Mathematical Programming (2022) 194:159–190
https://doi.org/10.1007/s10107-021-01622-5

FULL LENGTH PAPER

Series A

Rates of superlinear convergence for classical quasi-Newton methods

Anton Rodomanov[1] · **Yurii Nesterov**[2]

Received: 30 March 2020 / Accepted: 18 January 2021 / Published online: 8 February 2021

Abstract
We study the local convergence of classical quasi-Newton methods for nonlinear optimization. Although it was well established a long time ago that asymptotically these methods converge superlinearly, the corresponding rates of convergence still remain unknown. In this paper, we address this problem. We obtain first explicit non-asymptotic rates of superlinear convergence for the standard quasi-Newton methods, which are based on the updating formulas from the convex Broyden class. In particular, for the well-known DFP and BFGS methods, we obtain the rates of the form $(\frac{nL^2}{\mu^2 k})^{k/2}$ and $(\frac{nL}{\mu k})^{k/2}$ respectively, where k is the iteration counter, n is the dimension of the problem, μ is the strong convexity parameter, and L is the Lipschitz constant of the gradient.

Keywords Quasi-Newton methods · Convex Broyden class · DFP · BFGS · Superlinear convergence · Local convergence · Rate of convergence

Mathematics Subject Classification 90C53 · 90C30 · 68Q25

Research results in this paper were obtained with support of ERC Advanced Grant 788368.

✉ Anton Rodomanov
anton.rodomanov@uclouvain.be

Yurii Nesterov
yurii.nesterov@uclouvain.be

1 Institute of Information and Communication Technologies, Electronics and Applied Mathematics, Catholic University of Louvain, Louvain-la-Neuve, Belgium

2 Center for Operations Research and Econometrics, Catholic University of Louvain, Louvain-la-Neuve, Belgium

1 Introduction

Motivation In this work, we investigate the classical quasi-Newton algorithms for smooth unconstrained optimization, the main examples of which are the *Davidon–Fletcher–Powell (DFP) method* [1,2] and the *Broyden–Fletcher–Goldfarb–Shanno (BFGS)* method [3–7]. These algorithms are based on the idea of replacing the exact Hessian in the Newton method with some approximation, that is updated in iterations according to certain formulas, involving only the gradients of the objective function. For an introduction into the topic, see [8] and [9, Chapter 6]; also see [10] for the treatment of quasi-Newton algorithms in the context of nonsmooth optimization and [11–13] for randomized variants of quasi-Newton methods.

One of the questions about quasi-Newton methods, that has been extensively studied in the literature, is their *superlinear convergence*. First theoretical results here were obtained for the methods with exact line search, first by Powell [14], who analyzed the DFP method, and then by Dixon [15,16], who showed that with the exact line search all quasi-Newton algorithms in the Broyden family [17] coincide. Soon after that Broyden, Dennis and Moré [18] considered the quasi-Newton algorithms without line search and proved the local superlinear convergence of DFP, BFGS and several other methods. Their analysis was based on the Frobenius-norm potential function. Later, Dennis and Moré [19] unified the previous proofs by establishing the necessary and sufficient condition of superlinear convergence. This condition together with the original analysis of Broyden, Dennis and Moré have been applied since then in almost every work on quasi-Newton methods for proving superlinear convergence (see e.g. [20–27]). Finally, one should mention that an important contribution to the theoretical analysis of quasi-Newton methods has been made by Byrd, Liu, Nocedal and Yuan in the series of works [28–30], where they introduced a new potential function by combining the trace with the logarithm of determinant.

However, the theory of superlinear convergence of quasi-Newton methods is still far from being complete. The main reason for this is that all currently existing results on superlinear convergence of quasi-Newton methods are only asymptotic: they simply show that the ratio of successive residuals in the method tends to zero as the number of iterations goes to infinity, without providing any specific bounds on the corresponding *rate* of convergence. It is therefore important to obtain some *explicit* and *non-asymptotic* rates of superlinear convergence for quasi-Newton methods.

This observation was the starting point for a recent work [31], where the *greedy* analogs of the classical quasi-Newton methods have been developed. As opposed to the classical quasi-Newton methods, which use the difference of successive iterates for updating Hessian approximations, these methods employ basis vectors, greedily selected to maximize a certain measure of progress. As shown in [31], greedy quasi-Newton methods have superlinear convergence rate of the form $(1 - \frac{\mu}{nL})^{k^2/2}(\frac{nL}{\mu})^k$, where k is the iteration counter, n is the dimension of the problem, μ is the strong convexity parameter, and L is the Lipschitz constant of the gradient.

In this work, we continue the same line of research but now we study the *classical* quasi-Newton methods. Namely, we consider the methods, based on the updates from the convex Broyden class, which is formed by all convex combinations of the DFP and BFGS updates. For this class, we derive explicit bounds on the rate of superlinear convergence of standard quasi-Newton methods without line search. In particular, for the standard DFP and BFGS methods, we obtain the rates of the form $(\frac{nL^2}{\mu^2 k})^{k/2}$ and $(\frac{nL}{\mu k})^{k/2}$ respectively.

Contents This paper is organized as follows. First, in Sect. 2, we study the convex Broyden class of updating rules for approximating a self-adjoint positive definite linear operator, and establish several important properties of this class. Then, in Sect. 3, we analyze the standard quasi-Newton scheme, based on the updating rules from the convex Broyden class, as applied to minimizing a quadratic function. We show that this scheme has the same rate of linear convergence as that of the classical gradient method, and also a superlinear convergence rate of the form $(\frac{Q}{k})^{k/2}$, where $Q \geq 1$ is a certain constant, related to the condition number, and depending on the method. After that, in Sect. 4, we consider the general problem of unconstrained minimization and the corresponding quasi-Newton scheme for solving it. We show that, for this scheme, it is possible to prove absolutely the same results as for the quadratic function, provided that the starting point is sufficiently close to the solution. In Sect. 5, we compare the rates of superlinear convergence, that we obtain for the classical quasi-Newton methods, with the corresponding rates of the greedy quasi-Newton methods. Sect. 6 contains some auxiliary results, that we use in our analysis.

Notation In what follows, $\mathbb{E}$ denotes an arbitrary n-dimensional real vector space. Its dual space, composed by all linear functionals on $\mathbb{E}$, is denoted by $\mathbb{E}^*$. The value of a linear function $s \in \mathbb{E}^*$, evaluated at point $x \in \mathbb{E}$, is denoted by $\langle s, x \rangle$.

For a smooth function $f : \mathbb{E} \to \mathbb{R}$, we denote by $\nabla f(x)$ and $\nabla^2 f(x)$ its gradient and Hessian respectively, evaluated at a point $x \in \mathbb{E}$. Note that $\nabla f(x) \in \mathbb{E}^*$, and $\nabla^2 f(x)$ is a self-adjoint linear operator from $\mathbb{E}$ to $\mathbb{E}^*$.

The partial ordering of self-adjoint linear operators is defined in the standard way. We write $A \preceq A_1$ for $A, A_1 : \mathbb{E} \to \mathbb{E}^*$ if $\langle (A_1 - A)x, x \rangle \geq 0$ for all $x \in \mathbb{E}$, and $W \preceq W_1$ for $W, W_1 : \mathbb{E}^* \to \mathbb{E}$ if $\langle s, (W_1 - W)s \rangle \geq 0$ for all $s \in \mathbb{E}^*$.

Any self-adjoint positive definite linear operator $A : \mathbb{E} \to \mathbb{E}^*$ induces in the spaces $\mathbb{E}$ and $\mathbb{E}^*$ the following pair of conjugate Euclidean norms:

$$\|h\|_A \overset{\text{def}}{=} \langle Ah, h \rangle^{1/2}, \quad h \in \mathbb{E}, \qquad \|s\|_A^* \overset{\text{def}}{=} \langle s, A^{-1}s \rangle^{1/2}, \quad s \in \mathbb{E}^*. \tag{1.1}$$

When $A = \nabla^2 f(x)$, where $f : \mathbb{E} \to \mathbb{R}$ is a smooth function with positive definite Hessian, and $x \in \mathbb{E}$, we prefer to use notation $\| \cdot \|_x$ and $\| \cdot \|_x^*$, provided that there is no ambiguity with the reference function f.

Sometimes, in the formulas, involving products of linear operators, it is convenient to treat $x \in \mathbb{E}$ as a linear operator from $\mathbb{R}$ to $\mathbb{E}$, defined by $x\alpha = \alpha x$, and x^* as a linear operator from $\mathbb{E}^*$ to $\mathbb{R}$, defined by $x^* s = \langle s, x \rangle$. Likewise, any $s \in \mathbb{E}^*$ can be treated as a linear operator from $\mathbb{R}$ to $\mathbb{E}^*$, defined by $s\alpha = \alpha s$, and s^* as a linear operator from $\mathbb{E}$ to $\mathbb{R}$, defined by $s^* x = \langle s, x \rangle$. In this case, xx^* and ss^* are rank-one self-adjoint linear operators from $\mathbb{E}^*$ to $\mathbb{E}$ and from $\mathbb{E}^*$ to $\mathbb{E}$ respectively, acting as follows:

$$(xx^*)s = \langle s, x \rangle x, \qquad (ss^*)x = \langle s, x \rangle s, \qquad x \in \mathbb{E}, \ s \in \mathbb{E}^*.$$

Given two self-adjoint linear operators $A : \mathbb{E} \to \mathbb{E}^*$ and $W : \mathbb{E}^* \to \mathbb{E}$, we define the trace and the determinant of A with respect to W as follows:

$$\langle W, A\rangle \stackrel{\text{def}}{=} \mathrm{Tr}(WA), \qquad \mathrm{Det}(W, A) \stackrel{\text{def}}{=} \mathrm{Det}(WA).$$

Note that WA is a linear operator from $\mathbb{E}$ to itself, and hence its trace and determinant are well-defined real numbers (they coincide with the trace and determinant of the matrix representation of WA with respect to an arbitrary chosen basis in the space $\mathbb{E}$, and the result is independent of the particular choice of the basis). In particular, if W is positive definite, then $\langle W, A\rangle$ and $\mathrm{Det}(W, A)$ are respectively the sum and the product of the eigenvalues[1] of A relative to W^{-1}. Observe that $\langle \cdot, \cdot\rangle$ is a bilinear form, and for any $x \in \mathbb{E}$, we have

$$\langle Ax, x\rangle = \langle xx^*, A\rangle. \tag{1.2}$$

When A is invertible, we also have

$$\langle A^{-1}, A\rangle = n, \qquad \mathrm{Det}(A^{-1}, \delta A) = \delta^n. \tag{1.3}$$

for any $\delta \in \mathbb{R}$. Also recall the following multiplicative formula for the determinant:

$$\mathrm{Det}(W, A) = \mathrm{Det}(W, G)\mathrm{Det}(G^{-1}, A), \tag{1.4}$$

which is valid for any invertible linear operator $G : \mathbb{E} \to \mathbb{E}^*$. If the operator W is positive semidefinite, and $A \preceq A_1$ for some self-adjoint linear operator $A_1 : \mathbb{E} \to \mathbb{E}^*$, then $\langle W, A\rangle \le \langle W, A_1\rangle$ and $\mathrm{Det}(W, A) \le \mathrm{Det}(W, A_1)$. Similarly, if A is positive semidefinite and $W \preceq W_1$ for some self-adjoint linear operator $W_1 : \mathbb{E}^* \to \mathbb{E}$, then $\langle W, A\rangle \le \langle W_1, A\rangle$ and $\mathrm{Det}(W, A) \le \mathrm{Det}(W_1, A)$.

