Electronic Transactions on Numerical Analysis.
Volume 55, pp. 285–309, 2022.
Copyright © 2022, Kent State University.
ISSN 1068–9613.
DOI: 10.1553/etna_vol55s285

ETNA
Kent State University and
Johann Radon Institute (RICAM)

ONE-STEP CONVERGENCE OF INEXACT ANDERSON ACCELERATION FOR CONTRACTIVE AND NON-CONTRACTIVE MAPPINGS*

FEI XUE†

Abstract. We give a one-step convergence analysis of inexact Anderson acceleration for the fixed point iteration $x_{k+1} = g(x_k)$ with a potentially non-contractive mapping g, where $g(x_k)$ is evaluated approximately and the minimization of the nonlinear residual norms is performed in the vector 2-norm by the linear least-squares method. If g is non-contractive, then the original fixed point iteration does not converge, but a recent analysis by S. Pollock and L. Rebholz [IMA J. Numer. Anal., 41 (2021), pp. 2841–2872] shows that Anderson acceleration may still converge provided that the minimization at each step has a sufficient gain. In this paper, we show that inexact Anderson acceleration exhibits essentially the same convergence behavior as the exact algorithm if each $g(x_k)$ is evaluated with an error proportional to the nonlinear residual norm $\|w_k\| = \|g(x_k) - x_k\|$, regardless of whether g is contractive or not. This means that the existing relationship between exact and inexact Anderson acceleration can be generalized in a unified framework for both contractive and non-contractive mappings. Numerical experiments show that the inexact algorithm can converge as rapidly as the exact counterpart while it can lower the computational cost.

Key words. fixed point iteration, inexact Anderson acceleration, non-contractive mapping, one-step convergence

AMS subject classifications. 65N22, 65H10, 65F50

1. Introduction.

Anderson acceleration is a computational technique designed for accelerating the convergence of fixed point iterations [1, 2]. Variants of this strategy are also referred to as nonlinear GMRES [28, 29] or Anderson mixing, Pulay mixing [19, 20], and direct inversion in the iterative subspace (DIIS) [21] in the community of quantum mechanics. The idea of this approach is to find a proper linear combination of successive previous iterates with coefficients obtained from a constrained minimization to obtain a new iterate that potentially yields a smaller nonlinear residual norm than the new iterate computed from the original fixed point iteration. Since the solution of systems of nonlinear equations by fixed point iterations is ubiquitous in science and engineering, Anderson acceleration has been widely used in a variety of applications; see, e.g., [13, 17, 28] and the references therein.

To better understand the behavior of Anderson acceleration, the connections between this method and other algorithms for solving nonlinear systems of equations have been explored. In particular, it is shown in [10, 11, 21] that Anderson acceleration is equivalent to a special variant of the generalized Broyden's method, and in particular, the approximate inverse Jacobian is obtained implicitly from an optimization subject to secant equations involving the most recent iterates [11]. Such an equivalence naturally prompts one to consider exploring the convergence of Anderson acceleration from the large volume of existing literature on the convergence of quasi-Newton's methods; see, e.g., [5, 13] and the references therein. On the other hand, if the underlying fixed point iteration exhibits *linear* convergence, then extensive numerical evidence suggest that the minimization step adopted by Anderson acceleration usually helps to improve the robustness as well as the rate of convergence. Such a favorable property makes it reasonable to study the convergence of Anderson acceleration within its own framework based on minimization. An early major convergence result was given in [27] for *contractive* mappings, and this was later improved in [17] for contractive and *non-contractive* mappings. However, it is shown [9] that Anderson acceleration may not speed up quadratically convergent iterations.

*Received February 8, 2021. Accepted December 15, 2021. Published online on January 26, 2022. Recommended by Quiang Ye. This research was supported by the U.S. National Science Foundation under grants DMS-1719461 and DMS-1819097.

†School of Mathematical and Statistical Sciences, Clemson University, O-203 Martin Hall, Box 340975, Clemson, SC 29634 (fxue@clemson.edu).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

To save costs for iterative methods for solving large nonlinear systems, an economic strategy is to use *inexact* methods, which follow the process of their exact counterpart but evaluate each iterate only approximately. For instance, inexact Newton's method in the generic nonlinear system setting is a well-known example [5, 6]; for eigenvalue computations, inexact Newton-like methods have also been explored [12, 24, 25]. Inexact Anderson acceleration is another example in this line of research, explored both numerically [15, 26] and theoretically [26] for *contractive* mappings. The results show that inexact Anderson acceleration may converge as rapidly as the exact method, provided that the error introduced in the evaluation of each new iterate is proportional to the current nonlinear residual norm; also, the final accuracy of the approximate solution obtained from the inexact method is at worst proportional to the uniform upper bound for the errors in the evaluation of *all* iterates.

In this paper, we give a one-step convergence analysis of inexact Anderson acceleration for both contractive and *non-contractive* mappings. Note that the existing results in [26, 27] assumed that the mapping g is *contractive* and the coefficients obtained in the minimization step of Anderson acceleration are *uniformly bounded*. By contrast, we show that inexact Anderson acceleration *may* converge to the desired solution even if the mapping g is *non-contractive* near the solution, as long as the minimization at each step has sufficient gain and the least-squares problem that determine the coefficients of the linear combination of the previous iterates is not ill-conditioned. Assuming that each new iterate is evaluated with increasing accuracy as the iteration proceeds, we show that the inexact method closely follows the behavior of the exact variant. Note that unlike [26], we have no theories about the convergence for multiple iteration steps due to the non-contractiveness of the mapping g, though such a behavior is usually observed in numerical experiments.

Our work is motivated by a recent one-step convergence analysis of (exact) Anderson acceleration [17]. To understand the inexact method, we first develop a similar analysis of the exact variant. Our analysis is primarily based on linear algebra and on properties of orthogonal and oblique projectors in particular, to gain insight into the effect of the least-squares minimization in Anderson acceleration. Compared to the main conclusion in [17, Theorem 5.5], our result gives the same convergence factor for the linear term of the nonlinear residual, but we have a simpler upper bound for the higher-order terms that does not exhibit an explicit exponential growth with the acceleration depth and does not involve the squares of the most recent nonlinear residual norm if the minimization at the current step has the largest possible gain. More importantly, our analysis can be extended without difficulty to study the effects of an approximate evaluation of the fixed point mapping for the inexact method by exploring the impact of small perturbations in the relevant projectors, whereas it seems less clear how this idea could be realized based on the analysis in [17].

The rest of the paper is structured as follows: In Section 2, we consider the fixed point iteration $x_{k+1} = g(x_k)$ and state assumptions on the mapping g; we also outline Anderson acceleration and then present a few preliminary results for the subsequent analysis. In Section 3, we give a one-step convergence analysis of (exact) Anderson acceleration based on projectors and angles between subspaces, showing a result similar to that in [17] with a new bound for the higher-order terms. In Section 4, we provide an analysis of inexact Anderson acceleration, where each update $g(x_k)$ is allowed to be evaluated with an error proportional to the residual norm $\|w_k\| = \|g(x_k) - x_k\|$ without obviously affecting the convergence rate of the algorithm. Numerical results are shown in Section 5 to support our analysis. Finally, Section 6 gives conclusions.

Throughout the paper, the norm $\|\cdot\|$ refers to the 2-norm of matrices and vectors unless stated otherwise. The range (column space) of a matrix C is denoted as $\mathrm{col}(A) = \mathrm{range}(A)$, and $\mathrm{null}(C)$ refers to the null space of C.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

2. Problem settings and preliminaries. Consider the numerical iterative solution of the nonlinear system of equations

$$x = g(x).$$

We make the following assumptions on g and its derivative g' throughout the paper:

ASSUMPTION 1. *Let $g : X \to X$ be Fréchet differentiable, where $X \subset \mathbb{R}^n$ is convex, equipped with the inner product $(\cdot,\cdot)$ and the induced 2-norm $\|\cdot\|$. The corresponding matrix 2-norm is defined as $\|A\| = \sup_{u\in\mathbb{R}^n\setminus\{0\}} \frac{\|Au\|}{\|u\|}$. Assume that there exists a unique fixed point $x^* \in X$ such that $x^* = g(x^*)$ and there are positive constants κ_g, L_g, and σ_g such that*

1. *$\|g'(x)\| \le \kappa_g$ for all $x \in X$;*
2. *$\|g'(x) - g'(y)\| \le L_g\|x - y\|$ for all $x, y \in X$;*
3. *$\min_{u\in\mathbb{R}^n\setminus\{0\}} \frac{\left\|\left(\int_0^1 g'(z(t))dt - I\right)u\right\|}{\|u\|} \ge \sigma_g$, i.e., $\left\|\left(\int_0^1 g'(z(t))\,dt - I\right)^{-1}\right\| \le \frac{1}{\sigma_g}$, where $z(t) = (1-t)x + ty$ $(0 \le t \le 1)$, for all $x, y \in X$.*

The outline of Anderson acceleration is given in Algorithm 1.

Algorithm 1 Anderson acceleration for solving the nonlinear system $x = g(x)$.

Input: function $g : X \to X$, $x_0 \in X$, integer $m > 0$, $\beta \in (0, 1]$, and tolerance $\delta > 0$.
Output: an approximate solution x_k such that $x_k \approx g(x_k)$.

1: Compute $x_1 = g(x_0)$, and $w_0 = g(x_0) - x_0$.
2: **for** $k = 1, 2, \ldots,$ **do**
3: Compute $g(x_k)$ and $w_k = g(x_k) - x_k$.
4: **if** $\|w_k\| \le \delta$ **then**
5: terminate the algorithm and return x_k.
6: **end if**
7: Let $\ell = \max\{0, k - m\}$, and solve $\min_{\sum_{i=\ell}^k \alpha_i^{(k)}=1} \left\|\sum_{i=\ell}^k \alpha_i^{(k)} w_i\right\|$ for $\{\alpha_i^{(k)}\}$.
8: Evaluate the new iterate $x_{k+1} = (1-\beta)\sum_{j=\ell}^k \alpha_j^{(k)} x_j + \beta \sum_{j=\ell}^k \alpha_j^{(k)} g(x_j)$.
9: **end for**

REMARK 2.1. We note that at each step, Algorithm 1 performs only one function evaluation (namely in line 3). At each step $k \ge m$, the algorithm keeps a record of the recent iterates $x_{k-m}, x_{k-m+1}, \ldots, x_k$, their function evaluations $g(x_{k-m}), g(x_{k-m+1}), \ldots, g(x_k)$, and their nonlinear residuals $w_{k-m}, w_{k-m+1}, \ldots, w_k$, all of which have been computed from step $k - m$ to k. At the end of step k, the old vectors x_{k-m}, $g(x_{k-m})$, and w_{k-m} will not be used in the future and should be discarded, and the most recent vectors x_k, $g(x_k)$, and w_k should be added to the three sets of vectors to prepare for the new step $k + 1$.

