98	0.72	0.72	2e-159	1e-2	0.65	0.29
$(1,0,1)$	1e-320	2e-3	2.0	0.62	3e157	1.01	0.32	0.32	5e-162	4e-4	0.19	2e-3
$(2,1,1)$	1e-320	0.24	1.95	0.72	1e160	0.99	0.72	0.72	4e-159	2e-3	0.6	0.24
$(2,0,1)$	1e-320	8e-4	2.0	0.62	3e157	1.01	0.09	0.09	3e-165	4e-5	0.01	1e-3
$(10,0,1)$	1e-320	8e-12	2.0	0.60	3e157	1.01	8e-6	8e-6	2e-188	4e-21	7e-14	2e-9
$(1,1,0.9)$	1e-320	0.23	1.94	0.76	3e162	0.98	0.76	0.76	3e-160	8e-3	0.67	0.25
$(2,1,0.9)$	1e-320	0.27	1.95	0.73	2e159	0.98	0.73	0.73	2e-159	9e-3	0.48	0.27
$(1,1,X)$	1e-320	0.23	1.96	0.72	5e162	0.98	0.72	0.72	2.4e-160	1e-2	0.58	0.26
$(2,1,X)$	1e-320	0.19	1.94	0.71	5e161	0.98	0.71	0.71	5e-159	2e-2	0.56	0.19

bly always) if the standard assumptions for q-superlinear convergence of the iterates are satisfied. More precisely, the updates $(B_{k+1} - B_k)$ converge at least r-linearly to zero and possibly with an r-order larger than one. The iterates (u^k) were found to be convergent with a q-order larger than one. If the Jacobian at the root is singular, we were able to prove that some of the new conditions for convergence of (B_k) are actually satisfied. We proposed the conjecture that for Broyden's method the sequence $(\|B_{k+1} - B_k\|)$ converges multi–step q-quadratically to zero under the standard assumptions for q-superlinear convergence; this would imply $\sum_k \|B_{k+1} - B_k\| < \infty$.

References

1. Adly, S., Ngai, H.V.: Quasi-Newton methods for solving nonsmooth equations: Generalized Dennis-Moré theorem and Broyden's update. J. Convex Anal. **25**(4), 1075–1104 (2018)
2. Al-Baali, M., Spedicato, E., Maggioni, F.: Broyden's quasi-Newton methods for a nonlinear system of equations and unconstrained optimization: a review and open problems. Optim. Methods Softw. **29**(5), 937–954 (2014). DOI 10.1080/10556788.2013.856909
3. Aragón Artacho, F., Belyakov, A., Dontchev, A., López, M.: Local convergence of quasi-Newton methods under metric regularity. Comput. Optim. Appl. **58**(1), 225–247 (2014). DOI 10.1007/s10589-013-9615-y
4. Arguillère, S.: Approximation of sequences of symmetric matrices with the symmetric rank-one algorithm and applications. SIAM J. Matrix Anal. Appl. **36**(1), 329–347 (2015). DOI 10.1137/130926584
5. Broyden, C.: A class of methods for solving nonlinear simultaneous equations. Math. Comput. **19**, 577–593 (1965). DOI 10.2307/2003941
6. Conn, A.R., Gould, N.I.M., Toint, P.L.: Convergence of quasi-Newton matrices generated by the symmetric rank one update. Math. Program. **50**(2 (A)), 177–195 (1991). DOI 10.1007/BF01594934
7. Decker, D.W., Kelley, C.T.: Broyden's method for a class of problems having singular Jacobian at the root. SIAM J. Numer. Anal. **22**, 566–574 (1985). DOI 10.1137/0722034
8. Dennis, J., More, J.J.: A characterization of superlinear convergence and its application to quasi-Newton methods. Math. Comput. **28**, 549–560 (1974). DOI 10.2307/2005926
9. Dennis, J., More, J.J.: Quasi-Newton methods, motivation and theory. SIAM Rev. **19**, 46–89 (1977). DOI 10.1137/1019005
10. Dennis, J., Schnabel, R.B.: Numerical methods for unconstrained optimization and nonlinear equations. Repr. Philadelphia, PA: SIAM (1996). DOI 10.1137/1.9781611971200
11. jun. Dennis, J.E., More, J.J.: Quasi-Newton methods, motivation and theory. SIAM Rev. **19**, 46–89 (1977). DOI 10.1137/1019005