2 Convex Broyden class

Let A and G be two self-adjoint positive definite linear operators from $\mathbb{E}$ to $\mathbb{E}^*$, where A is the target operator, which we want to approximate, and G is the current approximation of the operator A. The *Broyden family* of quasi-Newton updates of G with respect to A along a direction $u \in \mathbb{E} \setminus \{0\}$, is the following class of updating formulas, parameterized by a scalar $\phi \in \mathbb{R}$:

$$\begin{aligned} \mathrm{Broyd}_\phi(A, G, u) \stackrel{\text{def}}{=} \phi \left[G - \frac{Auu^*G + Guu^*A}{\langle Au, u\rangle} + \left(\frac{\langle Gu, u\rangle}{\langle Au, u\rangle} + 1 \right) \frac{Auu^*A}{\langle Au, u\rangle} \right] \\ + (1 - \phi) \left[G - \frac{Guu^*G}{\langle Gu, u\rangle} + \frac{Auu^*A}{\langle Au, u\rangle} \right]. \end{aligned} \tag{2.1}$$

[1] Recall that, for linear operators $A, B : \mathbb{E} \to \mathbb{E}^*$, a scalar $\lambda \in \mathbb{R}$ is called a (relative) eigenvalue of A with respect to B if $Ax = \lambda Bx$ for some $x \in \mathbb{E} \setminus \{0\}$. If A, B are self-adjoint and B is positive definite, it is known that there exist eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ and a basis $x_1, \ldots, x_n \in \mathbb{E}$, such that $Ax_i = \lambda_i Bx_i$, $\|x_i\|_B = 1$, $\langle Bx_i, x_j\rangle = 0$ for all $1 \le i, j \le n$, $i \ne j$.

Note that $\mathrm{Broyd}_\phi(A, G, u)$ depends on A only through the product Au. For the sake of convenience, we also define $\mathrm{Broyd}_\phi(A, G, u) = G$ when $u = 0$.

Two important members of the Broyden family, DFP and BFGS updates, correspond to the values $\phi = 1$ and $\phi = 0$ respectively:

$$\mathrm{DFP}(A, G, u) \overset{\text{def}}{=} G - \frac{Auu^*G + Guu^*A}{\langle Au, u\rangle} + \left(\frac{\langle Gu, u\rangle}{\langle Au, u\rangle} + 1\right)\frac{Auu^*A}{\langle Au, u\rangle},$$
$$\mathrm{BFGS}(A, G, u) \overset{\text{def}}{=} G - \frac{Guu^*G}{\langle Gu, u\rangle} + \frac{Auu^*A}{\langle Au, u\rangle}. \tag{2.2}$$

Thus, the Broyden family consists of all affine combinations of DFP and BFGS updates:

$$\mathrm{Broyd}_\phi(A, G, u) \overset{(2.1)}{=} \phi\mathrm{DFP}(A, G, u) + (1 - \phi)\mathrm{BFGS}(A, G, u). \tag{2.3}$$

The subclass of the Broyden family, corresponding to $\phi \in [0, 1]$, is known as the *convex Broyden class* (or the *restricted Broyden class* in some texts).

Our subsequent developments will be based on two properties of the convex Broyden class. The first property states that each update from this class preserves the bounds on the relative eigenvalues with respect to the target operator.

Lemma 2.1 *Let $A, G : \mathbb{E} \to \mathbb{E}^*$ be self-adjoint positive definite linear operators such that*

$$\frac{A}{\xi} \preceq G \preceq \eta A, \tag{2.4}$$

where $\xi, \eta \geq 1$. Then, for any $u \in \mathbb{E}$, and any $\phi \in [0, 1]$, we have

$$\frac{A}{\xi} \preceq \mathrm{Broyd}_\phi(A, G, u) \preceq \eta A. \tag{2.5}$$

Proof Suppose that $u \neq 0$ since otherwise the claim is trivial. In view of (2.3), it suffices to prove (2.5) only for the DFP and BFGS updates independently.

For the DFP update, we have

$$\mathrm{DFP}(A, G, u) \overset{(2.2)}{=} G - \frac{Auu^*G + Guu^*A}{\langle Au, u\rangle} + \left(\frac{\langle Gu, u\rangle}{\langle Au, u\rangle} + 1\right)\frac{Auu^*A}{\langle Au, u\rangle}$$
$$= \left(I_{\mathbb{E}^*} - \frac{Auu^*}{\langle Au, u\rangle}\right) G \left(I_{\mathbb{E}} - \frac{uu^*A}{\langle Au, u\rangle}\right) + \frac{Auu^*A}{\langle Au, u\rangle},$$

where $I_{\mathbb{E}}$, $I_{\mathbb{E}^*}$ are the identity operators in the spaces $\mathbb{E}$, $\mathbb{E}^*$ respectively. Hence,

$$\begin{aligned}\mathrm{DFP}(A, G, u) &\overset{(2.4)}{\preceq} \eta \left(I_{\mathbb{E}^*} - \frac{Auu^*}{\langle Au, u \rangle} \right) A \left(I_{\mathbb{E}} - \frac{uu^*A}{\langle Au, u \rangle} \right) + \frac{Auu^*A}{\langle Au, u \rangle} \\ &= \eta \left(A - \frac{Auu^*A}{\langle Au, u \rangle} \right) + \frac{Auu^*A}{\langle Au, u \rangle} = \eta A - (\eta - 1) \frac{Auu^*A}{\langle Au, u \rangle} \preceq \eta A, \\ \mathrm{DFP}(A, G, u) &\overset{(2.4)}{\succeq} \frac{1}{\xi} \left(I_{\mathbb{E}^*} - \frac{Auu^*}{\langle Au, u \rangle} \right) A \left(I_{\mathbb{E}} - \frac{uu^*A}{\langle Au, u \rangle} \right) + \frac{Auu^*A}{\langle Au, u \rangle} \\ &= \frac{1}{\xi} \left(A - \frac{Auu^*A}{\langle Au, u \rangle} \right) + \frac{Auu^*A}{\langle Au, u \rangle} = \frac{1}{\xi} A + \left(1 - \frac{1}{\xi} \right) \frac{Auu^*A}{\langle Au, u \rangle} \succeq \frac{1}{\xi} A.\end{aligned}$$

For the BFGS update, we apply Lemma 6.1 (see Appendix):

$$\begin{aligned}\mathrm{BFGS}(A, G, u) &\overset{(2.2)}{=} G - \frac{Guu^*G}{\langle Gu, u \rangle} + \frac{Auu^*A}{\langle Au, u \rangle} \overset{(2.4)}{\preceq} \eta \left(A - \frac{Auu^*A}{\langle Au, u \rangle} \right) \\ &\quad + \frac{Auu^*A}{\langle Au, u \rangle} \\ &= \eta A - (\eta - 1) \frac{Auu^*A}{\langle Au, u \rangle} \preceq \eta A, \\ \mathrm{BFGS}(A, G, u) &\overset{(2.2)}{=} G - \frac{Guu^*G}{\langle Gu, u \rangle} + \frac{Auu^*A}{\langle Au, u \rangle} \overset{(2.4)}{\succeq} \frac{1}{\xi} \left(A - \frac{Auu^*A}{\langle Au, u \rangle} \right) \\ &\quad + \frac{Auu^*A}{\langle Au, u \rangle} \\ &= \frac{1}{\xi} A + \left(1 - \frac{1}{\xi} \right) \frac{Auu^*A}{\langle Au, u \rangle} \succeq \frac{1}{\xi} A.\end{aligned}$$

The proof is finished □

Remark 2.1 Lemma 2.1 has first been established in [5] in a slightly stronger form and using a different argument. It was also shown there that one of the relations in (2.5) may no longer be valid if $\phi \in \mathbb{R} \setminus [0, 1]$.

The second property of the convex Broyden class, which we need, is related to the question of convergence of the approximations G to the target operator A. Note that without any restrictions on the choice of the update directions u, one cannot guarantee any convergence of G to A in the usual sense (see [19,31] for more details). However, for our goals it will be sufficient to show that, independently of the choice of u, it is still possible to ensure that G converges to A *along the update directions* u, and estimate the corresponding rate of convergence.

Let us define the following measure of the closeness of G to A along the direction u:

$$\theta(A, G, u) \overset{\text{def}}{=} \left[\frac{\langle (G - A) A^{-1} (G - A) u, u \rangle}{\langle G A^{-1} G u, u \rangle} \right]^{1/2}, \tag{2.6}$$

where, for the sake of convenience, we define $\theta(A, G, u) = 0$ if $u = 0$. Note that $\theta(A, G, u) = 0$ if and only if $Gu = Au$. Thus, our goal now is to establish some upper bounds on θ, which will help us to estimate the rate, at which this measure goes to zero. For this, we will study how certain potential functions change after one update from the convex Broyden class, and estimate this change from below by an appropriate monotonically increasing function of θ. We will consider two potential functions.

The first one is a simple trace potential function, that we will use only when we can guarantee that $A \preceq G$:

$$\sigma(A, G) \overset{\text{def}}{=} \langle A^{-1}, G - A\rangle \ \geq \ 0. \tag{2.7}$$

Lemma 2.2 *Let $A, G : \mathbb{E} \to \mathbb{E}^*$ be self-adjoint positive definite linear operators such that*

$$A \preceq G \preceq \eta A \tag{2.8}$$

for some $\eta \geq 1$. Then, for any $\phi \in [0, 1]$ and any $u \in \mathbb{E}$, we have

$$\sigma(A, G) - \sigma(A, \mathrm{Broyd}_\phi(A, G, u)) \geq \left(\phi \frac{1}{\eta} + 1 - \phi\right) \theta^2(A, G, u). \tag{2.9}$$

Proof We can assume that $u \neq 0$ since otherwise the claim is trivial. Denote $G_+ \overset{\text{def}}{=} \mathrm{Broyd}_\phi(A, G, u)$ and $\theta \overset{\text{def}}{=} \theta(A, G, u)$. Then,

$$\begin{aligned}
\sigma(A, G) - \sigma(A, G_+) &\overset{(2.7)}{=} \langle A^{-1}, G - G_+\rangle \\
&\overset{(2.1)}{=} 2\phi \frac{\langle Gu, u\rangle}{\langle Au, u\rangle} - \left[\phi \frac{\langle Gu, u\rangle}{\langle Au, u\rangle} + 1\right] + (1 - \phi) \frac{\langle GA^{-1}Gu, u\rangle}{\langle Gu, u\rangle} \\
&= \phi \frac{\langle Gu, u\rangle}{\langle Au, u\rangle} + (1 - \phi) \frac{\langle GA^{-1}Gu, u\rangle}{\langle Gu, u\rangle} - 1 \\
&= \phi \frac{\langle (G - A)u, u\rangle}{\langle Au, u\rangle} + (1 - \phi) \frac{\langle G(A^{-1} - G^{-1})Gu, u\rangle}{\langle Gu, u\rangle}.
\end{aligned} \tag{2.10}$$

Note that

$$0 \overset{(2.8)}{\preceq} G - A \overset{(2.8)}{\preceq} (\eta - 1)A \ \preceq \ \eta A.$$

Therefore[2],

$$(G - A)A^{-1}(G - A) \preceq \eta(G - A). \tag{2.11}$$

[2] This is evident when $G - A$ is non-degenerate. The general case then follows by continuity.