2.1. Preliminaries.

(a) The connection between iterates and residuals. Note that each iterate x_i can be connected with its residual $w_i = g(x_i) - x_i$, and a similar relation holds for $x_i - x_j$. In fact, let $z_{*i}(t) = x^* + (x_i - x^*)t$ and $z_{ji}(t) = x_j + (x_i - x_j)t$. We have

$$\begin{aligned}\int_0^1 (g'(z_{*i}(t)) - I)\, dz_{*i}(t) &= \int_0^1 (g'(z_{*i}(t)) - I)\,(x_i - x^*)dt \\ &= g(x_i) - g(x^*) - (x_i - x^*) = g(x_i) - x_i = w_i,\end{aligned}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

from which we obtain

$$x_i - x^* = \Big(\int_0^1 g'(z_{*i}(t))dt - I \Big)^{-1} w_i. \tag{2.1}$$

Also, let $z_{j*}(t) = x_j + (x^* - x_j)t$. Since g' satisfies a Lipschitz condition, by (2.1),

$$\begin{aligned} &\left\| \int_0^1 \left(g'(z_{ji}(t)) - g'(z_{j*}(t)) \right) dt \right\| \\ &\qquad \le \int_0^1 \| g'(z_{ji}(t)) - g'(z_{j*}(t)) \| \, dt \\ &\qquad \le \int_0^1 L_g \| z_{ji}(t) - z_{j*}(t) \| dt = \int_0^1 t L_g \| x_i - x^* \| dt \\ &\qquad \le \frac{L_g}{2} \left\| \Big(\int_0^1 g'(z_{i*}(t))dt - I \Big)^{-1} \right\| \|w_i\| \le \frac{L_g}{2\sigma_g} \|w_i\|. \end{aligned} \tag{2.2}$$

Similarly, note that

$$\begin{aligned} w_i - w_j &= (g(x_i) - x_i) - (g(x_j) - x_j) = g(x_i) - g(x_j) - (x_i - x_j) \\ &= \int_0^1 g'(z_{ji}(t)) \, (x_i - x_j) dt - (x_i - x_j) = \int_0^1 \left(g'(z_{ji}(t)) - I \right) (x_i - x_j) dt, \end{aligned}$$

which leads to

$$x_i - x_j = \Big(\int_0^1 g'(z_{ji}(t)) \, dt - I \Big)^{-1} (w_i - w_j). \tag{2.3}$$

(b) Bounding the difference in the residuals. Next we explore, for each *fixed* index j ($k - m \le j \le k$) and each *running* index i ($k - m \le i \le k$, $i \ne j$), a connection between $\left\| \alpha_i^{(k)} (w_j - w_i) \right\|$ and $\|w_j\|$. To this end, recall that the coefficients $\alpha_i^{(k)}$ ($k - m \le i \le k$) are defined as the solution to

$$\min_{\sum_{i=k-m}^{k} \alpha_i^{(k)} = 1} \left\| \sum_{i=k-m}^{k} \alpha_i^{(k)} w_i \right\| = \min \left\| w_j - \sum_{i=k-m, i \ne j}^{k} \alpha_i^{(k)} (w_j - w_i) \right\|, \tag{2.4}$$

for each j ($k - m \le j \le k$). The right-hand side of (2.4) indicates that this minimization can be done by the linear least-squares approach.

To study the least-squares problem, we define two blocks of vectors and their column spaces,

$$\begin{aligned} U_j^{(k)} &= [w_j - w_k, \ldots, w_j - w_{j+1}, w_j - w_{j-1}, \ldots, w_j - w_{k-m}] \in \mathbb{R}^{n \times m}, \\ U_{j[i]}^{(k)} &= [w_j - w_k, \ldots, w_j - w_{i+1}, w_j - w_{i-1}, \ldots, \\ &\qquad w_j - w_{j+1}, w_j - w_{j-1}, \ldots, w_j - w_{k-m}] \in \mathbb{R}^{n \times (m-1)}, \\ \mathcal{U}_j^{(k)} &= \operatorname{col}\big(U_j^{(k)}\big) = \operatorname{span} \left\{ w_j - w_\ell \right\}_{k-m \le \ell \le k, \ell \ne j}, \qquad \text{and} \\ \mathcal{U}_{j[i]}^{(k)} &= \operatorname{col}\big(U_{j[i]}^{(k)}\big) = \operatorname{span} \left\{ w_j - w_\ell \right\}_{k-m \le \ell \le k, \ell \ne j, i}, \end{aligned} \tag{2.5}$$

where $k - m \le i, j \le k$ and $i \ne j$, such that

$$\mathcal{U}_j^{(k)} = \mathcal{U}_{j[i]}^{(k)} \cup \operatorname{span} \left\{ w_j - w_i \right\}. \tag{2.6}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

We furthermore define $\mathcal{U}_{j[i]}^{(k)} = \{\mathbf{0}\}$ and $\mathcal{U}_k^{(k)} = \operatorname{span}\{w_k - w_{k-1}\}$ if $m = 1$.

Assume that $U_j^{(k)}$ and $U_{j[i]}^{(k)}$ have full column rank, and define $\varphi_j^{(k)} = \angle(w_j, \mathcal{U}_j^{(k)})$. It then follows from (2.4) and the properties of the linear least-squares problem that

$$\left\| \sum_{i=k-m, i\neq j}^{k} \alpha_i^{(k)}(w_j - w_i) \right\| = \left\| U_j^{(k)} \left(U_j^{(k)T} U_j^{(k)} \right)^{-1} U_j^{(k)T} w_j \right\| = \cos\varphi_j^{(k)} \|w_j\|.$$

To quantify $\left\|\alpha_i^{(k)}(w_j - w_i)\right\|$, we define the following orthogonal projectors:

$$\begin{aligned}
Q_j^{(k)} &= U_j^{(k)} \big(U_j^{(k)T} U_j^{(k)}\big)^{-1} U_j^{(k)T} && (m \geq 1),\\
Q_{j[i]}^{(k)} &= U_{j[i]}^{(k)} \big(U_{j[i]}^{(k)T} U_{j[i]}^{(k)}\big)^{-1} U_{j[i]}^{(k)T} && (m \geq 2),\\
Q_{j[i]}^{(k)\perp} &= I && (m = 1),\\
Q_{j[i]}^{(k)\perp} &= I - Q_{j[i]}^{(k)} = I - U_{j[i]}^{(k)} \big(U_{j[i]}^{(k)T} U_{j[i]}^{(k)}\big)^{-1} U_{j[i]}^{(k)T} && (m \geq 2),
\end{aligned} \tag{2.7}$$

and the rank-1 projector (typically oblique for $m \geq 2$),

$$P_{j[i]}^{(k)} = \frac{(w_j - w_i)\,(w_j - w_i)^T\, Q_{j[i]}^{(k)\perp}}{(w_j - w_i)^T\, Q_{j[i]}^{(k)\perp}\,(w_j - w_i)}, \tag{2.8}$$

which projects any vector $u \in \mathcal{U}_j^{(k)}$ along $\mathcal{U}_{j[i]}^{(k)}$ onto $\operatorname{span}\{w_j - w_i\}$; see (2.6). Therefore,

$$\mathcal{U}_{j[i]}^{(k)} \subset \operatorname{null}\big(P_{j[i]}^{(k)}\big) \qquad \text{and} \qquad \operatorname{range}\big(P_{j[i]}^{(k)}\big) = \operatorname{span}\{w_j - w_i\}.$$

Moreover, by (2.7), it is not hard to see that

$$\cos\angle\big(w_j - w_i, Q_{j[i]}^{(k)\perp}(w_j - w_i)\big) = \begin{cases} 1 & (m = 1),\\ \sin\angle\big(w_j - w_i, \mathcal{U}_{j[i]}^{(k)}\big) & (m \geq 2), \end{cases}$$

and it follows from (2.8) that

$$P_{j[i]}^{(k)} = \begin{cases} \frac{(w_j - w_i)(w_j - w_i)^T}{\|w_j - w_i\|^2} & (m = 1),\\ \frac{1}{\sin\angle\left(w_j - w_i, \mathcal{U}_{j[i]}^{(k)}\right)} \frac{w_j - w_i}{\|w_j - w_i\|} \frac{(w_j - w_i)^T Q_{j[i]}^{(k)\perp}}{\left\|(w_j - w_i)^T Q_{j[i]}^{k\perp}\right\|} & (m \geq 2). \end{cases} \tag{2.9}$$

The orthogonal projection of w_j onto $\mathcal{U}_j^{(k)}$ is $u = \sum_{i=k-m, i\neq j}^{k} \alpha_i^{(k)}(w_j - w_i)$, and the projection of u along $\mathcal{U}_{j[i]}^{(k)}$ onto $\operatorname{span}\{w_j - w_i\}$ is $\alpha_i^{(k)}(w_j - w_i)$. This means that

$$\alpha_i^{(k)}(w_j - w_i) = P_{j[i]}^{(k)} Q_j^{(k)} w_j. \tag{2.10}$$

Here, note that $\left\|Q_j^{(k)} w_j\right\| = \cos\varphi_j^{(k)} \|w_j\|$ as a result of the linear least-squares problem.

For $m = 1$, it follows from (2.9) that $\alpha_i^{(k)}(w_j - w_i) = P_{j[i]}^{(k)} Q_j^{(k)} w_j = Q_j^{(k)} w_j$. Taking the norm of both sides of (2.10), we have

$$\left\|\alpha_i^{(k)}(w_j - w_i)\right\| = \left\|P_{j[i]}^{(k)} Q_j^{(k)} w_j\right\| = \left\|Q_j^{(k)} w_j\right\| = \cos\varphi_j^{(k)} \|w_j\|.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

For $m \geq 2$, we have $\cos\angle\big(Q_j^{(k)}w_j, Q_{j[i]}^{(k)\perp}(w_j - w_i)\big) = \sin\angle\big(Q_j^{(k)}w_j, Q_{j[i]}^{(k)}(w_j - w_i)\big)$ due to (2.7). Therefore, by (2.9),

$$\begin{aligned}
\big\|\alpha_i^{(k)}(w_j - w_i)\big\| &= \big\|P_{j[i]}^{(k)}Q_j^{(k)}w_j\big\| \\
&= \left\| \frac{1}{\sin\angle\big(w_j - w_i, \mathcal{U}_{j[i]}^{(k)}\big)} \frac{w_j - w_i}{\|w_j - w_i\|} \frac{(w_j - w_i)^T Q_{j[i]}^{(k)\perp}}{\big\|Q_{j[i]}^{(k)\perp}(w_j - w_i)\big\|} Q_j^{(k)}w_j \right\| \\
&= \frac{1}{\sin\angle\big(w_j - w_i, \mathcal{U}_{j[i]}^{(k)}\big)} \left| \frac{(w_j - w_i)^T Q_{j[i]}^{(k)\perp}}{\big\|Q_{j[i]}^{(k)\perp}(w_j - w_i)\big\|} Q_j^{(k)}w_j \right| \left\| \frac{w_j - w_i}{\|w_j - w_i\|} \right\| \\
&= \frac{1}{\sin\angle\big(w_j - w_i, \mathcal{U}_{j[i]}^{(k)}\big)} \left\| \frac{(w_j - w_i)^T Q_{j[i]}^{(k)\perp}}{\big\|Q_{j[i]}^{(k)\perp}(w_j - w_i)\big\|} \right\| \\
&\qquad \times \big\|Q_j^{(k)}w_j\big\| \cos\angle\big(Q_j^{(k)}w_j, Q_{j[i]}^{(k)\perp}(w_j - w_i)\big) \\
&= \frac{\sin\angle\big(Q_j^{(k)}w_j,\, Q_{j[i]}^{(k)}(w_j - w_i)\big)}{\sin\angle\big(w_j - w_i, \mathcal{U}_{j[i]}^{(k)}\big)} \cos\varphi_j^{(k)} \|w_j\|.
\end{aligned}$$