12. Fayez Khalfan, H., Byrd, R.H., Schnabel, R.B.: A theoretical and experimental study of the symmetric rank-one update. SIAM J. Optim. **3**(1), 1–24 (1993). DOI 10.1137/0803001
13. Gay, D.M.: Some convergence properties of Broyden's method. SIAM J. Numer. Anal. **16**, 623–630 (1979). DOI 10.1137/0716047
14. Ge, R., Powell, M.J.D.: The convergence of variable metric matrices in unconstrained optimization. Math. Program. **27**, 123–143 (1983). DOI 10.1007/BF02591941
15. Gower, R.M., Richtárik, P.: Randomized quasi-Newton updates are linearly convergent matrix inversion algorithms. SIAM J. Matrix Anal. Appl. **38**(4), 1380–1409 (2017). DOI 10.1137/16M1062053
16. Griewank, A.: Broyden updating, the good and the bad! Doc. Math. (Bielefeld) pp. 301–315 (2012). URL `https://www.emis.de/journals/DMJDMV/vol-ismp/45_griewank-andreas-broyden.pdf`
17. Horn, R.A., Johnson, C.R.: Matrix analysis, 2nd edn. Cambridge: Cambridge University Press (2013). DOI 10.1017/9781139020411
18. Jankowska, J.: Theory of multivariate secant methods. SIAM J. Numer. Anal. **16**, 547–562 (1979). DOI 10.1137/0716042
19. Li, D., Fukushima, M.: A derivative-free line search and global convergence of Broyden-like method for nonlinear equations. Optim. Methods Softw. **13**(3), 181–201 (2000). DOI 10.1080/10556780008805782
20. Li, D., Zeng, J., Zhou, S.: Convergence of Broyden-like matrix. Appl. Math. Lett. **11**(5), 35–37 (1998). DOI 10.1016/S0893-9659(98)00076-7
21. Mannel, F.: Improved local convergence of Broyden's method for mixed linear–nonlinear systems of equations. Submitted (2020). URL `https://imsc.uni-graz.at/mannel/ImprovGay.pdf`
22. Mannel, F.: On some convergence properties of the Broyden–like method for mixed linear–nonlinear systems of equations. Submitted (2020). URL `https://imsc.uni-graz.at/mannel/CBLMMLNSE.pdf`
23. Mannel, F.: On the convergence of the Broyden matrices for a class of singular problems. Submitted (2020). URL `https://imsc.uni-graz.at/mannel/CGB_sing.pdf`
24. Mannel, F.: On the rate of convergence of the Broyden–like method. In preparation (2020). URL `https://imsc.uni-graz.at/mannel/BLMFTQS.pdf`
25. Mannel, F., Rund, A.: A hybrid semismooth quasi-Newton method for nonsmooth optimal control with PDEs. Optim. Eng. **Online First** (2020). DOI 10.1007/s11081-020-09523-w
26. Martínez, J.M.: Practical quasi-Newton methods for solving nonlinear systems. J. Comput. Appl. Math. **124**(1-2), 97–121 (2000). DOI 10.1016/S0377-0427(00)00434-9
27. More, J., Trangenstein, J.: On the global convergence of Broyden's method. Math. Comput. **30**, 523–540 (1976). DOI 10.2307/2005323
28. Muoi, P.Q., Hào, D.N., Maass, P., Pidcock, M.: Semismooth Newton and quasi-Newton methods in weighted ℓ^1-regularization. J. Inverse Ill-Posed Probl. **21**(5), 665–693 (2013). DOI 10.1515/jip-2013-0031
29. Ortega, J.M., Rheinboldt, W.C.: Iterative solution of nonlinear equations in several variables., vol. 30. Philadelphia, PA: SIAM, Society for Industrial and Applied Mathematics (2000). DOI 10.1137/1.9780898719468
30. Powell, M.J.D.: A new algorithm for unconstrained optimization. Nonlinear Programming, Proc. Sympos. Math. Res. Center, Univ. Wisconsin, Madison 1970, 31-65 (1970). DOI 10.1016/B978-0-12-597050-1.50006-3
31. Rodomanov, A., Nesterov, Y.: Greedy quasi-newton methods with explicit superlinear convergence (2020)
32. Sachs, E.: Convergence rates of quasi-Newton algorithms for some nonsmooth optimization problems. SIAM J. Control Optim. **23**, 401–418 (1985). DOI 10.1137/0323026
33. Stoer, J.: The convergence of matrices generated by rank-2 methods from the restricted β-class of Broyden. Numer. Math. **44**, 37–52 (1984). DOI 10.1007/BF01389753
34. Ulbrich, M., Ulbrich, S.: Nichtlineare Optimierung. Basel: Birkhäuser (2012). DOI 10.1007/978-3-0346-0654-7