Consequently,

$$\begin{aligned}\frac{\langle (G-A)u,u\rangle}{\langle Au,u\rangle} &\overset{(2.11)}{\geq} \frac{1}{\eta}\frac{\langle (G-A)A^{-1}(G-A)u,u\rangle}{\langle Au,u\rangle}\\ &\overset{(2.8)}{\geq} \frac{1}{\eta}\frac{\langle (G-A)A^{-1}(G-A)u,u\rangle}{\langle GA^{-1}Gu,u\rangle} \overset{(2.6)}{=} \frac{1}{\eta}\theta^2. \end{aligned} \tag{2.12}$$

At the same time,

$$\begin{aligned}(G-A)A^{-1}(G-A) &= GA^{-1}G - 2G + A\\ &\overset{(2.8)}{\preceq} GA^{-1}G - G = G(A^{-1}-G^{-1})G.\end{aligned}$$

Hence,

$$\begin{aligned}\frac{\langle G(A^{-1}-G^{-1})Gu,u\rangle}{\langle Gu,u\rangle} &\geq \frac{\langle (G-A)A^{-1}(G-A)u,u\rangle}{\langle Gu,u\rangle}\\ &\overset{(2.8)}{\geq} \frac{\langle (G-A)A^{-1}(G-A)u,u\rangle}{\langle GA^{-1}Gu,u\rangle} \overset{(2.6)}{=} \theta^2. \end{aligned} \tag{2.13}$$

Substituting now (2.12) and (2.13) into (2.10), we obtain (2.9). □

The second potential function is more universal since we can work with it even if the condition $A \preceq G$ is violated. This function was first introduced in [29], and is defined as follows:

$$\psi(A,G) \overset{\text{def}}{=} \langle A^{-1}, G-A\rangle - \ln \mathrm{Det}(A^{-1},G). \tag{2.14}$$

In fact, ψ is nothing else but the Bregman divergence, generated by the strictly convex function $d(G) \overset{\text{def}}{=} -\ln \mathrm{Det}(B^{-1},G)$, defined on the set of self-adjoint positive definite linear operators from $\mathbb{E}$ to $\mathbb{E}^*$, where $B : \mathbb{E} \to \mathbb{E}^*$ is an arbitrary fixed self-adjoint positive definite linear operator. Indeed,

$$\begin{aligned}\psi(A,G) &\overset{(1.3)}{=} -\ln \mathrm{Det}(B^{-1},G) + \ln \mathrm{Det}(B^{-1},A) - \langle -A^{-1}, G-A\rangle\\ &= d(G) - d(A) - \langle \nabla d(A), G-A\rangle.\end{aligned}$$

Thus, $\psi(A,G) \geq 0$ and $\psi(A,G)=0$ if and only if $G=A$.

Let $\omega : (-1,+\infty) \to \mathbb{R}$ be the univariate function

$$\omega(t) \overset{\text{def}}{=} t - \ln(1+t) \;\geq\; 0. \tag{2.15}$$

Clearly, ω is a convex function, which is decreasing on $(-1, 0]$ and increasing on $[0, +\infty)$. Also, on the latter interval, it satisfies the following bounds (see [32, Lemma 5.1.5]):

$$\frac{t^2}{2(1+t)} \le \frac{t^2}{2\left(1+\frac{2}{3}t\right)} \le \omega(t) \le \frac{t^2}{2+t}, \qquad t \ge 0. \tag{2.16}$$

Thus, for large values of t, the function $\omega(t)$ is approximately linear in t, while for small values of t, it is quadratic.

There is a close relationship between ω and the potential function ψ. Indeed, if $\lambda_1, \ldots, \lambda_n \ge 0$ are the relative eigenvalues of G with respect to A, then

$$\psi(A, G) \overset{(2.14)}{=} \sum_{i=1}^{n}(\lambda_i - 1 - \ln \lambda_i) \overset{(2.15)}{=} \sum_{i=1}^{n}\omega(\lambda_i - 1).$$

We are going to use the function ω to estimate from below the change in the potential function ψ, which is achieved after one update from the convex Broyden class, via the closeness measure θ. However, first of all, we need an auxiliary lemma.

Lemma 2.3 *For any real $\alpha \ge \beta > 0$, we have*

$$\alpha - \ln \beta - 1 \ge \omega(\sqrt{\alpha\beta - 2\beta + 1}).$$

Proof Equivalently, we need to prove that

$$\alpha - 1 \ge \omega(\sqrt{\alpha\beta - 2\beta + 1}) + \ln \beta. \tag{2.17}$$

Let us show that the right-hand side of (2.17) is increasing in β. This is evident if $\alpha \ge 2$ because ω is increasing on $[0, +\infty)$, so suppose that $\alpha < 2$. Denote

$$t \overset{\text{def}}{=} \sqrt{\alpha\beta - 2\beta + 1} = \sqrt{1 - (2-\alpha)\beta} \in [0, 1). \tag{2.18}$$

Note that t is decreasing in β. Therefore, it suffices to prove that the right-hand side of (2.17) is decreasing in t. But

$$\begin{aligned} \omega(\sqrt{\alpha\beta - 2\beta + 1}) + \ln \beta &\overset{(2.18)}{=} \omega(t) + \ln \frac{1-t^2}{2-\alpha} \\ &= \omega(t) + \ln(1-t^2) - \ln(2-\alpha) \\ &\overset{(2.15)}{=} t - \ln(1+t) + \ln(1-t^2) - \ln(2-\alpha) \\ &= t + \ln(1-t) - \ln(2-\alpha) \\ &\overset{(2.15)}{=} -\omega(-t) - \ln(2-\alpha), \end{aligned}$$

which is indeed decreasing in t since ω is decreasing on $(-1, 0]$.

Thus, it suffices to prove (2.17) only in the boundary case $\beta = \alpha$:

$$\alpha - 1 \geq \omega(\sqrt{\alpha^2 - 2\alpha + 1}) + \ln \alpha \ = \ \omega(|\alpha - 1|) + \ln \alpha,$$

or, equivalently, in view of (2.15), that

$$\omega(\alpha - 1) \geq \omega(|\alpha - 1|)$$

For $\alpha \geq 1$, this is obvious, so suppose that $\alpha \leq 1$. It now remains to justify that

$$\omega(-t) \geq \omega(t), \tag{2.19}$$

for all $t \in [0, 1)$. But this easily follows by integration from the fact that

$$\frac{d}{dt}\omega(-t) = -\omega'(-t) \overset{(2.15)}{=} \frac{t}{1-t} \geq \frac{t}{1+t} \overset{(2.15)}{=} \omega'(t)$$

for all $t \in [0, 1)$. □

Now we are ready to prove the main result.

Lemma 2.4 *Let $A, G : \mathbb{E} \to \mathbb{E}^*$ be self-adjoint positive definite linear operators such that*

$$\frac{1}{\xi}A \preceq G \preceq \eta A \tag{2.20}$$

for some $\xi, \eta \geq 1$. Then, for any $\phi \in [0, 1]$ and any $u \in \mathbb{E}$, we have

$$\psi(A, G) - \psi(A, \text{Broyd}_\phi(A, G, u)) \geq \phi\, \omega\left(\frac{\theta(A, G, u)}{\xi^{3/2}\sqrt{\eta}}\right) + (1 - \phi)\omega\left(\frac{\theta(A, G, u)}{\xi}\right).$$

Proof Suppose that $u \neq 0$ since otherwise the claim is trivial. Let us denote $G_+ \overset{\text{def}}{=} \text{Broyd}_\phi(A, G, u)$ and $\theta \overset{\text{def}}{=} \theta(A, G, u)$. We already know that

$$\langle A^{-1}, G - G_+\rangle \overset{(2.10)}{=} \phi\frac{\langle Gu, u\rangle}{\langle Au, u\rangle} + (1 - \phi)\frac{\langle GA^{-1}Gu, u\rangle}{\langle Gu, u\rangle} - 1.$$

Applying now Lemma 6.2, we obtain

$$\text{Det}(G^{-1}, G_+) = \phi\frac{\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle} + (1 - \phi)\frac{\langle Au, u\rangle}{\langle Gu, u\rangle}.$$

Thus,

$$\psi(A, G) - \psi(A, G_+)$$

$$
\begin{aligned}
&\overset{(2.14)}{=} \langle A^{-1}, G - G_+\rangle + \ln \mathrm{Det}(A^{-1}, G_+) - \ln \mathrm{Det}(A^{-1}, G) \\
&\overset{(1.4)}{=} \langle A^{-1}, G - G_+\rangle + \ln \mathrm{Det}(G^{-1}, G_+) \\
&= \phi \frac{\langle Gu, u\rangle}{\langle Au, u\rangle} + (1-\phi)\frac{\langle GA^{-1}Gu, u\rangle}{\langle Gu, u\rangle} \\
&\quad -1 + \ln\left[\phi \frac{\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle} + (1-\phi)\frac{\langle Au, u\rangle}{\langle Gu, u\rangle}\right] \\
&\geq \phi\left[\frac{\langle Gu, u\rangle}{\langle Au, u\rangle} + \ln \frac{\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle}\right] + (1-\phi)\left[\frac{\langle GA^{-1}Gu, u\rangle}{\langle Gu, u\rangle} + \ln \frac{\langle Au, u\rangle}{\langle Gu, u\rangle}\right] - 1 \\
&= \phi\left[\frac{\langle Gu, u\rangle}{\langle Au, u\rangle} - \ln \frac{\langle Au, u\rangle}{\langle AG^{-1}Au, u\rangle} - 1\right] \\
&\quad +(1-\phi)\left[\frac{\langle GA^{-1}Gu, u\rangle}{\langle Gu, u\rangle} - \ln \frac{\langle Gu, u\rangle}{\langle Au, u\rangle} - 1\right], \qquad (2.21)
\end{aligned}
$$

where we have used the concavity of the logarithm.

Denote

$$
\begin{aligned}
\alpha_1 &\overset{\text{def}}{=} \frac{\langle Gu, u\rangle}{\langle Au, u\rangle}, \qquad \beta_1 \overset{\text{def}}{=} \frac{\langle Au, u\rangle}{\langle AG^{-1}Au, u\rangle}, \\
\alpha_0 &\overset{\text{def}}{=} \frac{\langle GA^{-1}Gu, u\rangle}{\langle Gu, u\rangle}, \qquad \beta_0 \overset{\text{def}}{=} \frac{\langle Gu, u\rangle}{\langle Au, u\rangle}. \qquad (2.22)
\end{aligned}
$$

Clearly, $\alpha_1 \geq \beta_1$ and $\alpha_0 \geq \beta_0$ by the Cauchy–Schwartz inequality. Also,

$$
\begin{aligned}
&\alpha_1\beta_1 - 2\beta_1 + 1 \overset{(2.22)}{=} \frac{\langle Gu, u\rangle}{\langle AG^{-1}Au, u\rangle} - 2\frac{\langle Au, u\rangle}{\langle AG^{-1}Au, u\rangle} + 1 \\
&\quad = \frac{\langle (G-A)G^{-1}(G-A)u, u\rangle}{\langle AG^{-1}Au, u\rangle} \\
&\quad \overset{(2.20)}{\geq} \frac{1}{\eta}\frac{\langle (G-A)A^{-1}(G-A)u, u\rangle}{\langle AG^{-1}Au, u\rangle} \overset{(2.20)}{\geq} \frac{1}{\xi^3\eta}\frac{\langle (G-A)A^{-1}(G-A)u, u\rangle}{\langle GA^{-1}Gu, u\rangle} \\
&\quad \overset{(2.6)}{=} \frac{\theta^2}{\xi^3\eta}, \\
&\alpha_0\beta_0 - 2\beta_0 + 1 \overset{(2.22)}{=} \frac{\langle GA^{-1}Gu, u\rangle}{\langle Au, u\rangle} - 2\frac{\langle Gu, u\rangle}{\langle Au, u\rangle} + 1 \\
&\quad = \frac{\langle (G-A)A^{-1}(G-A)u, u\rangle}{\langle Au, u\rangle} \\
&\quad \overset{(2.20)}{\geq} \frac{1}{\xi^2}\frac{\langle (G-A)A^{-1}(G-A)u, u\rangle}{\langle GA^{-1}Gu, u\rangle} \overset{(2.6)}{=} \frac{\theta^2}{\xi^2}.
\end{aligned}
$$

Therefore, by Lemma 2.3 and the fact that ω is increasing on $[0, +\infty)$, we have

$$\frac{\langle Gu, u\rangle}{\langle Au, u\rangle} - \ln \frac{\langle Au, u\rangle}{\langle AG^{-1}Au, u\rangle} - 1 \geq \omega\left(\left[\frac{\langle (G-A)G^{-1}(G-A)u, u\rangle}{\langle AG^{-1}Au, u\rangle}\right]^{1/2}\right)$$
$$\geq \omega\left(\frac{\theta}{\xi^{3/2}\sqrt{\eta}}\right),$$
$$\frac{\langle GA^{-1}Gu, u\rangle}{\langle Gu, u\rangle} - \ln \frac{\langle Gu, u\rangle}{\langle Au, u\rangle} - 1 \geq \omega\left(\left[\frac{\langle (G-A)A^{-1}(G-A)u, u\rangle}{\langle Au, u\rangle}\right]^{1/2}\right)$$
$$\geq \omega\left(\frac{\theta}{\xi}\right).$$

Combining these inequalities with (2.21), we obtain the claim. □

3 Unconstrained quadratic minimization

In this section, we study the classical quasi-Newton methods, based on the updating formulas from the convex Broyden class, as applied to minimizing the quadratic function

$$f(x) \overset{\text{def}}{=} \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle, \tag{3.1}$$

where $A : \mathbb{E} \to \mathbb{E}^*$ is a self-adjoint positive definite operator, and $b \in \mathbb{E}^*$.