Thus, if we define

$$\eta_{ij}^{(k)} = \begin{cases} 1 & (m = 1), \\ \dfrac{\sin\angle\big(Q_j^{(k)}w_j, Q_{j[i]}^{(k)}(w_j - w_i)\big)}{\sin\angle\big(w_j - w_i, \mathcal{U}_{j[i]}^{(k)}\big)} & (m \geq 2), \end{cases} \tag{2.11}$$

then

$$\big\|\alpha_i^{(k)}(w_j - w_i)\big\| = \eta_{ij}^{(k)} \cos\varphi_j^{(k)} \|w_j\|. \tag{2.12}$$

(c) The inverse of a matrix sum and a norm inequality. Consider $A, B \in \mathbb{R}^{n\times n}$, where A is nonsingular and $\|B\|$ is sufficiently small such that $\|BA^{-1}\| \leq \|A^{-1}\|\|B\| < 1$. Then

$$\begin{aligned}
(A + B)^{-1} &= \big((I + BA^{-1})A\big)^{-1} = A^{-1}(I + BA^{-1})^{-1} \\
&= A^{-1}\left(I - BA^{-1}(I + BA^{-1})^{-1}\right) \\
&= A^{-1} - A^{-1}BA^{-1}(I + BA^{-1})^{-1},
\end{aligned} \tag{2.13}$$

where

$$\begin{aligned}
\|A^{-1}BA^{-1}(I + BA^{-1})^{-1}\| &\leq \|A^{-1}\|^2\|B\| \left\|(I + BA^{-1})^{-1}\right\| \\
&\leq \frac{\|A^{-1}\|^2\|B\|}{1 - \|A^{-1}\|\|B\|}.
\end{aligned} \tag{2.14}$$

The last inequality is based on $\|(I + C)^{-1}\| \leq \frac{1}{1-\|C\|}$ provided that $\|C\| < 1$; this can be derived without difficulty by the singular values of these matrices.

(d) The continuous dependence for least-squares problems. Finally, we highlight the fact that the result of a well-posed linear least-squares problem depends *continuously* on its data. Specifically, assume that $U \in \mathbb{R}^{n\times m}$ $(m < n)$ is of full column rank m and $w \in \mathbb{R}^n \backslash \{0\}$. Then,

$$f(U, w) = \min_y \|w - Uy\| = \big\|(I - U(U^TU)^{-1}U^T)w\big\|$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

is a continuous (in fact, differentiable) function of U and w. Therefore, there exist some positive constant C_u and C_w (which depend on U and w) such that

$$|f(U+\Delta U, w+\Delta w) - f(U,w)| \le C_u\|\Delta U\| + C_w\|\Delta w\| \tag{2.15}$$

for all ΔU and Δw sufficiently small in the norm.

In fact, one can obtain such C_u and C_w by evaluating $|f(U+\Delta U, w+\Delta w) - f(U,w)|$. For a given small constant $\epsilon \in (0,1)$, let us assume that $\|\Delta U\|$ is sufficiently small such that

$$(2\|U\|\|\Delta U\| + \|\Delta U\|^2)\|(U^TU)^{-1}\| \le \epsilon.$$

Using (2.13) and (2.14), with some algebraic work, we can show that

$$\begin{aligned}
&|f(U+\Delta U, w+\Delta w) - f(U,w)| \\
&\quad \le \big\| \big(I - (U+\Delta U)\big((U+\Delta U)^T(U+\Delta U)\big)^{-1}(U+\Delta U)^T\big)(w+\Delta w) \\
&\qquad\qquad - (I - U(U^TU)^{-1}U^T)w \big\| \\
&\quad \le 2\left(1 + \frac{\|(U^TU)^{-1}\|\|U\|^2}{1-\epsilon}\right)\|U(U^TU)^{-1}\|\|w\|\|\Delta U\| \\
&\qquad\qquad + \|(I - U(U^TU)^{-1}U^T)\|\|\Delta w\| + \mathcal{O}(\|\Delta U\|^2) + \mathcal{O}(\|\Delta U\|\|\Delta w\|).
\end{aligned}$$

Here, we only need a rough estimate of C_u and C_w. Let $\epsilon = \frac{1}{2}$, and adopt larger values for C_u and C_w to absorb the quadratic terms of $\|\Delta U\|$ and $\|\Delta w\|$ into the linear terms:

$$\begin{aligned}
C_u &= 3\left(1 + 2\|(U^TU)^{-1}\|\|U\|^2\right)\|U(U^TU)^{-1}\|\|w\|, \qquad \text{and} \\
C_w &= 2\|(I - U(U^TU)^{-1}U^T)\| = 2.
\end{aligned} \tag{2.16}$$

Note that C_u is proportional to $\|w\|$ and could be large if U is ill-conditioned, whereas C_w is bounded independent of U. Then (2.15) holds for all sufficiently small $\|\Delta U\|$ and $\|\Delta w\|$.

3. A new analysis of (exact) Anderson acceleration. In this section, we derive a one-step convergence analysis of Algorithm 1 under Assumption 1. This section serves as a basis to derive our subsequent analysis of the inexact Anderson acceleration under these assumptions, which are more relaxed than those adopted in [26, 27].

3.1. Linear convergence. With the above preliminaries, we can give our convergence analysis of Algorithm 1. We shall show that this algorithm typically converges linearly with a factor that could be significantly smaller than κ_g (see Assumption 1). We start with the expression of x_{k+1} defined by Algorithm 1:

$$\begin{aligned}
x_{k+1} &= (1-\beta)\sum_{j=k-m}^{k}\alpha_j^{(k)}x_j + \beta\sum_{j=k-m}^{k}\alpha_j^{(k)}g(x_j) \\
&= \sum_{j=k-m}^{k}\alpha_j^{(k)}x_j + \beta\sum_{j=k-m}^{k}\alpha_j^{(k)}\left(g(x_j) - x_j\right).
\end{aligned} \tag{3.1}$$

We subtract on both sides $x_k = \sum_{j=k-m}^{k}\alpha_j^{(k)}x_k$ (since $\sum_{j=k-m}^{k}\alpha_j^{(k)} = 1$) to obtain

$$x_{k+1} - x_k = \sum_{j=k-m}^{k-1}\alpha_j^{(k)}(x_j - x_k) + \beta\sum_{j=k-m}^{k}\alpha_j^{(k)}w_j. \tag{3.2}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Then we multiply both sides of (3.2) with $\int_0^1 g'\left(z_{k(k+1)}(t)\right) dt$ from the right and get

(3.3)
$$\begin{aligned} g(x_{k+1}) - g(x_k) &= \int_0^1 g'\left(z_{k(k+1)}(t)\right)(x_{k+1} - x_k)\, dt \\ &= \sum_{j=k-m}^{k-1} \alpha_j^{(k)} \int_0^1 g'\left(z_{k(k+1)}(t)\right)(x_j - x_k) dt \\ &\qquad + \beta \int_0^1 g'\left(z_{k(k+1)}(t)\right)\left(\sum_{j=k-m}^{k} \alpha_j^{(k)} w_j\right) dt \\ &= \sum_{j=k-m}^{k-1} \alpha_j^{(k)} \int_0^1 g'(z_{kj}(t))(x_j - x_k) dt \\ &\qquad + \beta \int_0^1 g'\left(z_{k(k+1)}(t)\right)\left(\sum_{j=k-m}^{k} \alpha_j^{(k)} w_j\right) dt \\ &\qquad + \sum_{j=k-m}^{k-1} \alpha_j^{(k)} \int_0^1 \left(g'\left(z_{k(k+1)}(t)\right) - g'(z_{kj}(t))\right)(x_j - x_k) dt \\ &= \sum_{j=k-m}^{k-1} \alpha_j^{(k)} \left(g(x_j) - g(x_k)\right) \\ &\qquad + \beta \int_0^1 g'\left(z_{k(k+1)}(t)\right)\left(\sum_{j=k-m}^{k} \alpha_j^{(k)} w_j\right) dt + q_k, \end{aligned}$$

where the quadratic term is

$$q_k = \sum_{j=k-m}^{k-1} \alpha_j^{(k)} \int_0^1 \left(g'\left(z_{k(k+1)}(t)\right) - g'(z_{kj}(t))\right)(x_j - x_k) dt. \tag{3.4}$$

Adding $g(x_k)$ on both sides of (3.3) and using $1 - \sum_{j=k-m}^{k-1} \alpha_j^{(k)} = \alpha_k^{(k)}$, we have

$$g(x_{k+1}) = \sum_{j=k-m}^{k} \alpha_j^{(k)} g(x_j) + \beta \int_0^1 g'\left(z_{k(k+1)}(t)\right)\left(\sum_{j=k-m}^{k} \alpha_j^{(k)} w_j\right) dt + q_k. \tag{3.5}$$

Then we subtract (3.5) from (3.1) and obtain

$$\begin{aligned} g(x_{k+1}) - x_{k+1} &= (1-\beta) \sum_{j=k-m}^{k} \alpha_j^{(k)} w_j \\ &\qquad + \beta \int_0^1 g'\left(z_{k(k+1)}(t)\right)\left(\sum_{j=k-m}^{k} \alpha_j^{(k)} w_j\right) dt + q_k, \end{aligned} \tag{3.6}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

where the first two (linear) terms on the right-hand side satisfy

$$\begin{aligned}\left\|(1-\beta)\sum_{j=k-m}^{k}\alpha_j^{(k)}w_j+\beta\int_0^1 g'\big(z_{k(k+1)}(t)\big)\left(\sum_{j=k-m}^{k}\alpha_j^{(k)}w_j\right)dt\right\|&\\ \le\left\|(1-\beta)I+\beta\int_0^1 g'\big(z_{k(k+1)}(t)\big)\,dt\right\|\left\|\sum_{j=k-m}^{k}\alpha_j^{(k)}w_j\right\|&\\ \le((1-\beta)+\beta\kappa_g)\sin\varphi_k^{(k)}\|w_k\|.&\end{aligned}$$

To sum up, the above derivation shows that

$$\|w_{k+1}\|\le((1-\beta)+\beta\kappa_g)\sin\varphi_k^{(k)}\|w_k\|+\|q_k\|. \tag{3.7}$$

Suppose that the quadratic term $\|q_k\|$ is much smaller than the linear term involving $\|w_k\|$. Then this means that Algorithm 1 converges at least linearly with a factor not larger than

$$\widehat{\kappa}_g^{(k,m)}:=((1-\beta)+\beta\kappa_g)\sin\varphi_k^{(k)}\le(1-\beta)+\beta\kappa_g.$$

Note that κ_g is an upper bound for the factor of convergence of the simple fixed point iteration $x_{k+1}=g(x_k)$. Since $\min\{1,\kappa_g\}\le(1-\beta)+\beta\kappa_g\le\max\{1,\kappa_g\}$, the "acceleration" capability comes from the factor $\theta_k=\sin\varphi_k^{(k)}$, which could be much smaller than 1.

3.2. The quadratic term. Let us now investigate the quadratic term q_k. The expression (3.4) for q_k shows that we need to bound $g'(z_{k(k+1)}(t))-g'(z_{kj}(t))$ and $\alpha_j^{(k)}(x_j-x_k)$ $(k-m\le j\le k-1)$ in the norm. Since g' satisfies a Lipschitz condition, we obtain

$$\big\|g'\big(z_{k(k+1)}(t)\big)-g'(z_{kj}(t))\big\|\le L_g\big\|z_{k(k+1)}(t)-z_{kj}(t)\big\|=tL_g\|x_{k+1}-x_j\|, \tag{3.8}$$

where

$$x_{k+1}-x_j=\sum_{i=k-m,i\ne j}^{k}\alpha_i^{(k)}(x_i-x_j)+\beta\sum_{i=k-m}^{k}\alpha_i^{(k)}w_i \tag{3.9}$$

by (3.1) for each j $(k-m\le j\le k)$.