Let $B : \mathbb{E} \to \mathbb{E}^*$ be a self-adjoint positive definite linear operator, that we will use to initialize our methods. Denote by $\mu > 0$ the strong convexity parameter of f, and by $L > 0$ the Lipschitz constant of the gradient of f, both measured with respect to B:

$$\mu B \preceq A \preceq LB. \tag{3.2}$$

Consider the following standard quasi-Newton scheme for minimizing (3.1). For the sake of simplicity, we assume that the constant L is available.

Initialization: Choose $x_0 \in \mathbb{E}$. Set $G_0 = LB$.	
For $k \geq 0$ **iterate:** 1. Update $x_{k+1} = x_k - G_k^{-1}\nabla f(x_k)$. 2. Set $u_k = x_{k+1} - x_k$ and choose $\phi_k \in [0, 1]$. 3. Compute $G_{k+1} = \text{Broyd}_{\phi_k}(A, G_k, u_k)$.	(3.3)

Remark 3.1 In an actual implementation of scheme (3.3), it is typical to store in memory and update in iterations the matrix $H_k \overset{\text{def}}{=} G_k^{-1}$ instead of G_k (or, alternatively, the Cholesky decomposition of G_k). This allows one to compute $G_{k+1}^{-1}\nabla f(x_k)$ in $O(n^2)$ operations. Note that, due to a low-rank structure of the update (2.1), H_k can be updated into H_{k+1} also in $O(n^2)$ operations (for specific formulas, see e.g. [8, Section 8]).

To measure the convergence rate of scheme (3.3), we look at the norm of the gradient, measured with respect to A:

$$\lambda_f(x) \overset{\text{def}}{=} \|\nabla f(x)\|_A^* \overset{(1.1)}{=} \langle \nabla f(x), A^{-1}\nabla f(x)\rangle^{1/2}. \tag{3.4}$$

The following lemma shows that the measure $\theta(A, G_k, u_k)$, that we introduced in (2.6) to measure the closeness of G_k to A along the direction u_k, is directly related to the progress of one step of the scheme (3.3). Note that it is important here that the updating direction $u_k = x_{k+1} - x_k$ is chosen as the difference of the iterates, and, for other choices of u_k, this result is no longer true.

Lemma 3.1 *In scheme* (3.3)*, for all* $k \geq 0$*, we have*

$$\lambda_f(x_{k+1}) = \theta(A, G_k, u_k)\lambda_f(x_k). \tag{3.5}$$

Proof Indeed,

$$\nabla f(x_{k+1}) \overset{(3.1)}{=} \nabla f(x_k) + A(x_{k+1} - x_k) \overset{(3.3)}{=} -G_k u_k + A u_k = -(G_k - A)u_k.$$

Hence, denoting $\theta_k \overset{\text{def}}{=} \theta(A, G_k, u_k)$, we get

$$\begin{aligned}\lambda_f(x_{k+1}) &\overset{(3.4)}{=} \langle (G_k - A)A^{-1}(G_k - A)u_k, u_k\rangle^{1/2} \overset{(2.6)}{=} \theta_k \langle G_k A^{-1} G_k u_k, u_k\rangle^{1/2} \\ &\overset{(3.3)}{=} \theta_k \langle \nabla f(x_k), A^{-1}\nabla f(x_k)\rangle^{1/2} \overset{(3.4)}{=} \theta_k \lambda_f(x_k).\end{aligned}$$

The proof is finished □

Let us show that the scheme (3.3) has global linear convergence, and that the corresponding rate is at least as good as that of the standard gradient method.

Theorem 3.1 *In scheme* (3.3)*, for all* $k \geq 0$*, we have*

$$A \preceq G_k \preceq \frac{L}{\mu} A, \tag{3.6}$$

and

$$\lambda_f(x_k) \leq \left(1 - \frac{\mu}{L}\right)^k \lambda_f(x_0). \tag{3.7}$$

Proof For $k = 0$, (3.6) follows from the fact that $G_0 = LB$ and (3.2). For all other $k \geq 1$, it follows by induction using Lemma 2.1.

Thus, we have

$$0 \overset{(3.6)}{\preceq} A^{-1} - G_k^{-1} \overset{(3.6)}{\preceq} \left(1 - \frac{\mu}{L}\right) A^{-1}. \tag{3.8}$$

Therefore,

$$\begin{aligned}(G_k - A)A^{-1}(G_k - A) &= G_k(A^{-1} - G_k^{-1})A(A^{-1} - G_k^{-1})G_k \\ &\preceq \left(1 - \frac{\mu}{L}\right)^2 G_k A^{-1} G_k,\end{aligned}$$

and so

$$\theta(A, G_k, u_k) \overset{(2.6)}{\leq} 1 - \frac{\mu}{L}.$$

Applying now Lemma 3.5, we obtain (3.7). □

Now, let us establish the superlinear convergence of the scheme (3.3). First, we do this by working with the trace potential function σ, defined by (2.7). Note that this is possible since $A \preceq G_k$ in view of (3.6).

Theorem 3.2 *In scheme* (3.3)*, for all* $k \geq 1$*, we have*

$$\lambda_f(x_k) \leq \frac{1}{\prod_{i=0}^{k-1} \left(\phi_i \frac{\mu}{L} + 1 - \phi_i\right)^{1/2}} \left(\frac{nL}{\mu k}\right)^{k/2} \lambda_f(x_0). \tag{3.9}$$

Proof Denote $\sigma_i \overset{\text{def}}{=} \sigma(A, G_i)$, $\theta_i \overset{\text{def}}{=} \theta(A, G_i, u_i)$, and $p_i \overset{\text{def}}{=} \phi_i \frac{\mu}{L} + 1 - \phi_i$ for any $i \geq 0$. Let $k \geq 1$ be arbitrary. From (3.6) and Lemma 2.2, it follows that

$$\sigma_i - \sigma_{i+1} \geq p_i \theta_i^2$$

for all $0 \leq i \leq k - 1$. Summing up these inequalities, we obtain

$$\begin{aligned}\sum_{i=0}^{k-1} p_i \theta_i^2 &\leq \sigma_0 - \sigma_k \overset{(2.7)}{\leq} \sigma_0 \overset{(3.3)}{=} \sigma(A, LB) \overset{(2.7)}{=} \langle A^{-1}, LB - A\rangle \\ &\overset{(3.2)}{\leq} \langle A^{-1}, \frac{L}{\mu}A - A\rangle \overset{(1.3)}{=} n\left(\frac{L}{\mu} - 1\right) \leq \frac{nL}{\mu}.\end{aligned} \tag{3.10}$$

Hence, by Lemma 3.1 and the arithmetic-geometric mean inequality,

$$\lambda_f(x_k) = \lambda_f(x_0)\prod_{i=0}^{k-1}\theta_i \;=\; \frac{1}{\prod_{i=0}^{k-1} p_i^{1/2}}\left[\prod_{i=0}^{k-1} p_i\theta_i^2\right]^{1/2}\lambda_f(x_0)$$

$$\leq \frac{1}{\prod_{i=0}^{k-1} p_i^{1/2}}\left(\frac{1}{k}\sum_{i=0}^{k-1} p_i\theta_i^2\right)^{k/2}\lambda_f(x_0) \overset{(3.10)}{\leq} \frac{1}{\prod_{i=0}^{k-1} p_i^{1/2}}\left(\frac{nL}{\mu k}\right)^{k/2}\lambda_f(x_0).$$

The proof is finished □

Remark 3.2 As can be seen from (3.10), the factor $\frac{nL}{\mu}$ in the efficiency estimate (3.9) can be improved up to $\langle A^{-1}, LB - A\rangle = \sum_{i=1}^{n}(\frac{L}{\lambda_i} - 1)$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A relative to B. This improved factor can be significantly smaller than the original one if the majority of the eigenvalues λ_i are much larger than μ. However, for the sake of simplicity, we prefer to work directly with constants n, L and μ. This corresponds to the worst-case analysis. The same remark applies to all other theorems on superlinear convergence, that will follow.

Let us discuss the efficiency estimate (3.9). Note that its *maximal* value over all $\phi_i \in [0, 1]$ is achieved at $\phi_i = 1$ for all $0 \leq i \leq k-1$. This corresponds to the *DFP method*. In this case, the efficiency estimate (3.9) looks as follows:

$$\lambda_f(x_k) \leq \left(\frac{nL^2}{\mu^2 k}\right)^{k/2}\lambda_f(x_0).$$

Hence, the moment, when the superlinear convergence starts, can be described as follows:

$$\frac{nL^2}{\mu^2 k} \leq 1 \qquad \Longleftrightarrow \qquad k \geq \frac{nL^2}{\mu^2}.$$

In contrast, the *minimal* value of the efficiency estimate (3.9) over all $\phi_i \in [0, 1]$ is achieved at $\phi_i = 0$ for all $0 \leq i \leq k-1$. This corresponds to the *BFGS method*. In this case, the efficiency estimate (3.9) becomes

$$\lambda_f(x_k) \leq \left(\frac{nL}{\mu k}\right)^{k/2}\lambda_f(x_0), \tag{3.11}$$

and the moment, when the superlinear convergence begins, can be described as follows:

$$\frac{nL}{\mu k} \leq 1 \qquad \Longleftrightarrow \qquad k \geq \frac{nL}{\mu}.$$

Thus, we see that, compared to DFP, the superlinear convergence of BFGS starts in $\frac{L}{\mu}$ times earlier, and its rate is much faster.

Let us present for the scheme (3.3) another justification of the superlinear convergence rate in the form (3.9). For this, instead of σ, we will work with the potential function ω, defined by (2.15). The advantage of this analysis is that it is extendable onto general nonlinear functions.

Theorem 3.3 *In scheme* (3.3)*, for all* $k \geq 1$*, we have*

$$\lambda_f(x_k) \leq \frac{1}{\prod_{i=0}^{k-1}\left(\phi_i \frac{\mu}{L} + 1 - \phi_i\right)^{1/2}} \left(\frac{4nL}{\mu k}\right)^{k/2} \lambda_f(x_0). \tag{3.12}$$

Proof Denote $\theta_i \overset{\text{def}}{=} \theta(A, G_i, u_i)$, $\psi_i \overset{\text{def}}{=} \psi(A, G_i)$, and $p_i \overset{\text{def}}{=} \phi_i \frac{\mu}{L} + 1 - \phi_i$ for any $i \geq 0$. Let $k \geq 1$ and $0 \leq i \leq k-1$ be arbitrary. In view of (3.6) and Lemma 2.4, we have

$$\psi_i - \psi_{i+1} \overset{(2.21)}{\geq} \phi_i \omega\left(\sqrt{\frac{\mu}{L}}\theta_i\right) + (1-\phi_i)\omega(\theta_i). \tag{3.13}$$

Note that $\theta_i \leq 1$. Indeed, if $u_i = 0$, then $\theta_i = 0$ by definition. Otherwise,

$$\theta_i^2 \overset{(2.6)}{=} 1 - \frac{\langle (2G_i - A)u_i, u_i\rangle}{\langle G_i A^{-1} G_i u_i, u_i\rangle} \overset{(3.6)}{\leq} 1.$$