It then follows from (3.9), (2.3), and (2.12) that

$$\begin{aligned}\|x_{k+1}-x_j\|&=\Big\|\sum_{i=k-m,i\ne j}^{k}\alpha_i^{(k)}(x_i-x_j)+\beta\sum_{i=k-m}^{k}\alpha_i^{(k)}w_i\Big\|\\ &\le\Big\|\sum_{i=k-m,i\ne j}^{k}\Big(\int_0^1 g'(z_{ji}(t))dt-I\Big)^{-1}\big(\alpha_i^{(k)}(w_i-w_j)\big)\Big\|+\beta\Big\|\sum_{i=k-m}^{k}\alpha_i^{(k)}w_i\Big\|\\ &\le\sum_{i=k-m,i\ne j}^{k}\Big\|\Big(\int_0^1 g'(z_{ji}(t))dt-I\Big)^{-1}\Big\|\big\|\alpha_i^{(k)}(w_i-w_j)\big\|+\beta\sin\varphi_k^{(k)}\|w_k\|\\ &\le\frac{1}{\sigma_g}\sum_{i=k-m,i\ne j}^{k}\eta_{ij}^{(k)}\cos\varphi_j^{(k)}\|w_j\|+\beta\sin\varphi_k^{(k)}\|w_k\|.\end{aligned} \tag{3.10}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

We then substitute the above relations into (3.8) and then into (3.4) to obtain

$$\begin{aligned}
\|q_k\| &\le \sum_{j=k-m}^{k-1} \Big\| \int_0^1 \big(g'(z_{k(k+1)}(t)) - g'(z_{kj}(t))\big)\, dt \Big\| \big\|\alpha_j^{(k)}(x_j - x_k)\big\| \\
&\le \sum_{j=k-m}^{k-1} \left(\int_0^1 tL_g \|x_{k+1} - x_j\| dt \right) \frac{1}{\sigma_g} \big\|\alpha_j^{(k)}(w_j - w_k)\big\| \quad \text{(see (2.3) and Ass. 1)} \\
&\le \sum_{j=k-m}^{k-1} \left(\int_0^1 tL_g \|x_{k+1} - x_j\| dt \right) \frac{\eta_{jk}^{(k)}}{\sigma_g} \cos\varphi_k^{(k)} \|w_k\| \\
&\le \sum_{j=k-m}^{k-1} \frac{L_g \eta_{jk}^{(k)} \cos\varphi_k^{(k)}}{2\sigma_g} \left(\frac{1}{\sigma_g} \sum_{i=k-m, i\ne j}^{k} \eta_{ij}^{(k)} \cos\varphi_j^{(k)} \|w_j\| \right. \\
&\qquad \left. + \beta \sin\varphi_k^{(k)} \|w_k\| \right) \|w_k\|.
\end{aligned}$$

For $m \ge 2$, by (2.11), if $\sin\angle\big(w_j - w_i, \mathcal{U}_{j[i]}^{(k)}\big)$ is small for *any* pair of indices (i,j) ($k-m \le j \le k-1$ and $k-m \le i \le k$, $i \ne j$), then $\eta_{ij}^{(k)}$ is large, and the quadratic term $\|q_k\|$ will contain a large quadratic term $\mathcal{O}\big(\|w_j\|\|w_k\|\big)$ for this index j, unless $\eta_{jk}^{(k)}$ is very small.

Fortunately, if all w_i ($k-m \le i \le k$) are sufficiently small in the norm, such a potential issue can be alleviated. To show this, let us define

$$A_j = \int_0^1 g'(z_{j*}(t))dt - I, \qquad B_{ij} = \int_0^1 \big(g'(z_{ji}(t)) - g'(z_{j*}(t))\big)\, dt, \qquad \text{and}$$
$$C_{ij} = A_j^{-1} B_{ij} A_j^{-1} (I + B_{ij} A_j^{-1})^{-1},$$

such that, by (2.13),

$$\begin{aligned}
\Big(\int_0^1 g'(z_{ji}(t))dt - I \Big)^{-1} &= (A_j + B_{ij})^{-1} \\
&= A_j^{-1} - A_j^{-1} B_{ij} A_j^{-1} (I + B_{ij} A_j^{-1})^{-1} = A_j^{-1} - C_{ij}.
\end{aligned} \tag{3.11}$$

Let us assume that $\|w_i\| < \frac{2\sigma_g^2}{L_g}$. Then by (2.2),

$$\|B_{ij}\| = \Big\| \int_0^1 \big(g'(z_{ji}(t)) - g'(z_{j*}(t))\big) dt \Big\| \le \frac{L_g}{2\sigma_g} \|w_i\|, \quad \text{and}$$
$$\big\|B_{ij} A_j^{-1}\big\| \le \|B_{ij}\| \big\|A_j^{-1}\big\| \le \frac{L_g}{2\sigma_g} \|w_i\| \cdot \frac{1}{\sigma_g} < 1.$$

Consequently, by (2.14),

$$\begin{aligned}
\|C_{ij}\| &= \big\|A_j^{-1} B_{ij} A_j^{-1} (I + B_{ij} A_j^{-1})^{-1}\big\| \\
&\le \frac{\big\|A_j^{-1}\big\|^2 \|B_{ij}\|}{1 - \big\|A_j^{-1}\big\| \|B_{ij}\|} \le \frac{\frac{1}{\sigma_g^2} \cdot \frac{L_g}{2\sigma_g} \|w_i\|}{1 - \frac{1}{\sigma_g} \cdot \frac{L_g}{2\sigma_g} \|w_i\|} = \frac{1}{\sigma_g} \frac{L_g \|w_i\|}{2\sigma_g^2 - L_g \|w_i\|}.
\end{aligned} \tag{3.12}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

To simplify the notation, we define $\nu_\beta^{(k)} = \beta\big\|\sum_{i=k-m}^{k}\alpha_i^{(k)}w_i\big\| = \beta\sin\varphi_k^{(k)}\|w_k\|$. From the first two lines in (3.10), by (3.11), (3.12), and (2.12), we have

$$\begin{aligned}
\|x_{k+1}-x_j\| &\le \Big\|\sum_{i=k-m,i\neq j}^{k}\Big(\int_0^1 g'(z_{ji}(t))dt - I\Big)^{-1}\big(\alpha_i^{(k)}(w_i-w_j)\big)\Big\| + \nu_\beta^{(k)}\\
&= \Big\|A_j^{-1}\sum_{i=k-m,i\neq j}^{k}\big(\alpha_i^{(k)}(w_i-w_j)\big) - \sum_{i=k-m,i\neq j}^{k}C_{ij}\alpha_i^{(k)}(w_i-w_j)\Big\| + \nu_\beta^{(k)}\\
&\le \|A_j^{-1}\|\Big\|\sum_{i=k-m,i\neq j}^{k}\big(\alpha_i^{(k)}(w_i-w_j)\big)\Big\|\\
&\qquad + \sum_{i=k-m,i\neq j}^{k}\|C_{ij}\|\big\|\alpha_i^{(k)}(w_i-w_j)\big\| + \nu_\beta^{(k)}\\
&\le \frac{1}{\sigma_g}\cos\varphi_j^{(k)}\|w_j\| + \sum_{i=k-m,i\neq j}^{k}\frac{1}{\sigma_g}\frac{L_g\|w_i\|}{2\sigma_g^2-L_g\|w_i\|}\big\|\alpha_i^{(k)}(w_i-w_j)\big\| + \nu_\beta^{(k)}\\
&\le \frac{1}{\sigma_g}\left(1+\sum_{i=k-m,i\neq j}^{k}\frac{L_g\eta_{ij}^{(k)}\|w_i\|}{2\sigma_g^2-L_g\|w_i\|}\right)\cos\varphi_j^{(k)}\|w_j\| + \nu_\beta^{(k)}.
\end{aligned}$$

Let us define $\gamma_{ij}^{(k)} = \frac{L_g\eta_{ij}^{(k)}\|w_i\|}{2\sigma_g^2-L_g\|w_i\|}$, and substitute the above inequality into (3.8) and (3.4) to obtain

$$\begin{aligned}
\|q_k\| &\le \sum_{j=k-m}^{k-1}\Big\|\int_0^1\big(g'(z_{k(k+1)}(t)) - g'(z_{kj}(t))\big)\,dt\Big\|\big\|\alpha_j^{(k)}(w_j-w_k)\big\|\\
&\le \sum_{j=k-m}^{k-1}\left(\int_0^1 tL_g\|x_{k+1}-x_j\|dt\right)\eta_{jk}^{(k)}\cos\varphi_k^{(k)}\|w_k\|\\
&\le \sum_{j=k-m}^{k-1}\frac{L_g\eta_{jk}^{(k)}\cos\varphi_k^{(k)}}{2}\left(\frac{1}{\sigma_g}\left(1+\sum_{i=k-m,i\neq j}^{k}\gamma_{ij}^{(k)}\right)\cos\varphi_j^{(k)}\|w_j\| + \nu_\beta^{(k)}\right)\|w_k\|.
\end{aligned}$$

The above results can be summarized in the following main theorem.

THEOREM 3.1. *Under Assumption 1, consider Algorithm 1, which generates the approximate solutions $x_0, x_1, \ldots$ and the residuals $w_k = g(x_k) - x_k$. With the subspaces $\mathcal{U}_j^{(k)}$ and $\mathcal{U}_{j[i]}^{(k)}$ defined in (2.5), the orthogonal projectors $Q_j^{(k)}$ and $Q_{j[i]}^{(k)}$ defined in (2.7), $\eta_{ij}^{(k)}$ defined in (2.11),*

$$\varphi_j^{(k)} = \angle\big(w_j, \mathcal{U}_j^{(k)}\big), \qquad \gamma_{ij}^{(k)} = \frac{L_g\eta_{ij}^{(k)}\|w_i\|}{2\sigma_g^2-L_g\|w_i\|}, \quad \textit{and} \quad \nu_\beta^{(k)} = \beta\sin\varphi_k^{(k)}\|w_k\|,$$

we have

$$\|w_{k+1}\| \le ((1-\beta)+\beta\kappa_g)\sin\varphi_k^{(k)}\|w_k\| + \|q_k\|,$$

where

$$\|q_k\| \le \sum_{j=k-m}^{k-1}\frac{L_g\eta_{jk}^{(k)}\cos\varphi_k^{(k)}}{2}\left(\frac{1}{\sigma_g}\sum_{i=k-m,i\neq j}^{k}\eta_{ij}^{(k)}\cos\varphi_j^{(k)}\|w_j\| + \nu_\beta^{(k)}\right)\|w_k\|. \tag{3.13}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

In addition, if $\|w_i\| < \frac{2\sigma_g^2}{L_g}$ *for all* $k-m \le i \le k$, *then*

$$\|q_k\| \le \sum_{j=k-m}^{k-1} \frac{L_g \eta_{jk}^{(k)} \cos\varphi_k^{(k)}}{2} \left(\frac{1}{\sigma_g} \left(1 + \sum_{i=k-m, i\neq j}^{k} \gamma_{ij}^{(k)} \right) \cos\varphi_j^{(k)} \|w_j\| + \nu_\beta^{(k)} \right) \|w_k\|.$$

REMARK 3.2. Compared to the classical fixed point iteration $x_{k+1} = g(x_k)$ with the convergence factor bounded by κ_g, Algorithm 1 has the smaller convergence factor $\big((1-\beta)+\beta\kappa_g\big)\sin\varphi_k^{(k)}$ (disregarding the high-order term q_k), which has been obtained in [17]. Our new insight here concerns the factors that influence q_k in the upper bound for $\|w_{k+1}\|$.

- From the discussion about the least-squares problem in Section 2.1 part (b), namely, (2.5), (2.6), and (2.11), we see that the condition number of $U_j^{(k)}$ $(k-m \le j \le k)$ has an impact on the one-step convergence of Anderson acceleration. If $U_j^{(k)}$ is well-conditioned, then all angles $\angle(w_j - w_i, \mathcal{U}_{j[i]}^{(k)})$ $(k-m \le i \le k, i \neq j)$ cannot be small, and hence all $\eta_{ij}^{(k)}$ in (2.11) will not be large, which guarantees that the bound on q_k (3.13) is modest. As a result, it is more likely to have $\|w_{k+1}\| < \|w_k\|$ if $\big((1-\beta)+\beta\kappa_g\big)\sin\varphi_k^{(k)} < 1$ than in the case of an ill-conditioned $U_j^{(k)}$.
- The upper bound for the higher-order term q_k given in [17] grows *exponentially* with the acceleration depth m, whereas there is no reason to assume that such a rapid growth is the common case for our bound (3.13). In practice, the condition number of $U_j^{(k)}$ does grow with m but apparently not as rapidly as the exponential pattern.
- Generally, a smaller L_g and a larger σ_g in Assumption 1 lead to a smaller $\|q_k\|$, which is of the form $\mathcal{O}(\|w_{k-m}\|\|w_k\|) + \ldots + \mathcal{O}(\|w_{k-1}\|\|w_k\|) + \mathcal{O}(\|w_k\|^2)$.
- An optimization with no gain at step k $(\cos\varphi_k^{(k)} = 0)$ sets $q_k = 0$, and one with maximum gain $(\sin\varphi_k^{(k)} = 0)$ eliminates $\mathcal{O}\left(\|w_k\|^2\right)$ but not the $\mathcal{O}\left(\|w_j\|\|w_k\|\right)$-terms in q_k (the latter point regarding $\sin\varphi_k^{(k)} = 0$ is not seen in [17]).
- Suppose that Algorithm 1 converges so that $\gamma_{ij}^{(k)} = \frac{L_g \eta_{ij}^{(k)} \|w_i\|}{2\sigma_g^2 - L_g\|w_i\|} \to 0$. Then, in the asymptotic phase of convergence, if $\sin\varphi_k^{(k)}\cos\varphi_k^{(k)}$ is not close to 0,
  - for $m = 1$, the quadratic terms are moderate since $\eta_{jk}^{(k)} = 1$;
  - for $m \ge 2$, if $\sin\angle(w_k - w_j, \mathcal{U}_{k[j]}^{(k)})$ is small for some index j $(k-m \le j \le k-1)$ and $\sin\angle(Q_k w_k, Q_{k[j]}^{(k)}(w_k - w_j))$ is not small, then $\eta_{jk}^{(k)}$ is large, and consequently the $\mathcal{O}(\|w_j\|\|w_k\|)$- and $\mathcal{O}(\|w_k\|^2)$-terms will be large in q_k. Whether this happens depends primarily on the direction of the vectors involved in the linear least-squares problem *at step* k, not so much on the previous steps because $\gamma_{ij}^{(k)} = \frac{L_g \eta_{ij}^{(k)} \|w_i\|}{2\sigma_g^2 - L_g\|w_i\|} \to 0$, assuming convergence.

We shall see in the next section that our above framework from the linear algebra perspectives based on projectors and angles between subspaces can easily accommodate the analysis of the effects of the inexact evaluation of each $g(x_k)$, whereas it is less clear how this can be done for the result of [17] due to its entry-wise or column vector-wise analysis of the QR factorization used to solve the minimization by the linear least-squares method.

4. Inexact Anderson acceleration.

Our main goal in this paper is the investigation of inexact Anderson acceleration, based on our analytic framework presented in Section 3. To this end, recall that given $m+1$ approximate solutions to $x = g(x)$, namely, $x_{k-m}, \ldots, x_{k-1}, x_k$,

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Algorithm 1 gives the new iterate

$$x_{k+1} = (1-\beta)\sum_{j=k-m}^{k} \alpha_j^{(k)} x_j + \beta \sum_{j=k-m}^{k} \alpha_j^{(k)} g(x_j),$$

where the coefficients $\{\alpha_j^{(k)}\}_{j=k-m}^{k}$ solve the minimization problem

$$\min_{\sum_{j=k-m}^{k} \alpha_j^{(k)}=1} \left\| \sum_{j=k-m}^{k} \alpha_j^{(k)} (g(x_j) - x_j) \right\|.$$

In this section, our goal is to show that each $g(x_j)$ can be computed approximately with some errors kept under control, without obviously impacting the one-step linear convergence factor $\big((1-\beta)+\beta\kappa_g\big)\sin\varphi_k^{(k)}$ of Algorithm 1; see (3.7).

Algorithm 2 Inexact Anderson acceleration for solving the nonlinear system $x = g(x)$.

Input: function $g : X \to X$, $x_0 \in X$, integer $m > 0$, $\beta \in (0, 1]$, a sufficiently small fixed constant $\tau \in (0, 1)$, and tolerance $\delta > 0$.
Output: an approximate solution x_k such that $x_k \approx g(x_k)$.

1: Compute $x_1 = \widehat{g}_0 \approx g(x_0)$ such that the error $\delta g_0 = \widehat{g}_0 - g(x_0)$ satisfies $\|\delta g_0\| \le \tau\|g(x_0) - x_0\|$ (achieved without exact evaluation of $g(x_0)$), and let $\widehat{w}_0 = \widehat{g}_0 - x_0$.
2: **for** $k = 1, 2, \ldots,$ **do**
3: Compute $\widehat{g}_k \approx g(x_k)$, such that $\delta g_k = \widehat{g}_k - g(x_k)$ satisfies $\|\delta g_k\| \le \tau\|g(x_k) - x_k\|$ (achieved without exact evaluation of $g(x_k)$), and let $\widehat{w}_k = \widehat{g}_k - x_k$.
4: **if** $\|\widehat{w}_k\| \le \delta$ **then**
5: terminate the algorithm and return x_k.
6: **end if**
7: Let $\ell = \max\{0, k-m\}$, and solve $\min_{\sum_{i=\ell}^{k} \widehat{\alpha}_i^{(k)}=1} \big\| \sum_{i=\ell}^{k} \widehat{\alpha}_i^{(k)} \widehat{w}_i \big\|$ for $\{\widehat{\alpha}_i^{(k)}\}$.
8: Evaluate the new iterate $x_{k+1} = (1-\beta)\sum_{j=\ell}^{k} \widehat{\alpha}_j^{(k)} x_j + \beta \sum_{j=\ell}^{k} \widehat{\alpha}_j^{(k)} \widehat{g}_j$.
9: **end for**

The inexact variant of Anderson acceleration is outlined in Algorithm 2, where the evaluation of $g(x_k)$ at each step k is performed only *approximately and never exactly*. Specifically, at step k, let $\widehat{g}_k \approx g(x_k)$ be the actual computed approximation to $g(x_k)$ such that the error $\delta g_k = \widehat{g}_k - g(x_k)$ satisfies $\|\delta g_k\| \le \tau\|g(x_k) - x_k\| = \tau\|w_k\|$, where $\tau \in (0, 1)$ is a sufficiently small constant predetermined and fixed throughout the algorithm. Since τ is fixed but the residual norm $\|w_k\|$ tends to decrease as the algorithm proceeds, so does $\|\delta g_k\|$. This means that $\widehat{g}_k$ should be an increasingly accurate approximation to $g(x_k)$ as the algorithm approaches convergence. The bound for $\|\delta g_k\|$ can be achieved without an exact evaluation of $g(x_k)$. The *computed* (approximate) nonlinear residual is

$$\widehat{w}_k = \widehat{g}_k - x_k = g(x_k) - x_k + \widehat{g}_k - g(x_k) = w_k + \delta g_k$$

such that

$$\|\widehat{w}_k - w_k\| = \|\delta g_k\| \le \tau\|w_k\|, \tag{4.1}$$

which should also be achieved without having the exact $g(x_k)$ or w_k at hand. Then we solve the perturbed least-squares problem $\min_{\sum_{j=k-m}^{k} \widehat{\alpha}_j^{(k)}=1} \big\| \sum_{j=k-m}^{k} \widehat{\alpha}_j^{(k)} \widehat{w}_j \big\|$ and construct

ETNA
Kent State University and
Johann Radon Institute (RICAM)

the new approximate solution

$$
\begin{aligned}
x_{k+1} &= (1-\beta)\sum_{j=k-m}^{k}\widehat{\alpha}_j^{(k)}x_j + \beta\sum_{j=k-m}^{k}\widehat{\alpha}_j^{(k)}\widehat{g}_j \\
&= \sum_{j=k-m}^{k}\widehat{\alpha}_j^{(k)}x_j + \beta\sum_{j=k-m}^{k}\widehat{\alpha}_j^{(k)}(\widehat{g}_j - x_j).
\end{aligned} \tag{4.2}
$$

The approximate solutions computed by both algorithms are denoted as $\{x_k\}$ because it should be clear from the context if they are obtained from $\{g(x_\ell)\}$ and $\{w_\ell\}$ or from $\{\widehat{g}_\ell\}$ and $\{\widehat{w}_\ell\}$. To compute the new iterate x_{k+1}, both algorithms use *their own* recent iterates $\{x_\ell\}_{\ell=k-m}^{k}$. Algorithm 1 needs the exact function evaluations $\{g(x_\ell)\}_{\ell=k-m}^{k}$ and the exact residuals $\{w_\ell\} = \{g(x_\ell) - x_\ell\}_{\ell=k-m}^{k}$, whereas Algorithm 2 uses the approximate evaluations $\{\widehat{g}_\ell\}_{\ell=k-m}^{k}$ and the approximate residuals $\{\widehat{w}_\ell\} = \{\widehat{g}_\ell - x_\ell\}_{\ell=k-m}^{k}$, where $\|\widehat{g}_k - g(x_k)\| = \|\widehat{w}_k - w_k\| \le \tau\|w_k\|$ for a small tolerance $\tau > 0$. Both methods need only one function evaluation (exact or approximate) at each step because they keep a record of $\{g(x_\ell)\}$ (or $\{\widehat{g}_\ell\}$) and $\{w_\ell\}$ (or $\{\widehat{w}_\ell\}$) ($k-m \le \ell \le k-1$) computed from the previous steps. The inexact algorithm does not compute, use, or rely on any information generated by the exact algorithm, including the recent iterates, exact function evaluations, and exact residuals.