Therefore,

$$\omega\left(\sqrt{\frac{\mu}{L}}\theta_i\right) \overset{(2.16)}{\geq} \frac{\mu}{L}\frac{\theta_i^2}{2\left(1+\sqrt{\frac{\mu}{L}}\theta_i\right)} \geq \frac{\mu}{L}\frac{\theta_i^2}{4}, \qquad \omega(\theta_i) \overset{(2.16)}{\geq} \frac{\theta_i^2}{2(1+\theta_i)} \geq \frac{\theta_i^2}{4},$$

and we conclude that

$$\psi_i - \psi_{i+1} \overset{(3.13)}{\geq} \frac{1}{4} p_i \theta_i^2.$$

Summing this inequality and using the fact that $\psi_k \geq 0$, we obtain

$$\begin{aligned}
\frac{1}{4}\sum_{i=0}^{k-1} p_i \theta_i^2 \leq \psi_0 - \psi_k \;\leq\; \psi_0 \;&\overset{(3.3)}{=}\; \psi(A, LB) \\
&\overset{(2.14)}{=} \langle A^{-1}, LB - A\rangle - \ln \operatorname{Det}(A^{-1}, LB) \\
&\overset{(3.2)}{\leq} \langle A^{-1}, \frac{L}{\mu}A - A\rangle - \ln \operatorname{Det}(A^{-1}, \frac{L}{\mu}A) \\
&\overset{(1.3)}{=} n\left(\frac{L}{\mu} - 1 - \ln\frac{L}{\mu}\right) \;\leq\; \frac{nL}{\mu}.
\end{aligned} \tag{3.14}$$

Hence, by Lemma 3.1 and the arithmetic-geometric mean inequality,

$$\lambda_f(x_k) = \lambda_f(x_0)\prod_{i=0}^{k-1}\theta_i = \frac{1}{\prod_{i=0}^{k-1} p_i^{1/2}}\left[\prod_{i=0}^{k-1} p_i\theta_i^2\right]^{1/2}\lambda_f(x_0)$$
$$\leq \frac{1}{\prod_{i=0}^{k-1} p_i^{1/2}}\left(\frac{1}{k}\sum_{i=0}^{k-1} p_i\theta_i^2\right)^{k/2}\lambda_f(x_0)$$
$$\overset{(3.14)}{\leq} \frac{1}{\prod_{i=0}^{k-1} p_i^{1/2}}\left(\frac{4nL}{\mu k}\right)^{k/2}\lambda_f(x_0).$$

The proof is finished □

Comparing our new efficiency estimate (3.12) with the previous one (3.9), we see that they differ only in a constant. Thus, for the quadratic function, we do not gain anything by working with the potential function ω instead of σ. Nevertheless, our second proof is more universal, and, in contrast to the first one, can be generalized onto general nonlinear functions, as we will see in the next section.

4 Minimization of general functions

Consider now a general unconstrained minimization problem:

$$\min_{x\in\mathbb{E}} f(x), \tag{4.1}$$

where $f : \mathbb{E} \to \mathbb{R}$ is a twice differentiable function with positive definite Hessian.

To write down the standard quasi-Newton scheme for (4.1), we fix some self-adjoint positive definite linear operator $B : \mathbb{E} \to \mathbb{E}^*$ and a constant $L > 0$, that we use to define the initial Hessian approximation.

Initialization: Choose $x_0 \in \mathbb{E}$. Set $G_0 = LB$.	
For $k \geq 0$ **iterate:** 1. Update $x_{k+1} = x_k - G_k^{-1}\nabla f(x_k)$. 2. Set $u_k = x_{k+1} - x_k$ and choose $\phi_k \in [0, 1]$. 3. Denote $J_k = \int_0^1 \nabla^2 f(x_k + tu_k)dt$. 4. Set $G_{k+1} = \mathrm{Broyd}_{\phi_k}(J_k, G_k, u_k)$.	(4.2)

Remark 4.1 Similarly to Remark 3.1, when implementing scheme (4.2), it is common to work directly with the inverse $H_k \stackrel{\text{def}}{=} G_k^{-1}$ instead of G_k. Also note that it is not necessary to compute J_k explicitly. Indeed, for implementing the Hessian approximation update at Step 4 (or the corresponding update for its inverse), one only needs the product

$$J_k u_k = \nabla f(x_{k+1}) - \nabla f(x_k),$$

which is just the difference of the successive gradients.

In what follows, we make the following assumptions about the problem (4.1). First, we assume that, with respect to the operator B, the objective function f is *strongly convex* with parameter $\mu > 0$ and its gradient is *Lipschitz continuous* with constant L, i.e.

$$\mu B \preceq \nabla^2 f(x) \preceq LB \tag{4.3}$$

for all $x \in \mathbb{E}$. Second, we assume that the objective function f is *strongly self-concordant* with some constant $M \geq 0$, i.e.

$$\nabla^2 f(y) - \nabla^2 f(x) \preceq M\|y - x\|_z \nabla^2 f(w) \tag{4.4}$$

for all $x, y, z, w \in \mathbb{E}$. The class of strongly self-concordant functions was recently introduced in [31], and contains at least all strongly convex functions with Lipschitz continuous Hessian (see [31, Example 4.1]). It gives us the the following convenient relations between the Hessians of the objective function:

Lemma 4.1 *(see [31, Lemma 4.1]) Let $x, y \in \mathbb{E}$, and let $r \stackrel{\text{def}}{=} \|y - x\|_x$. Then,*

$$\frac{\nabla^2 f(x)}{1 + Mr} \preceq \nabla^2 f(y) \preceq (1 + Mr)\nabla^2 f(x). \tag{4.5}$$

Also, for $J \stackrel{\text{def}}{=} \int_0^1 \nabla^2 f(x + t(y - x))dt$, we have

$$\frac{\nabla^2 f(x)}{1 + \frac{Mr}{2}} \preceq J \preceq \left(1 + \frac{Mr}{2}\right) \nabla^2 f(x), \tag{4.6}$$

$$\frac{\nabla^2 f(y)}{1 + \frac{Mr}{2}} \preceq J \preceq \left(1 + \frac{Mr}{2}\right) \nabla^2 f(y). \tag{4.7}$$

As a particular example of a nonquadratic function, satisfying assumptions (4.3), (4.4), one can consider the regularized log-sum-exp function, defined by $f(x) \stackrel{\text{def}}{=} \ln(\sum_{i=1}^m e^{\langle a_i, x\rangle + b_i}) + \frac{\mu}{2}\|x\|^2$, where $a_i \in \mathbb{E}^*$, $b_i \in \mathbb{R}$ for $i = 1, \ldots, m$, and $\mu > 0$, $\|x\| \stackrel{\text{def}}{=} \langle Bx, x\rangle^{1/2}$.

Remark 4.2 Since we are interested in *local* convergence, it is possible to relax our assumptions by requiring that (4.3), (4.4) hold only in some neighborhood of a minimizer x^*. For this, it suffices to assume that the Hessian of f is Lipschitz continuous in this neighborhood with $\nabla^2 f(x^*)$ being positive definite. These are exactly the standard assumptions, used in [8] and many other works, studying local convergence of quasi-Newton methods. However, to avoid excessive technicalities, we do not do this.

Let us now analyze the process (4.2). For measuring its convergence, we look at the local norm of the gradient:

$$\lambda_f(x) \stackrel{\text{def}}{=} \|\nabla f(x)\|_x^* \stackrel{(1.1)}{=} \langle \nabla f(x), \nabla^2 f(x)^{-1} \nabla f(x)\rangle^{1/2}, \qquad x \in \mathbb{E}. \tag{4.8}$$

First, let us estimate the progress of one step of the scheme (4.2). Recall that $\theta(J_k, G_k, u_k)$ is the measure of closeness of G_k to J_k along the direction u_k (see (2.6)).

Lemma 4.2 *In scheme* (4.2)*, for all $k \geq 0$ and $r_k \stackrel{\text{def}}{=} \|u_k\|_{x_k}$, we have*

$$\lambda_f(x_{k+1}) \leq \left(1 + \frac{Mr_k}{2}\right) \theta(J_k, G_k, u_k) \lambda_f(x_k).$$

Proof Denote $\theta_k \stackrel{\text{def}}{=} \theta(J_k, G_k, u_k)$. In view of Taylor's formula,

$$\nabla f(x_{k+1}) = \nabla f(x_k) + J_k(x_{k+1} - x_k) \stackrel{(4.2)}{=} -(G_k - J_k)u_k. \tag{4.9}$$

Therefore,

$$\begin{aligned}
\lambda_f(x_{k+1}) &\stackrel{(4.8)}{=} \langle \nabla f(x_{k+1}), \nabla^2 f(x_{k+1})^{-1} \nabla f(x_{k+1})\rangle^{1/2} \\
&\stackrel{(4.7)}{\leq} \sqrt{1 + \frac{Mr_k}{2}} \langle \nabla f(x_{k+1}), J_k^{-1} \nabla f(x_{k+1})\rangle^{1/2} \\
&\stackrel{(4.9)}{=} \sqrt{1 + \frac{Mr_k}{2}} \langle (G_k - J_k) J_k^{-1} (G_k - J_k) u_k, u_k\rangle^{1/2} \\
&\stackrel{(2.6)}{=} \sqrt{1 + \frac{Mr_k}{2}} \theta_k \langle G_k J_k^{-1} G_k u_k, u_k\rangle^{1/2} \\
&\stackrel{(4.2)}{=} \sqrt{1 + \frac{Mr_k}{2}} \theta_k \langle \nabla f(x_k), J_k^{-1} \nabla f(x_k)\rangle^{1/2} \\
&\stackrel{(4.6)}{\leq} \left(1 + \frac{Mr_k}{2}\right) \theta_k \langle \nabla f(x_k), \nabla^2 f(x_k)^{-1} \nabla f(x_k)\rangle^{1/2} \\
&\stackrel{(4.8)}{=} \left(1 + \frac{Mr_k}{2}\right) \theta_k \lambda_f(x_k).
\end{aligned}$$

The proof is finished □

Our next result states that, if the starting point in scheme (4.2) is chosen sufficiently close to the solution, then the relative eigenvalues of the Hessian approximations G_k with respect to both the Hessians $\nabla^2 f(x_k)$ and the integral Hessians J_k are always located between 1 and $\frac{L}{\mu}$, up to some small numerical constant. As a consequence, the process (4.2) has at least the linear convergence rate of the gradient method.

Theorem 4.1 *Suppose that, in scheme (4.2),*

$$M\lambda_f(x_0) \le \frac{\ln \frac{3}{2}}{4} \frac{\mu}{L}. \tag{4.10}$$

Then, for all $k \ge 0$, we have

$$\frac{1}{\xi_k}\nabla^2 f(x_k) \preceq G_k \preceq \xi_k \frac{L}{\mu}\nabla^2 f(x_k), \tag{4.11}$$

$$\frac{1}{\xi_k'} J_k \preceq G_k \preceq \xi_k' \frac{L}{\mu} J_k, \tag{4.12}$$

$$\xi_k \lambda_f(x_k) \le \left(1 - \frac{\mu}{2L}\right)^k \lambda_f(x_0), \tag{4.13}$$

where[3]

$$\xi_k \overset{\text{def}}{=} e^{M\sum_{i=0}^{k-1} r_i} \le \left(1 + \frac{Mr_k}{2}\right) e^{M\sum_{i=0}^{k-1} r_i} \overset{\text{def}}{=} \xi_k' \le \sqrt{\frac{3}{2}}, \tag{4.14}$$

and $r_i \overset{\text{def}}{=} \|u_i\|_{x_i}$ for any $i \ge 0$.

Proof Note that $\xi_0 = 1$ and $G_0 = LB$. Therefore, for $k = 0$, both (4.11), (4.13) are satisfied. Indeed, the first one reads $\nabla^2 f(x_0) \preceq LB \preceq \frac{L}{\mu}\nabla^2 f(x_0)$ and follows from (4.3), while the second one reads $\lambda_f(x_0) \le \lambda_f(x_0)$ and is obviously true.