In this section, we consider a *one-step* analysis of the convergence of Algorithm 2 and compare it with the one-step analysis of Algorithm 1 in the previous section. This is not a multi-step analysis of the residual w_k of Algorithm 2 that starts with the same initial iterate as Algorithm 1, which seems rather complex for a non-contractive mapping g, and this is beyond the scope of this paper (to the best of our knowledge, no multi-step analysis of Algorithm 1 for non-contractive g has been performed). To explore the one-step difference between Algorithms 1 and 2, we shall show that at step k, if the error control constant $\tau \in (0,1)$ is sufficiently small, the true residual $\|w_{k+1}\|$ of the new iterate x_{k+1} from the inexact method (not to be computed in practice) would be sufficiently close to the counterpart of the exact method, assuming that Algorithms 1 and 2 have the same set of previous iterates $x_{k-m}, x_{k-m+1}, \ldots, x_k$ but different corresponding function evaluations ($\{g(x_\ell)\}$ or $\widehat{g}_\ell$) and residuals ($\{w_\ell\}$ or $\{\widehat{w}_\ell\}$).

To this end, define a new block of vectors, associated subspaces, and projectors:

$$
\begin{aligned}
W_{[j]}^{(k)} &= [\widehat{w}_k, \ldots, \widehat{w}_{j+1}, \widehat{w}_{j-1}, \ldots, \widehat{w}_{k-m}], & \mathcal{W}_{[j]}^{(k)} &= \operatorname{col}\big(W_{[j]}^{(k)}\big), \\
S_{[j]}^{(k)} &= W_{[j]}^{(k)}\big(W_{[j]}^{(k)T}W_{[j]}^{(k)}\big)^{-1}W_{[j]}^{(k)T}, \quad \text{and} & S_{[j]}^{(k)\perp} &= I - S_{[j]}^{(k)}.
\end{aligned} \tag{4.3}
$$

Here, $S_{[j]}^{(k)\perp}$ is the orthogonal projector along $\mathcal{W}_{[j]}^{(k)}$ onto the orthogonal complement of $\mathcal{W}_{[j]}^{(k)}$. We also define the oblique projector

$$
T_{[j]}^{(k)} = \frac{\widehat{w}_j\widehat{w}_j^T S_{[j]}^{(k)\perp}}{\widehat{w}_j^T S_{[j]}^{(k)\perp}\widehat{w}_j} = \frac{1}{\sin\angle(\widehat{w}_j, S_{[j]}^{(k)}\widehat{w}_j)}\frac{\widehat{w}_j}{\|\widehat{w}_j\|}\frac{\widehat{w}_j^T S_{[j]}^{(k)\perp}}{\|\widehat{w}_j^T S_{[j]}^{(k)\perp}\|}.
$$

It is not difficult to see that $\operatorname{range}\big(T_{[j]}^{(k)}\big) = \operatorname{span}\{\widehat{w}_j\}$, $\mathcal{W}_{[j]}^{(k)} \subset \operatorname{null}\big(T_{[j]}^{(k)}\big)$, and that

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$\widehat{\alpha}_j^{(k)}\widehat{w}_j = T_{[j]}^{(k)}\left(\sum_{\ell=k-m}^{k}\widehat{\alpha}_\ell^{(k)}\widehat{w}_\ell\right)$. It follows that

$$
\begin{aligned}
\|\widehat{\alpha}_j^{(k)}\widehat{w}_j\| &= \left\| T_{[j]}^{(k)}\left(\sum_{\ell=k-m}^{k}\widehat{\alpha}_\ell^{(k)}\widehat{w}_\ell\right)\right\| \\
&= \frac{1}{\sin\angle(\widehat{w}_j,\, S_{[j]}^{(k)}\widehat{w}_j)}\left|\frac{\widehat{w}_j^T S_{[j]}^{(k)\perp}\left(\sum_{\ell=k-m}^{k}\widehat{\alpha}_\ell^{(k)}\widehat{w}_\ell\right)}{\left\|\widehat{w}_j^T S_{[j]}^{(k)\perp}\right\|}\right|\left\|\frac{\widehat{w}_j}{\|\widehat{w}_j\|}\right\| \\
&= \frac{1}{\sin\angle(\widehat{w}_j,\, S_{[j]}^{(k)}\widehat{w}_j)}\left\|\frac{\widehat{w}_j^T S_{[j]}^{(k)\perp}}{\|\widehat{w}_j^T S_{[j]}^{(k)\perp}\|}\right\| \\
&\qquad\qquad \times\left\|\sum_{\ell=k-m}^{k}\widehat{\alpha}_\ell^{(k)}\widehat{w}_\ell\right\|\cos\angle\big(S_{[j]}^{(k)\perp}\widehat{w}_j,\, \sum_{\ell=k-m}^{k}\widehat{\alpha}_\ell^{(k)}\widehat{w}_\ell\big) \\
&= \frac{\sin\widehat{\varphi}_k^{(k)}\|\widehat{w}_k\|}{\sin\angle(\widehat{w}_j,\, S_{[j]}^{(k)}\widehat{w}_j)}\cos\angle\big(S_{[j]}^{(k)\perp}\widehat{w}_j,\, \sum_{\ell=k-m}^{k}\widehat{\alpha}_\ell^{(k)}\widehat{w}_\ell\big) \\
&\le \frac{\sin\widehat{\varphi}_k^{(k)}}{\sin\angle(\widehat{w}_j,\, S_{[j]}^{(k)}\widehat{w}_j)}\|\widehat{w}_k\|,
\end{aligned}
\tag{4.4}
$$

where $\widehat{\varphi}^{(k)} = \angle(\widehat{w}_k, \widehat{\mathcal{U}}_k^{(k)})$, with $\widehat{\mathcal{U}}_k^{(k)} = \text{range}\,([\widehat{w}_k - \widehat{w}_{k-1}, \ldots, \widehat{w}_k - \widehat{w}_{k-m}])$.

Now we let $z_{k(k+1)} = x_k + (x_{k+1} - x_k)t$. Starting with (4.2), one can follow the derivations from (3.2) through (3.6) to obtain

$$
\begin{aligned}
&g(x_{k+1}) - x_{k+1} \\
&\quad = (1-\beta)\Big(\sum_{j=k-m}^{k}\widehat{\alpha}_j^{(k)}\widehat{w}_j\Big) \\
&\qquad + \beta\int_0^1 g'(z_{k(k+1)}(t))\Big(\sum_{j=k-m}^{k}\widehat{\alpha}_j^{(k)}\widehat{w}_j\Big)dt + \sum_{j=k-m}^{k}\widehat{\alpha}_j^{(k)}(w_j - \widehat{w}_j) + \widehat{q}_k,
\end{aligned}
\tag{4.5}
$$

where the quadratic term is

$$
\widehat{q}_k = \sum_{j=k-m}^{k}\widehat{\alpha}_j^{(k)}\int_0^1 \big(g'(z_{k(k+1)}(t) - g'(z_{kj}(t))\big)(x_j - x_k)dt.
$$

To analyze the inexact algorithm, we make the following assumption.

ASSUMPTION 2. *Assume that Algorithm 1 converges linearly but not with a higher order, so that there exists a constant $\mu \in (0,1)$ such that $\|w_{j+1}\| \ge \mu\|w_j\|$ for all j.*

Under such an assumption, we have $\|w_{k-\ell}\| \le \mu^{-\ell}\|w_k\|$ for all ℓ $(1 \le \ell \le k)$. Note that under certain circumstances, Algorithm 1 may exhibit asymptotic superlinear convergence; see, e.g., a discussion in [21]. This seems consistent with the superlinear convergence of GMRES for solving linear systems of equations, which occurs when convergence-delaying eigenvalues of the coefficient matrix have been 'resolved' after sufficiently many steps, so that the effective spectrum shrinks as the GMRES iteration proceeds; see, e.g., [14, 22] and the references therein. In practice, however, superlinear convergence might not be the most typical behavior of Algorithm 1 if the underlying fixed point iteration $x_{k+1} = g(x_k)$ converges no

ETNA
Kent State University and
Johann Radon Institute (RICAM)

more rapidly than linearly near the solution: In [21] it is claimed that the general experience with DIIS was a typical lack of superlinear convergence; see [21, Section 3.3, Figure 3]. In a recent study of Algorithm 1 for accelerating the Picard iteration to compute steady-state solutions of the incompressible Navier-Stokes equations [18], no superlinear convergence was observed.

We recall that $\big\| \sum_{j=k-m}^{k} \alpha_j^{(k)} w_j \big\|$ and $\big\| \sum_{j=k-m}^{k} \widehat{\alpha}_j^{(k)} \widehat{w}_j \big\|$ are the results of the original and the perturbed linear least-squares problem, respectively. In fact, in (2.15), if we let

$$\begin{aligned} w &= w_k, & U &= U_k^{(k)} = [w_k - w_{k-1}, w_k - w_{k-2}, \ldots, w_k - w_{k-m}], \\ \Delta w &= \delta g_k, & \Delta U &= [\delta g_k - \delta g_{k-1}, \delta g_k - \delta g_{k-2}, \ldots, \delta g_k - \delta g_{k-m}], \end{aligned} \tag{4.6}$$

then

$$f(U, w) = \big\| \sum_{j=k-m}^{k} \alpha_j^{(k)} w_j \big\| \qquad \text{and} \qquad f(U + \Delta U, w + \Delta w) = \big\| \sum_{j=k-m}^{k} \widehat{\alpha}_j^{(k)} \widehat{w}_j \big\|.$$

Under Assumption 2, for sufficiently small $\|\Delta U\|$ and $\|\Delta w\|$, by (2.15), we have

$$\begin{aligned} \big\| \sum_{j=k-m}^{k} \widehat{\alpha}_j^{(k)} \widehat{w}_j \big\| &\le \big\| \sum_{j=k-m}^{k} \alpha_j^{(k)} w_j \big\| + C_u^{(k)} \|\Delta U\| + C_w \|\Delta w\| \\ &\le \sin \varphi_k^{(k)} \|w_k\| + C_u^{(k)} \big(m \|\delta g_k\| + \sum_{j=k-m}^{k-1} \|\delta g_j\| \big) + C_w \|\delta g_k\| \\ &\le \sin \varphi_k^{(k)} \|w_k\| + (m C_u^{(k)} + C_w) \tau \|w_k\| + C_u^{(k)} \tau \sum_{j=k-m}^{k-1} \|w_j\| \\ &\le \sin \varphi_k^{(k)} \|w_k\| + (m C_u^{(k)} + C_w) \tau \|w_k\| + C_u^{(k)} \tau \sum_{\ell=1}^{m} \mu^{-\ell} \|w_k\| \\ &\le \left(\sin \varphi_k^{(k)} + \left((m + \frac{\mu^{-m} - 1}{1 - \mu}) C_u^{(k)} + C_w \right) \tau \right) \|w_k\|. \end{aligned} \tag{4.7}$$

Here, we use a superscript k for C_u but not for C_w because, as shown in (2.16), C_u depends on U (and on the iteration count k) but C_w can be bounded uniformly in all iterations.

To continue, note that (4.1) leads to $\|w_j\| - \|\widehat{w}_j\| \le \|w_j - \widehat{w}_j\| \le \tau \|w_j\|$. For $\tau \in [0, 1)$, we have $\|w_j\| \le \frac{1}{1-\tau} \|\widehat{w}_j\|$. Similarly, $\|\widehat{w}_j\| - \|w_j\| \le \tau \|w_j\|$. It follows that

$$\|w_j - \widehat{w}_j\| \le \tau \|w_j\| \le \frac{\tau}{1 - \tau} \|\widehat{w}_j\| \qquad \text{and} \qquad \|\widehat{w}_j\| \le (1 + \tau) \|w_j\|. \tag{4.8}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Starting with (4.5), by (4.1), (4.4), (4.7), and (4.8), we have for Algorithm 2 that

$$
\begin{aligned}
&\|g(x_{k+1}) - x_{k+1}\| \\
&\le ((1-\beta) + \beta\kappa_g) \Big\| \sum_{j=k-m}^{k} \widehat{\alpha}_j^{(k)} \widehat{w}_j \Big\| + \sum_{j=k-m}^{k} \big\| \widehat{\alpha}_j^{(k)} (w_j - \widehat{w}_j) \big\| + \|\widehat{q}_k\| \\
&\le ((1-\beta) + \beta\kappa_g) \left(\sin \varphi_k^{(k)} + \left(\big(m + \frac{\mu^{-m}-1}{1-\mu}\big) C_u^{(k)} + C_w \right) \tau \right) \|w_k\| \\
&\qquad + \frac{\tau}{1-\tau} \sum_{j=k-m}^{k} \|\widehat{\alpha}_j^{(k)} \widehat{w}_j\| + \|\widehat{q}_k\| \qquad (4.9) \\
&\le ((1-\beta) + \beta\kappa_g) \left(\sin \varphi_k^{(k)} + \zeta_1^{(k)} \tau \right) \|w_k\| \\
&\qquad + \frac{\tau}{1-\tau} \sum_{j=k-m}^{k} \frac{\sin \widehat{\varphi}_k^{(k)}}{\sin \angle(\widehat{w}_j, S_{[j]}^{(k)} \widehat{w}_j)} \|\widehat{w}_k\| + \|\widehat{q}_k\| \\
&\le \left(((1-\beta) + \beta\kappa_g) \left(\sin \varphi_k^{(k)} + \zeta_1^{(k)} \tau \right) + \frac{\tau(1+\tau)}{1-\tau} \zeta_2^{(k)} \right) \|w_k\| + \|\widehat{q}_k\|,
\end{aligned}
$$

where

$$
\zeta_1^{(k)} = \left(m + \frac{\mu^{-m}-1}{1-\mu} \right) C_u^{(k)} + C_w, \quad \zeta_2^{(k)} = \sum_{j=m-k}^{k} \frac{\sin \widehat{\varphi}_k^{(k)}}{\sin \angle(\widehat{w}_j, S_{[j]}^{(k)} \widehat{w}_j)}. \qquad (4.10)
$$

Assume that $\zeta_1^{(k)}$ and $\zeta_2^{(k)}$ are not large for a fixed m at each step k. This assumption is valid if both $U_k^{(k)} = [w_k - w_{k-1}, \ldots, w_k - w_{k-m}]$ and $[\widehat{w}_k, \ldots, \widehat{w}_{k-m}]$ are not ill-conditioned; see the definitions of C_u in (2.16), $S_{[j]}^{(k)}$ in (4.3), and (4.10). Since

$$
\lim_{\tau \to 0^+} \left(((1-\beta) + \beta\kappa_g) \left(\sin \varphi_k^{(k)} + \zeta_1^{(k)} \tau \right) + \frac{\tau(1+\tau)}{1-\tau} \zeta_2^{(k)} \right) = (1 - \beta + \beta\kappa_g) \sin \varphi_k^{(k)},
$$

the one-step convergence of Algorithm 2 would be sufficiently close to that of Algorithm 1 if both algorithms had the same set of recent iterates $x_{k-m}, x_{k-m+1}, \ldots, x_k$ and the relative tolerance $\tau \in (0,1)$ is sufficiently small. The quadratic term $\|\widehat{q}_k\|$ can be analyzed in detail as given in Theorem 3.1, but we will not further explore it here as it is supposed to have limited impact on the overall linear convergence when the algorithm is nearly convergent with very small residual norms. The major result of this section is as follows.

THEOREM 4.1. *Suppose that $g : X \to X$ satisfies Assumption 1 and Algorithm 1 satisfies Assumption 2. At a given step k, assume that in Algorithm 2 each $\widehat{g}_j \approx g(x_j)$ and the corresponding computed residuals $\widehat{w}_j = \widehat{g}_j - x_j$ satisfy*

$$
\|\widehat{w}_j - w_j\| = \|\widehat{g}_j - g(x_j)\| \le \tau \|w_j\| \qquad (k-m \le j \le k)
$$

for some small relative tolerance $\tau \in (0,1)$. Then, for the new iterate x_{k+1} of Algorithm 2 defined in (4.2), its true residual norm $\|g(x_{k+1}) - x_{k+1}\|$ satisfies (4.9) with $\zeta_1^{(k)}$ and $\zeta_2^{(k)}$ defined in (4.10). Assume that $\zeta_1^{(k)}$ and $\zeta_2^{(k)}$ are bounded at step k. With the same set of recent iterates $x_{k-m}, x_{k-m+1}, \ldots, x_k$, as $\tau \to 0^+$, the one-step convergence of Algorithm 2 is the same as Algorithm 1.

REMARK 4.2. From (2.16), (4.6), and (4.10), we conclude from Theorem 4.1 that the condition number of $U = U_k^{(k)}$ has an impact on the rate of convergence of the inexact

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Anderson acceleration. If $U_k^{(k)}$ is well-conditioned, then $C_u^{(k)}$ is small and so is $\zeta_2^{(k)}$. As a result, the inexact algorithm tends to exhibit one-step convergence that is closer to that of the exact method. If $U_k^{(k)}$ is ill-conditioned, then we may need a smaller tolerance τ to keep $\zeta_2^{(k)}$ under control to have the inexact algorithm match the behavior of the exact counterpart.

5. Numerical experiments. In this section, we solve a few nonlinear problems to illustrate the performance of exact and inexact Anderson acceleration. The behavior of the exact variant has been shown extensively in quite a few different problem settings; see, e.g., [13, 17, 21, 28] and the references therein. Our focus is to show that the inexact variant can exhibit essentially the same convergence behavior as the exact variant, whereas the former requires considerably less computational costs for evaluating $g(x_k)$. Our focus is on the comparison of Algorithms 1 and 2 for non-contractive mapping g because these problems are more challenging (the iteration $x_{k+1} = g(x_k)$ would not converge without Anderson acceleration), and there are no similar results for such a comparison in the literature to the best of our knowledge.

We let $\beta = 1$ in Algorithms 1 and 2 for all tests. Our experiments were performed in MATLAB R2018b, on a Macbook Pro with operating system OS X 10.11.6, a 2.9 GHz dual-core Intel Core i5 CPU, and 16 GB 1867 MHz DDR3 memory. The experiments were done in double precision, yet they could be done in single precision and may exhibit similar results if the matrices $U_k^{(k)} = [w_k - w_{k-1}, \ldots, w_k - w_{k-m}]$ and $[\widehat{w}_k, \ldots, \widehat{w}_{k-m}]$ are not ill-conditioned at each step k, which guarantees that $\zeta_1^{(k)}$ and $\zeta_2^{(k)}$ in (4.10) are modest.

Example 1. We seek the steady-state solution of a 1-D Burgers' equation

$$\begin{aligned} &u_t + uu_x = \nu u_{xx}, \qquad 0 < x < 1, \\ &u(0) = a, u(1) = b. \end{aligned} \tag{5.1}$$

In our experiments, we let $a = -1$, $b = 3$, and $\nu = 5 \times 10^{-6}$. To obtain a numerical steady-state solution of (5.1), we apply the finite difference discretization to (5.1) with $n = 2^{16}$ equispaced subintervals of length $h = \frac{1}{n}$ on $[0, 1]$, using the standard 2-point and 3-point centered differences to approximate the first- and second-order derivatives, respectively. Since u_t vanishes, we choose the Picard iteration as

$$u^{(k)} u_x^{(k)} = \nu u_{xx}^{(k+1)},$$

though other options are also possible. Let $u_j^{(k)}$ $(0 \le j \le n)$ be the approximation of $u(jh)$ at the k-th step of the Picard iteration, and define the vector $u^{(k)} = [u_1^{(k)}, u_2^{(k)}, \ldots, u_{n-1}^{(k)}]^T \in \mathbb{R}^{n-1}$ and the tridiagonal matrices

$$D = \frac{1}{2h}\mathrm{trig}[-1, 0, 1] \qquad \text{and} \qquad L = \frac{1}{h^2}\mathrm{trig}[-1, 2, -1] \in \mathbb{R}^{(n-1)\times(n-1)}.$$

This leads to

$$\begin{aligned} &\mathrm{diag}(u^{(k)}) \left(Du^{(k)} + \frac{1}{2h}[-u_0, 0, \ldots, 0, u_n]^T \right) \\ &\qquad = \nu \left(-Lu^{(k+1)} + \frac{1}{h^2}[u_0, 0, \ldots, 0, u_n]^T \right), \end{aligned}$$

which defines the relation $u^{(k+1)} = g(u^{(k)})$. We let $u_j^{(0)} = -4\cos(9\pi jh) - 4jh + 3$, which satisfies the boundary condition $u_0 = u(0) = -1$ and $u_n = u(1) = 3$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Note that with $\nu = 5\times10^{-6}$, the Picard iteration above defines a *non-contractive* mapping g near the desired fixed point u^*. The non-contractiveness has been verified numerically by running the Picard iteration with the initial approximation $u^{(0)}$ described above and with the approximate fixed point solution u^* (found by Anderson acceleration with $m \geq 2$), both of which quickly lead to a blowup in $u^{(k)}$.

The evaluation of $u^{(k+1)}$ from $u^{(k)}$ requires the solution of a linear system of the form $Lu^{(k+1)} = f$, where L is symmetric and positive definite. Solving this linear system is quite straightforward by a direct linear solver (e.g., MATLAB's backslash operator), but for illustration purposes here, we solve it iteratively by the preconditioned steepest descent (PSD) method with the preconditioner $M = \mathrm{trig}[-1, 2+10^{-7}, -1]$, whose action can be performed by a Cholesky factorization.

TABLE 5.1
Performance of exact and inexact Anderson acceleration for solving a 1-D nonlinear Burgers' equation.

	AA progress			total PSD iterations		
(ν, m)	$\max \kappa_2(U_k^{(k)})$	exact	inexact	exact	inexact	improvement
$(5\times10^{-6}, 2)$	1.61×10^{2}	27	27	7154	4184	41.5%
$(5\times10^{-6}, 4)$	2.26×10^{4}	21	21	5640	2690	52.3%
$(5\times10^{-6}, 8)$	2.18×10^{5}	17	19	4556	2512	44.9%

The exact Anderson acceleration uses a relative tolerance 10^{-9} for the PSD solves, whereas for the inexact variant the PSD tolerance at the k-th step of Anderson acceleration is specified to be $\max\left\{10^{-9}, \min\{10^{-5}, 10^{-5}\|g(u^{(k-1)}) - u^{(k-1)}\|_{\ell_2}\}\right\}$, where the ℓ_2-norm $\|u\|_{\ell_2} := \left(\sum_{j=1}^{n-1} u_j^2 h\right)^{1/2}$ of the vector $u \in \mathbb{R}^{n-1}$, with $u_j \approx u(jh)$, is an approximation of $\|u(x)\|_{L_2} = \left(\int_0^1 u^2(x)dx\right)^{1/2}$. This sets the PSD tolerance to be proportional to the nonlinear residual norm at the previous Anderson acceleration step, but also require it to be bounded between 10^{-5} and 10^{-9}. Both methods are terminated once $\|g(u^{(k)}) - u^{(k)}\|_{\ell_2}$ drops below 5×10^{-6}.