Now assume that $k \ge 0$, and that (4.11), (4.13) have already been proved for all $0 \le k' \le k$. Combining (4.11) with (4.6), using the definition of ξ_k', we obtain (4.12). Further, denote $\lambda_i \overset{\text{def}}{=} \lambda_f(x_i)$ for $0 \le i \le k$. Note that

$$\begin{aligned} r_k &\overset{(4.2)}{=} \|G_k^{-1}\nabla f(x_k)\|_{x_k} \overset{(1.1)}{=} \langle \nabla f(x_k), G_k^{-1}\nabla^2 f(x_k) G_k^{-1}\nabla f(x_k)\rangle^{1/2} \\ &\overset{(4.11)}{\le} \xi_k \langle \nabla f(x_k), \nabla^2 f(x_k)^{-1}\nabla f(x_k)\rangle^{1/2} \overset{(4.8)}{=} \xi_k \lambda_k. \end{aligned} \tag{4.15}$$

Therefore,

$$M\sum_{i=0}^{k} r_i \overset{(4.15)}{\le} M\sum_{i=0}^{k} \xi_i \lambda_i \overset{(4.13)}{\le} M\lambda_0 \sum_{i=0}^{k}\left(1 - \frac{\mu}{2L}\right)^i$$

[3] Here we follow the standard convention that the sum over the empty set is defined as zero. Thus, $\xi_0 = e^0 = 1$.

$$\leq \frac{2L}{\mu} M\lambda_0 \stackrel{(4.10)}{\leq} \frac{\ln \frac{3}{2}}{2}. \tag{4.16}$$

Consequently, by the definition of ξ_k and ξ_k',

$$\xi_k \leq \xi_k' \leq e^{\frac{Mr_k}{2}} e^{M\sum_{i=0}^{k-1} r_i} \leq e^{M\sum_{i=0}^{k} r_i} \stackrel{(4.16)}{\leq} \sqrt{\frac{3}{2}}.$$

Thus, (4.12), (4.14) are now proved. To finish the proof by induction, it remains to prove (4.11), (4.13) for $k' = k+1$.

We start with (4.11). Applying Lemma 2.1, using (4.12), we obtain

$$\frac{1}{\xi_k'} J_k \preceq G_{k+1} \preceq \xi_k' \frac{L}{\mu} J_k. \tag{4.17}$$

Consequently,

$$\begin{aligned} G_{k+1} &\stackrel{(4.7)}{\preceq} \left(1 + \frac{Mr_k}{2}\right) \xi_k' \frac{L}{\mu} \nabla^2 f(x_{k+1}) \stackrel{(4.14)}{=} \left(1 + \frac{Mr_k}{2}\right)^2 \xi_k \frac{L}{\mu} \nabla f(x_{k+1}) \\ &\preceq e^{Mr_k} \xi_k \frac{L}{\mu} \nabla^2 f(x_{k+1}) \stackrel{(4.14)}{=} \xi_{k+1} \frac{L}{\mu} \nabla^2 f(x_{k+1}), \end{aligned}$$

and

$$G_{k+1} \stackrel{(4.7)}{\succeq} \frac{\nabla^2 f(x_{k+1})}{\left(1 + \frac{Mr_k}{2}\right)\xi_k'} \stackrel{(4.14)}{=} \frac{\nabla^2 f(x_{k+1})}{\left(1 + \frac{Mr_k}{2}\right)^2 \xi_k} \succeq \frac{\nabla^2 f(x_{k+1})}{e^{Mr_k} \cdot \xi_k} \stackrel{(4.14)}{=} \frac{\nabla^2 f(x_{k+1})}{\xi_{k+1}}.$$

Thus, (4.11) is proved for $k' = k+1$.

It remains to prove (4.13) for $k' = k+1$. By Lemma 4.2,

$$\lambda_{k+1} \leq \left(1 + \frac{Mr_k}{2}\right) \theta_k \lambda_k, \tag{4.18}$$

where $\theta_k \stackrel{\text{def}}{=} \theta(J_k, G_k, u_k)$. Note that

$$-\left(1 - \frac{\mu}{\xi_k' L}\right) J_k^{-1} \stackrel{(4.12)}{\preceq} G_k^{-1} - J_k^{-1} \stackrel{(4.12)}{\preceq} (\xi_k' - 1) J_k^{-1}.$$

Hence,

$$(J_k^{-1} - G_k^{-1}) J_k (J_k^{-1} - G_k^{-1}) \preceq \rho_k^2 J_k^{-1},$$

where

$$\rho_k \stackrel{\text{def}}{=} \max\left\{1 - \frac{\mu}{\xi_k' L}, \xi_k' - 1\right\} \stackrel{(4.14)}{\geq} 0. \tag{4.19}$$

Therefore,

$$\theta_k^2 \overset{(2.6)}{=} \frac{\langle (J_k - G_k) J_k^{-1} (J_k - G_k) u_k, u_k \rangle}{\langle J_k G_k^{-1} J_k u_k, u_k \rangle}$$
$$= \frac{\langle G_k (J_k^{-1} - G_k^{-1}) J_k (J_k^{-1} - G_k^{-1}) G_k u_k, u_k \rangle}{\langle J_k G_k^{-1} J_k u_k, u_k \rangle} \leq \rho_k^2.$$

Thus,

$$\lambda_{k+1} \overset{(4.18)}{\leq} \left(1 + \frac{M r_k}{2}\right) \rho_k \lambda_k.$$

Consequently,

$$\xi_{k+1}\lambda_{k+1} \leq \xi_{k+1}\left(1 + \frac{M r_k}{2}\right) \rho_k \lambda_k \overset{(4.14)}{=} e^{M r_k}\left(1 + \frac{M r_k}{2}\right) \rho_k \xi_k \lambda_k$$
$$\leq e^{\frac{3 M r_k}{2}} \rho_k \xi_k \lambda_k \overset{(4.13)}{\leq} e^{\frac{3 M r_k}{2}} \rho_k \left(1 - \frac{\mu}{2L}\right)^k \lambda_0.$$

It remains to show that

$$e^{\frac{3 M r_k}{2}} \rho_k \leq 1 - \frac{\mu}{2L}. \tag{4.20}$$

Note that

$$\zeta_k \overset{\text{def}}{=} \frac{3 M r_k}{2} \overset{(4.15)}{\leq} \frac{3 M \xi_k \lambda_k}{2} \overset{(4.13)}{\leq} \frac{3 M \lambda_0}{2}$$
$$\overset{(4.10)}{\leq} \frac{3 \ln \frac{3}{2}}{8} \frac{\mu}{L} \leq \frac{3\mu}{16L} \leq \frac{\mu}{5L} \leq \frac{1}{5}. \tag{4.21}$$

Hence,

$$e^{\zeta_k} \leq \sum_{i=0}^{\infty} \zeta_k^i = 1 + \zeta_k \sum_{i=0}^{\infty} \zeta_k^i \overset{(4.21)}{\leq} 1 + \zeta_k \sum_{i=0}^{\infty} \left(\frac{1}{5}\right)^i$$
$$= 1 + \frac{5\zeta_k}{4} \overset{(4.21)}{\leq} 1 + \frac{\mu}{4L}. \tag{4.22}$$

Also,

$$\xi_k' \overset{(4.14)}{\leq} \sqrt{\frac{3}{2}} \leq \frac{4}{3}. \tag{4.23}$$

Combining (4.22) and (4.23), we obtain

$$e^{\frac{3Mr_k}{2}}\left(1-\frac{\mu}{\xi_k' L}\right) \leq \left(1+\frac{\mu}{4L}\right)\left(1-\frac{3\mu}{4L}\right) \leq 1-\left(\frac{3}{4}-\frac{1}{4}\right)\frac{\mu}{L} = 1-\frac{\mu}{2L},$$

and

$$e^{\frac{3Mr_k}{2}}(\xi_k'-1) \leq \left(1+\frac{1}{4}\right)\left(\sqrt{\frac{3}{2}}-1\right) = \frac{\frac{5}{4}\cdot\frac{1}{2}}{\sqrt{\frac{3}{2}}+1} \leq \frac{5}{16} \leq \frac{1}{2} \leq 1-\frac{\mu}{2L}.$$

Thus,

$$e^{\frac{3Mr_k}{2}}\rho_k \overset{(4.19)}{=} e^{\frac{3Mr_k}{2}} \max\left\{1-\frac{\mu}{\xi_k' L},\, \xi_k'-1\right\} \leq 1-\frac{\mu}{2L},$$

and (4.20) follows. □

Now we are ready to prove the main result of this section on the superlinear convergence of the scheme (4.2). In contrast to the quadratic case, now we cannot use the proof, based on the trace potential function σ, defined by (2.7), because we cannot longer guarantee that $J_k \preceq G_k$. However, the proof, based on the potential function ψ, defined by (2.14), still works.

Theorem 4.2 *Suppose that the initial point x_0 in scheme (4.2) is chosen sufficiently close to the solution, as specified by (4.10). Then, for all $k \geq 1$, we have*

$$\lambda_f(x_k) \leq \frac{1}{\prod_{i=0}^{k-1}\left(\phi_i \frac{\mu}{L}+1-\phi_i\right)^{1/2}}\left(\frac{11nL}{\mu k}\right)^{k/2}\lambda_f(x_0).$$

Proof Denote $r_i \overset{\text{def}}{=} \|u_i\|_{x_i}$, $\theta_i \overset{\text{def}}{=} \theta(J_i, G_i, u_i)$, $\psi_i \overset{\text{def}}{=} \psi(J_i, G_i)$, $\tilde{\psi}_{i+1} \overset{\text{def}}{=} \psi(J_i, G_{i+1})$, and $p_i \overset{\text{def}}{=} \phi_i \frac{\mu}{L}+1-\phi_i$ for any $i \geq 0$. Let $k \geq 1$ and $0 \leq i \leq k-1$ be arbitrary. By (4.12), (4.14) and Lemma 2.4, we have

$$\psi_i - \tilde{\psi}_{i+1} \geq \phi_i \omega\left(\frac{2}{3}\sqrt{\frac{\mu}{L}}\theta_i\right) + (1-\phi_i)\omega\left(\sqrt{\frac{2}{3}}\theta_i\right). \tag{4.24}$$

Moreover, since

$$\theta_i^2 \overset{(2.6)}{=} \frac{\langle (G_i-J_i)J_i^{-1}(G_i-J_i)u_i, u_i\rangle}{\langle G_i J_i^{-1} G_i u_i, u_i\rangle} = 1-\frac{\langle (2G_i-J_i)u_i, u_i\rangle}{\langle G_i J_i^{-1} G_i u_i, u_i\rangle} \overset{(4.12)}{\leq} 1,$$

we also have

$$\omega\left(\frac{2}{3}\sqrt{\frac{\mu}{L}}\theta_i\right) \overset{(2.16)}{\geq} \frac{\frac{4}{9}\frac{\mu}{L}\theta_i^2}{2\left(1+\frac{2}{3}\cdot\frac{2}{3}\sqrt{\frac{\mu}{L}}\theta_i\right)} \geq \frac{\frac{4}{9}}{2(1+\frac{4}{9})}\frac{\mu}{L}\theta_i^2 = \frac{2}{13}\frac{\mu}{L}\theta_i^2 \geq \frac{1}{7}\frac{\mu}{L}\theta_i^2,$$

$$\omega\left(\sqrt{\frac{2}{3}}\theta_i\right) \overset{(2.16)}{\geq} \frac{\frac{2}{3}\theta_i^2}{2\left(1+\sqrt{\frac{2}{3}}\theta_i\right)} \geq \frac{\frac{2}{3}}{2\left(1+\sqrt{\frac{2}{3}}\right)} \geq \frac{\frac{2}{3}}{4}\theta_i^2 = \frac{1}{6}\theta_i^2 \geq \frac{1}{7}\theta_i^2.$$

Thus,

$$\frac{1}{7}p_i\theta_i^2 \overset{(4.24)}{\leq} \psi_i - \tilde{\psi}_{i+1} = \psi_i - \psi_{i+1} + \Delta_i, \tag{4.25}$$

where

$$\Delta_i \overset{\text{def}}{=} \psi_{i+1} - \tilde{\psi}_{i+1} \overset{(2.14)}{=} \langle J_{i+1}^{-1} - J_i^{-1}, G_{i+1}\rangle + \ln \operatorname{Det}(J_i^{-1}, J_{i+1}). \tag{4.26}$$

Let us estimate $\sum_{i=0}^{k-1} \Delta_i$ from above. Note that

$$J_{i+1} \overset{(4.6)}{\succeq} \frac{\nabla^2 f(x_{i+1})}{1+\frac{Mr_{i+1}}{2}} \overset{(4.7)}{\succeq} \frac{1}{\delta_i}J_i, \tag{4.27}$$

where

$$\delta_i \overset{\text{def}}{=} \left(1+\frac{Mr_{i+1}}{2}\right)\left(1+\frac{Mr_i}{2}\right). \tag{4.28}$$

Hence,

$$\begin{aligned}\langle J_{i+1}^{-1} - J_i^{-1}, G_{i+1}\rangle &\overset{(4.27)}{\leq} (1-\delta_i^{-1})\langle J_{i+1}^{-1}, G_{i+1}\rangle \\ &\overset{(4.12)}{\leq} (1-\delta_i^{-1})\sqrt{\frac{3}{2}\frac{L}{\mu}}\langle J_{i+1}^{-1}, J_{i+1}\rangle \\ &\overset{(1.3)}{=} \sqrt{\frac{3}{2}\frac{nL}{\mu}}(1-\delta_i^{-1}) \overset{(4.23)}{\leq} \frac{4nL}{3\mu}(1-\delta_i^{-1}),\end{aligned}$$

and

$$\begin{aligned}\sum_{i=0}^{k-1}\Delta_i &\overset{(4.26)}{\leq} \frac{4nL}{3\mu}\sum_{i=0}^{k-1}(1-\delta_i^{-1}) + \sum_{i=0}^{k-1}\ln \operatorname{Det}(J_i^{-1}, J_{i+1}) \\ &= \frac{4nL}{3\mu}\sum_{i=0}^{k-1}(1-\delta_i^{-1}) + \ln \operatorname{Det}(J_0^{-1}, J_k).\end{aligned} \tag{4.29}$$

At the same time,

$$\sum_{i=0}^{k-1}(1-\delta_i^{-1}) \leq \sum_{i=0}^{k-1}\left(1-e^{-\frac{M(r_i+r_{i+1})}{2}}\right) \leq \frac{M}{2}\sum_{i=0}^{k-1}(r_i+r_{i+1}) \leq M\sum_{i=0}^{k} r_i \overset{(4.13)}{\leq} M\lambda_0 \sum_{i=0}^{k-1}\left(1-\frac{\mu}{2L}\right)^i \leq \frac{2L}{\mu}M\lambda_0 \overset{(4.10)}{\leq} \frac{\ln\frac{3}{2}}{2} \leq \frac{1}{4}.$$

Thus,

$$\sum_{i=0}^{k-1}\Delta_i \overset{(4.29)}{\leq} \frac{nL}{3\mu} + \ln \mathrm{Det}(J_0^{-1}, J_k). \tag{4.30}$$

Summing up (4.25) and using the fact that $\psi_k \geq 0$, we obtain

$$\begin{aligned}
\frac{1}{7}\sum_{i=0}^{k-1} p_i\theta_i^2 &\overset{(4.25)}{\leq} \psi_0 - \psi_k + \sum_{i=0}^{k-1}\Delta_i \leq \psi_0 + \sum_{i=0}^{k-1}\Delta_i \\
&\overset{(4.2)}{=} \psi(J_0, LB) + \sum_{i=0}^{k-1}\Delta_i \\
&\overset{(2.14)}{=} \langle J_0^{-1}, LB - J_0\rangle - \ln \mathrm{Det}(J_0^{-1}, LB) + \sum_{i=0}^{k-1}\Delta_i \\
&\overset{(4.30)}{\leq} \langle J_0^{-1}, LB - J_0\rangle + \frac{nL}{3\mu} - \ln \mathrm{Det}(J_k^{-1}, LB) \\
&\overset{(4.3)}{\leq} \langle J_0^{-1}, \frac{L}{\mu}J_0 - J_0\rangle + \frac{nL}{3\mu} \\
&\overset{(1.3)}{=} n\left(\frac{L}{\mu}-1\right) + \frac{nL}{3\mu} \leq \frac{4}{3}\frac{nL}{\mu}.
\end{aligned} \tag{4.31}$$

Since $(1+t)^p \leq 1 + pt$ for all $t \geq -1$ and $0 \leq p \leq 1$, we further have

$$1 + \frac{Mr_i}{2} \leq e^{\frac{Mr_i}{2}} \overset{(4.13)}{\leq} e^{\frac{M\lambda_0}{2}} \overset{(4.10)}{\leq} \left(\frac{3}{2}\right)^{1/8} = \sqrt{\left(\frac{3}{2}\right)^{1/4}} \leq \sqrt{1+\frac{1}{4}\cdot\frac{1}{2}} = \sqrt{\frac{9}{8}}.$$

Therefore, by Lemma 4.2 and the arithmetic-geometric mean inequality,

$$\lambda_f(x_k) \le \lambda_f(x_0) \prod_{i=0}^{k-1} \left[\sqrt{\frac{9}{8}\theta_i}\right] = \frac{1}{\prod_{i=0}^{k-1} p_i^{1/2}} \left[\left(\frac{9}{8}\right)^k \prod_{i=0}^{k-1} p_i \theta_i^2\right]^{1/2} \lambda_f(x_0)$$
$$\le \frac{1}{\prod_{i=0}^{k-1} p_i} \left(\frac{9}{8} \cdot \frac{1}{k} \sum_{i=0}^{k-1} p_i \theta_i^2\right)^{k/2} \lambda_f(x_0) \overset{(4.31)}{\le} \left(\frac{9}{8} \cdot 7 \cdot \frac{4}{3} \frac{nL}{\mu k}\right)^{k/2} \lambda_f(x_0)$$
$$\le \left(\frac{21nL}{2\mu k}\right)^{k/2} \lambda_f(x_0) \le \left(\frac{11nL}{\mu k}\right)^{k/2} \lambda_f(x_0).$$

The proof is finished □

5 Discussion

Let us compare the rates of superlinear convergence, that we have obtained for the classical quasi-Newton methods, with those of the greedy quasi-Newton methods [31]. For brevity, we discuss only the BFGS method. Moreover, since the complexity bounds for the general nonlinear case differ from those for the quadratic one only in some absolute constants (both for the classical and the greedy methods), we only consider the case, when the objective function f is quadratic.

As before, let n be the dimension of the problem, μ be the strong convexity parameter, L be the Lipschitz constant of the gradient of f, and $\lambda_f(x)$ be the local norm of the gradient of f at the point $x \in \mathbb{E}$ (as defined by (3.4)). Also, let us introduce the following condition number to simplify our notation:

$$Q \overset{\text{def}}{=} \frac{nL}{\mu} \ge 1. \tag{5.1}$$

The greedy BFGS method [31] is essentially the classical BFGS algorithm (scheme (3.3) with $\phi_k \equiv 0$) with the only difference that, at each iteration, the update direction u_k is chosen greedily according to the following rule:

$$u_k \overset{\text{def}}{=} \underset{u \in \{e_1, \ldots, e_n\}}{\operatorname{argmax}} \frac{\langle G_k u, u\rangle}{\langle Au, u\rangle},$$

where $e_1, \ldots, e_n$ is a basis in $\mathbb{E}$, such that $B^{-1} = \sum_{i=1}^n e_i e_i^*$. For this method, we have the following recurrence (see [31, Theorem 3.2]):

$$\lambda_f(x_{k+1}) \le \left(1 - \frac{1}{Q}\right)^k Q\lambda_f(x_k) \le e^{-\frac{k}{Q}} Q\lambda_f(x_k), \quad k \ge 0.$$

Hence, its rate of superlinear convergence is described by the expression

$$\lambda_f(x_k) \le \lambda_f(x_0) \prod_{i=0}^{k-1} \left[e^{-\frac{i}{Q}} Q \right] = e^{-\frac{k(k-1)}{2Q}} Q^k \lambda_f(x_0) \overset{\text{def}}{=} A_k, \quad k \ge 0. \quad (5.2)$$

Although the inequality (5.2) is valid for all $k \ge 0$, it is useful only when

$$e^{-\frac{k(k-1)}{2Q}} Q^k \le 1 \qquad \Longleftrightarrow \qquad k \ge 1 + 2Q \ln Q. \quad (5.3)$$

In other words, the relation (5.3) specifies the moment, starting from which it becomes meaningful to speak about the superlinear convergence of the greedy BFGS method.

For the classical BFGS method, we have the following bound (see (3.11)):

$$\lambda_f(x_k) \le \left(\frac{Q}{k} \right)^{k/2} \lambda_f(x_0) \overset{\text{def}}{=} B_k, \quad k \ge 1,$$

and the starting moment of its superlinear convergence is described as follows:

$$\left(\frac{Q}{k} \right)^{k/2} \le 1 \qquad \Longleftrightarrow \qquad k \ge Q. \quad (5.4)$$

Comparing (5.3) and (5.4), we see that, for the standard BFGS, the superlinear convergence may start slightly earlier than for the greedy one. However, the difference is only in the logarithmic factor.

Nevertheless, let us show that, very soon after the superlinear convergence of the greedy BFGS begins, namely, after

$$K \overset{\text{def}}{=} 1 + 6Q \ln(4Q) \qquad (\overset{(5.1)}{\ge} 7) \quad (5.5)$$

iterations, it will be significantly faster than the standard BFGS. Indeed,

$$\begin{aligned} \frac{A_k}{B_k} &= e^{-\frac{k(k-1)}{2Q}} Q^k \left(\frac{k}{Q} \right)^{k/2} = e^{-\frac{k(k-1)}{2Q}} (Qk)^{k/2} \\ &= e^{-\frac{k(k-1)}{2Q} + \frac{k}{2} \ln(Qk)} = e^{-\frac{k(k-1)}{2Q} \left[1 - \frac{Q \ln(Qk)}{k-1} \right]} \end{aligned} \quad (5.6)$$

for all $k \ge 1$. Note that the function $t \mapsto \frac{\ln t}{t}$ is decreasing on $[e, +\infty)$ (since its logarithm $\ln \ln t - \ln t$ is a decreasing function of $u = \ln t$ for $u \in [1, +\infty)$, which is easily verified by differentiation). Hence, for all $k \ge K$, we have (using first that $k \le 2(k-1)$ since $k \ge 2$)

$$\frac{Q \ln(Qk)}{k-1} \le \frac{Q \ln(2Q(k-1))}{k-1} \le \frac{Q \ln(2Q(K-1))}{K-1} \overset{(5.5)}{=} \frac{\ln\left(12 Q^2 \ln(4Q)\right)}{6 \ln(4Q)}$$

$$\leq \frac{\ln(48Q^3)}{6\ln(4Q)} \leq \frac{\ln(64Q^3)}{6\ln(4Q)} = \frac{3\ln(4Q)}{6\ln(4Q)} = \frac{1}{2}.$$

Consequently, for all $k \geq K$, we obtain

$$\frac{A_k}{B_k} \overset{(5.6)}{\leq} e^{-\frac{k(k-1)}{4Q}} \leq 1.$$

Thus, after K iterations, the rate of superlinear convergence of the greedy BFGS is always better than that of the standard BFGS. Moreover, as $k \to \infty$, the gap between these two rates grows as $e^{-k^2/Q}$. At the same time, the complexity of the Hessian update for the greedy BFGS method is more expensive than for the standard one.

6 Appendix

Lemma 6.1 *Let $A, B : \mathbb{E} \to \mathbb{E}^*$ be self-adjoint linear operators such that*

$$0 \prec A \preceq B. \tag{6.1}$$

Then, for any $u \in \mathbb{E} \setminus \{0\}$, we have

$$A - \frac{Auu^*A}{\langle Au, u\rangle} \preceq B - \frac{Buu^*B}{\langle Bu, u\rangle}.$$

Proof Indeed, for all $h \in \mathbb{E}$, we have

$$\begin{aligned}
\langle Ah, h\rangle - \frac{\langle Au, h\rangle^2}{\langle Au, u\rangle} &= \min_{\alpha\in\mathbb{R}} \left[\langle Ah, h\rangle - 2\alpha\langle Ah, u\rangle + \alpha^2\langle Au, u\rangle\right] \\
&= \min_{\alpha\in\mathbb{R}} \langle A(h-\alpha u), h-\alpha u\rangle \\
&\overset{(6.1)}{\leq} \min_{\alpha\in\mathbb{R}} \langle B(h-\alpha u), h-\alpha u\rangle \\
&= \min_{\alpha\in\mathbb{R}} \left[\langle Bh, h\rangle - 2\alpha\langle Bh, u\rangle + \alpha^2\langle Bu, u\rangle\right] \\
&= \langle Bh, h\rangle - \frac{\langle Bu, h\rangle^2}{\langle Bu, u\rangle}.
\end{aligned}$$

The proof is finished □

Lemma 6.2 *For any self-adjoint positive definite linear operators $A, G : \mathbb{E} \to \mathbb{E}^*$, any scalar $\phi \in \mathbb{R}$, and any direction $u \in \mathbb{E} \setminus \{0\}$, we have*

$$\mathrm{Det}(G^{-1}, \mathrm{Broyd}_\phi(A, G, u)) = \phi\frac{\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle} + (1-\phi)\frac{\langle Au, u\rangle}{\langle Gu, u\rangle}. \tag{6.2}$$

Remark 6.1 Note that formula (6.2) is known in the literature (see e.g. [30, eq. (1.9)]), although we do not know any reference, which contains an explicit proof of this result.

Proof Denote $G_+ \overset{\text{def}}{=} \text{Broyd}_\phi(A, G, u)$,

$$G_0 \overset{\text{def}}{=} G - \frac{Guu^*G}{\langle Gu, u\rangle} + \frac{Auu^*A}{\langle Au, u\rangle}, \qquad s \overset{\text{def}}{=} \frac{Au}{\langle Au, u\rangle} - \frac{Gu}{\langle Gu, u\rangle}. \tag{6.3}$$

Note that

$$\begin{aligned} G_+ &\overset{(2.1)}{=} G_0 + \phi\left[\frac{\langle Gu, u\rangle Auu^*A}{\langle Au, u\rangle^2} + \frac{Guu^*G}{\langle Gu, u\rangle} - \frac{Auu^*G + Guu^*A}{\langle Au, u\rangle}\right] \\ &= G_0 + \phi\langle Gu, u\rangle ss^*, \end{aligned}$$

and

$$\langle s, u\rangle = 0. \tag{6.4}$$

Let $Q \overset{\text{def}}{=} G + \frac{Auu^*A}{\langle Au,u\rangle}$. Note that

$$Qu = Gu + Au, \qquad QG^{-1}Au = \left(1 + \frac{\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle}\right) Au, \tag{6.5}$$

and $G_0 = Q - \frac{Guu^*G}{\langle Gu,u\rangle}$. Therefore, applying twice Lemma 6.3, we find that

$$\begin{aligned} \text{Det}(G^{-1}, Q) &= 1 + \frac{\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle}, \\ \text{Det}(Q^{-1}, G_0) &= 1 - \frac{\langle GQ^{-1}Gu, u\rangle}{\langle Gu, u\rangle} \overset{(6.5)}{=} 1 - \frac{\langle Gu - GQ^{-1}Au, u\rangle}{\langle Gu, u\rangle} \\ &= \frac{\langle GQ^{-1}Au, u\rangle}{\langle Gu, u\rangle} \overset{(6.5)}{=} \frac{\langle Au, u\rangle}{\langle Gu, u\rangle\left(1 + \frac{\langle AG^{-1}Au,u\rangle}{\langle Au,u\rangle}\right)}. \end{aligned}$$

Hence,

$$\text{Det}(G^{-1}, G_0) \overset{(1.4)}{=} \text{Det}(G^{-1}, Q)\text{Det}(Q^{-1}, G_0) = \frac{\langle Au, u\rangle}{\langle Gu, u\rangle}. \tag{6.6}$$

Further, note that

$$G_0u \overset{(6.3)}{=} Au, \qquad G_0G^{-1}Au \overset{(6.3)}{=} \frac{\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle}Au + Au - \frac{\langle Au, u\rangle}{\langle Gu, u\rangle}Gu. \tag{6.7}$$

So, applying Lemma 6.3 again, we obtain

$$\begin{aligned}
\mathrm{Det}(G_0^{-1}, G_+) &\overset{(6.3)}{=} 1 + \phi\langle Gu, u\rangle\langle s, G_0^{-1}s\rangle \\
&\overset{(6.3)}{=} 1 + \phi\frac{\langle Gu, u\rangle}{\langle Au, u\rangle}\langle s, G_0^{-1}Au - \frac{\langle Au, u\rangle}{\langle Gu, u\rangle}G_0^{-1}Gu\rangle \\
&\overset{(6.7)}{=} 1 + \phi\frac{\langle Gu, u\rangle}{\langle Au, u\rangle}\langle s, G^{-1}Au - \frac{\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle}G_0^{-1}Au\rangle \\
&\overset{(6.7)}{=} 1 + \phi\frac{\langle Gu, u\rangle}{\langle Au, u\rangle}\langle s, G^{-1}Au - \frac{\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle}u\rangle \\
&\overset{(6.4)}{=} 1 + \phi\frac{\langle Gu, u\rangle}{\langle Au, u\rangle}\langle s, G^{-1}Au\rangle \\
&\overset{(6.3)}{=} 1 + \phi\frac{\langle Gu, u\rangle}{\langle Au, u\rangle}\langle \frac{Au}{\langle Au, u\rangle} - \frac{Gu}{\langle Gu, u\rangle}, G^{-1}Au\rangle \\
&= \phi\frac{\langle Gu, u\rangle\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle^2} + 1 - \phi. \qquad (6.8)
\end{aligned}$$

Consequently,

$$\begin{aligned}
\mathrm{Det}(G^{-1}, G_+) &\overset{(1.4)}{=} \mathrm{Det}(G^{-1}, G_0)\mathrm{Det}(G_0^{-1}, G_+) \overset{(6.6)}{=} \frac{\langle Au, u\rangle}{\langle Gu, u\rangle}\mathrm{Det}(G_0^{-1}, G_+) \\
&\overset{(6.8)}{=} \frac{\langle Au, u\rangle}{\langle Gu, u\rangle}\left(\phi\frac{\langle Gu, u\rangle\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle^2} + 1 - \phi\right) \\
&= \phi\frac{\langle AG^{-1}Au, u\rangle}{\langle Au, u\rangle} + (1-\phi)\frac{\langle Au, u\rangle}{\langle Gu, u\rangle}.
\end{aligned}$$

The proof is finished □

Lemma 6.3 *(Determinant of rank-1 perturbation) Let $A : \mathbb{E} \to \mathbb{E}^*$ be a self-adjoint positive definite linear operator, $s \in \mathbb{E}^*$, and $\alpha \in \mathbb{R}$. Then,*

$$\mathrm{Det}(A^{-1}, A + \alpha ss^*) = 1 + \alpha\langle s, A^{-1}s\rangle.$$

Proof Indeed, with respect to A, the operator $A + \alpha ss^*$ has $n-1$ unit eigenvalues and one eigenvalue $1 + \alpha\langle s, A^{-1}s\rangle$ (corresponding to the eigenvector $A^{-1}s$). □

Acknowledgements The authors are thankful to two anonymous reviewers for their valuable time and useful feedback.

References

1. Davidon, W.: Variable metric method for minimization. Argonne Natl. Lab. Res. Dev. Rep. **5990**, (1959)
2. Fletcher, R., Powell, M.: A rapidly convergent descent method for minimization. Comput. J. **6**(2), 163–168 (1963)
3. Broyden, C.: The convergence of a class of double-rank minimization algorithms: 1. General considerations. IMA J. Appl. Math. **6**(1), 76–90 (1970)
4. Broyden, C.: The convergence of a class of double-rank minimization algorithms: 2. The new algorithm. IMA J. Appl. Math. **6**(3), 222–231 (1970)
5. Fletcher, R.: A new approach to variable metric algorithms. Comput. J. **13**(3), 317–322 (1970)
6. Goldfarb, D.: A family of variable-metric methods derived by variational means. Math. Comput. **24**(109), 23–26 (1970)
7. Shanno, D.: Conditioning of quasi-Newton methods for function minimization. Math. Comput. **24**(111), 647–656 (1970)
8. Dennis, J., Moré, J.: Quasi-Newton methods, motivation and theory. SIAM Rev. **19**(1), 46–89 (1977)
9. Nocedal, J., Wright, S.: Numerical Optimization. Springer, Berlin (2006)
10. Lewis, A., Overton, M.: Nonsmooth optimization via quasi-Newton methods. Math. Program. **141**(1–2), 135–163 (2013)
11. Gower, R., Goldfarb, D., Richtárik, P.: Stochastic block BFGS: squeezing more curvature out of data. In: International Conference on Machnie Learning, pp. 1869–1878 (2016)
12. Gower, R., Richtárik, P.: Randomized quasi-Newton updates are linearly convergent matrix inversion algorithms. SIAM J. Matrix Anal. Appl. **38**(4), 1380–1409 (2017)
13. Kovalev, D., Gower, R., Richtárik, P., Rogozin, A.: Fast linear convergence of randomized BFGS. (2020). arXiv:2002.11337
14. Powell, M.: On the convergence of the variable metric algorithm. IMA J. Appl. Math. **7**(1), 21–36 (1971)
15. Dixon, L.: Quasi-Newton algorithms generate identical points. Math. Program. **2**(1), 383–387 (1972)
16. Dixon, L.: Quasi Newton techniques generate identical points II: the proofs of four new theorems. Math. Program. **3**(1), 345–358 (1972)
17. Broyden, C.: Quasi-Newton methods and their application to function minimization. Math. Comput. **21**(99), 368–381 (1967)
18. Broyden, C., Dennis, J., Moré, J.: On the local and superlinear convergence of quasi-Newton methods. IMA J. Appl. Math. **12**(3), 223–245 (1973)
19. Dennis, J., Moré, J.: A characterization of superlinear convergence and its application to quasi-Newton methods. Math. Comput. **28**(126), 549–560 (1974)
20. Stachurski, A.: Superlinear convergence of Broyden's bounded θ-class of methods. Math. Program. **20**(1), 196–212 (1981)
21. Griewank, A., Toint, P.: Local convergence analysis for partitioned quasi-Newton updates. Numer. Math. **39**(3), 429–448 (1982)
22. Engels, J., Martínez, H.: Local and superlinear convergence for partially known quasi-Newton methods. SIAM J. Optim. **1**(1), 42–56 (1991)
23. Yabe, H., Yamaki, N.: Local and superlinear convergence of structured quasi-Newton methods for nonlinear optimization. J. Oper. Res. Soc. Jpn. **39**(4), 541–557 (1996)
24. Wei, Z., Yu, G., Yuan, G., Lian, Z.: The superlinear convergence of a modified BFGS-type method for unconstrained optimization. Comput. Optim. Appl. **29**(3), 315–332 (2004)
25. Yabe, H., Ogasawara, H., Yoshino, M.: Local and superlinear convergence of quasi-Newton methods based on modified secant conditions. J. Comput. Appl. Math. **205**(1), 617–632 (2007)
26. Mokhtari, A., Eisen, M., Ribeiro, A.: IQN: an incremental quasi-Newton method with local superlinear convergence rate. SIAM J. Optim. **28**(2), 1670–1698 (2018)
27. Gao, W., Goldfarb, D.: Quasi-Newton methods: superlinear convergence without line searches for self-concordant functions. Optim. Methods Softw. **34**(1), 194–217 (2019)
28. Byrd, R., Nocedal, J., Yuan, Y.: Global convergence of a class of quasi-Newton methods on convex problems. SIAM J. Numer. Anal. **24**(5), 1171–1190 (1987)
29. Byrd, R., Nocedal, J.: A tool for the analysis of quasi-Newton methods with application to unconstrained minimization. SIAM J. Numer. Anal. **26**(3), 727–739 (1989)

30. Byrd, R., Liu, D., Nocedal, J.: On the behavior of Broyden's class of quasi-Newton methods. SIAM J. Optim. **2**(4), 533–557 (1992)
31. Rodomanov, A., Nesterov, Y.: Greedy quasi-Newton methods with explicit superlinear convergence. In: CORE Discussion Papers 06, (2020)
32. Nesterov, Y.: Lectures on convex optimization. Springer, Berlin (2018)