The performance of Anderson acceleration is summarized in Table 5.1 with details illustrated in Figure 5.1. We see that the inexact variant converges as rapidly as the exact variant, but the former needs less computational cost. For example, Table 5.1 shows that with $m = 2$, it takes both for the inexact variant and the exact variant 27 steps to converge, but the former only needs 4184 PSD iterations in total compared to 7154 for the latter.

In Figure 5.1, the left column and the right columns display the nonlinear residual norms $\|g(u^{(k)}) - u^{(k)}\|_{\ell_2}$ and the PSD iterations, respectively, versus the Anderson acceleration steps. We see that in the first few Anderson acceleration steps, the exact and the inexact variants deliver approximately the same nonlinear residual norms, which corroborates our results given in Theorem 4.1 that the inexact algorithm can follow the exact variant's behavior if $g(x_k)$ is evaluated approximately with appropriate accuracy. Later on, though $\|g(u^{(k)}) - u^{(k)}\|_{\ell_2}$ obtained at each step of the two variants is not very close, both methods overall exhibit a similar convergence behavior until the stopping criterion is satisfied. Moreover, it is evident that the PSD step counts are much lower for the inexact variant than for the exact one thanks to the advantage of the former in the early steps, as our Theorem 4.1 shows.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

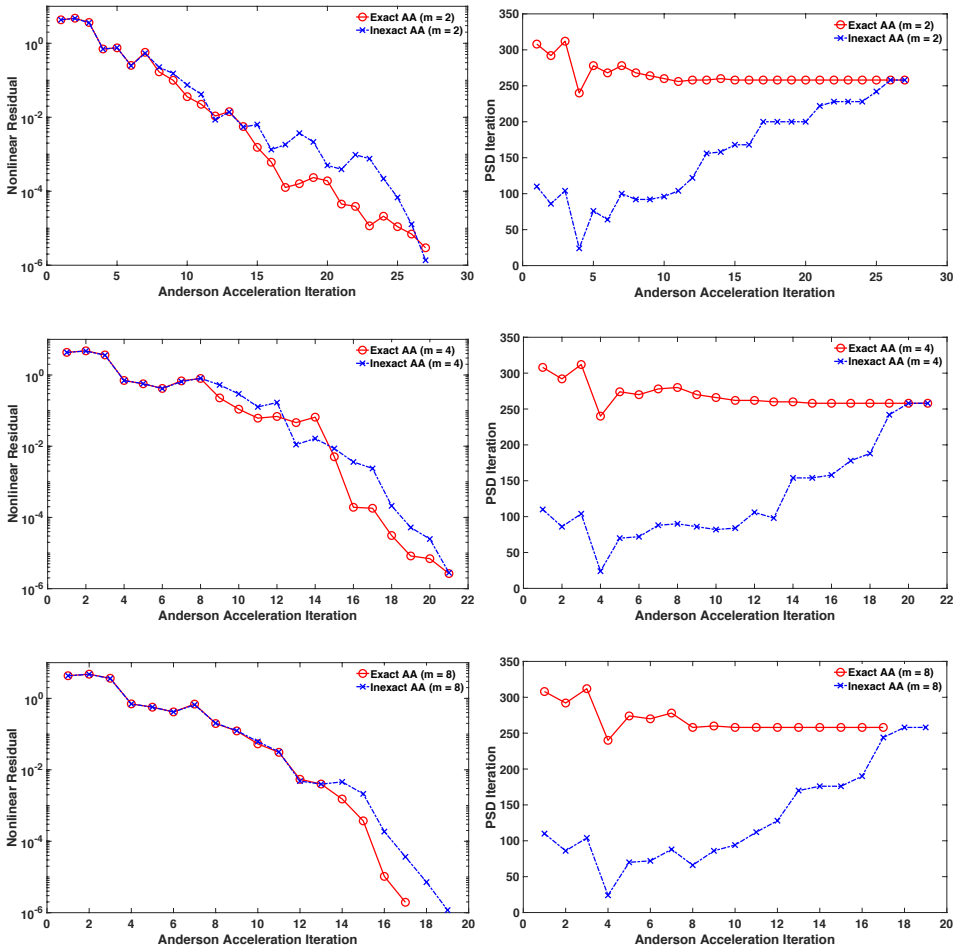

FIG. 5.1. *Performance of exact and inexact Anderson acceleration for solving a 1-D nonlinear Burgers' equation.*

Example 2. Consider the 1-D nonlinear Helmholtz (NLH) equation

$$\begin{aligned} u_{xx} + \kappa^2(1+\epsilon|u|^2)u = 0, &\qquad 0 < x < 10,\\ u_x + i\kappa u = 2i\kappa, &\qquad \text{at } x = 0,\\ u_x - i\kappa u = 0, &\qquad \text{at } x = 10, \end{aligned}$$

where κ is the wave number in the surrounding medium and $\epsilon \geq 0$ represents a material constant. This problem has been considered in [17] to test the *exact* Anderson acceleration. We discretize the domain into $n = 2^{15}$ equispaced subintervals and apply the second-order centered differences for the equation and the second-order forward/backward differences for the boundary conditions. The Picard iteration in the continuous form is defined as

$$u_{xx}^{(k+1)} + \kappa^2(1+\epsilon|u^{(k)}|^2)u^{(k+1)} = 0$$

with boundary conditions holding for each $u^{(k)}$. Since u on the boundary needs to be evaluated, we let $u^{(k)} = [u_0^{(k)}, u_1^{(k)}, \ldots, u_n^{(k)}]^T \in \mathbb{R}^{n+1}$, where $u_j^{(k)} \approx u(jh)$. The resulting scheme is

$$A_k u^{(k+1)} = b,$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

where

$$A_k = \begin{bmatrix} -3+2i\kappa h & 4 & -1 & & & \\ 1 & -2 & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & 1 & -2 & 1 & \\ & & 1 & -4 & 3-2i\kappa h \end{bmatrix} + \operatorname{diag}[0; \kappa^2 h^2 (1+\epsilon u^{(k)}(2\!:\!n)); 0],$$

$$b = [4i\kappa h, 0, \ldots, 0]^T,$$

which defines the relation $u^{(k+1)} = g(u^{(k)})$. The initial approximation $u^{(0)}$ contains the values of $e^{i\kappa x}$ (which satisfies the boundary conditions) at the nodes $x_j = jh$.

TABLE 5.2
Performance of exact and inexact Anderson acceleration for solving a 1-D nonlinear Helmholtz equation.

	AA progress			total GMRES iterations		
(ϵ, κ, m)	$\max \kappa_2(U_k^{(k)})$	exact	inexact	exact	inexact	improvement
$(0.17, 8, 5)$	2.90×10^2	60	58	3215	1790	44.3%
$(0.21, 10, 10)$	3.32×10^3	264	231	18314	7257	60.4%

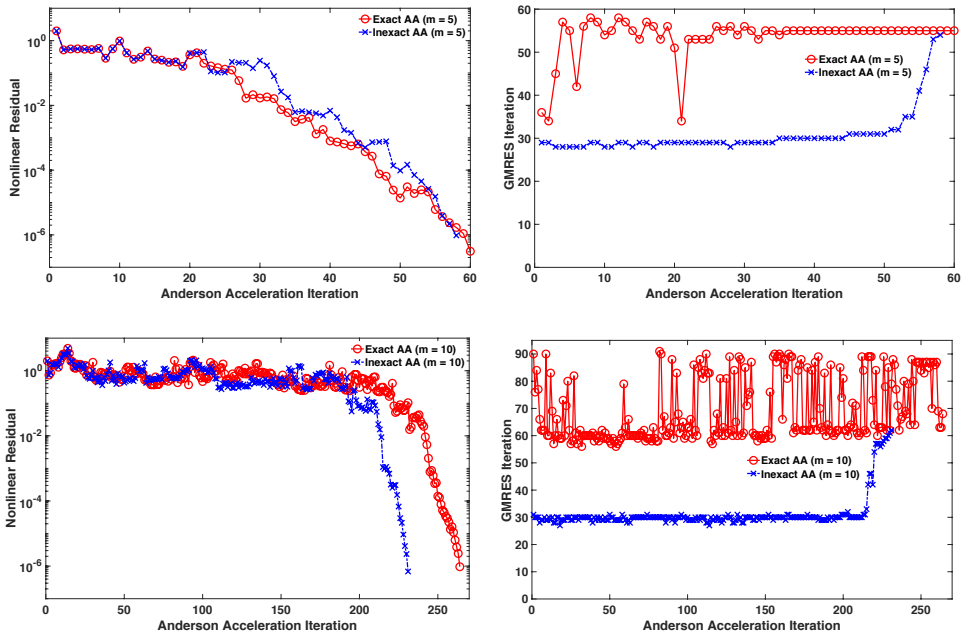

FIG. 5.2. *Performance of exact and inexact Anderson acceleration for solving a 1-D nonlinear Helmholtz equation.*

The parameters used here define two non-contractive mappings. With $\epsilon = 0.17$ and $\kappa = 8$, the fixed point iteration $u^{(k+1)} = g(u^{(k)})$ with the initial $u^{(0)}$ set as the approximate fixed point u^* (computed by Anderson acceleration with $m = 5$) slowly diverges from u^* with the

ETNA
Kent State University and
Johann Radon Institute (RICAM)

nonlinear residual norm $\|u^{(k)} - g(u^{(k)})\|$ gradually increasing to $\mathcal{O}(10^{-1})$. With $\epsilon = 0.21$ and $\kappa = 10$, such a divergence occurs much more rapidly with the nonlinear residual quickly going up to $\mathcal{O}(1)$ without blowup in $u^{(k)}$.

Similar to Example 1, a direct solution to these linear systems is easy, yet we solve them iteratively to illustrate our point. Since A_k is nonsymmetric, we use the right-side preconditioned GMRES(30) as the linear solver with the preconditioner M being the incomplete LU factorization with threshold and pivoting of the *tridiagonal* part of A_k with drop tolerance 3×10^{-3}. The relative tolerance for the GMRES solves is 10^{-9} for the exact method and $\max\left\{10^{-9}, \min\{10^{-4}, 10^{-3}\|g(u^{(k-1)}) - u^{(k-1)}\|_{\ell_2}\}\right\}$ for the inexact one; see the definition of $\|u\|_{\ell_2}$ in Example 1. Both methods stop when $\|g(u^{(k)}) - u^{(k)}\|_{\ell_2} \leq 10^{-6}$.

The performance of Algorithms 1 and 2 is illustrated in Figure 5.2. We note that the number of preconditioned GMRES iterations for the exact algorithm fluctuates widely for different outer iterations. This is probably due to the irregular convergence behavior of restarted GMRES; see, e.g., [3, 8, 16, 23]. Fortunately, we did not observe such irregularity with the inexact algorithm. Overall, the results are largely similar to those obtained for Example 1. Inexact Anderson acceleration takes slightly fewer steps to converge and considerably fewer GMRES iterations than the exact variant, especially in the early steps. For more difficult problems, more Anderson acceleration steps are needed, and the inexact method again seems to have a stronger advantage over the exact method in terms of the number of total GMRES iterations.

Example 3. Consider the 3-D steady-state Navier-Stokes equation (NSE)

$$\begin{aligned} -\nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p &= \mathbf{f}, && \text{in } \Omega \subset \mathbb{R}^3, \\ \nabla\cdot\mathbf{u} &= 0, && \text{in } \Omega, \\ \mathbf{u}|_{\partial\Omega} &= \mathbf{g}, \end{aligned}$$

which models incompressible flows in a lid-driven cavity. Here, ν is the kinematic viscosity, which is inversely proportional to the Reynolds number Re, $\mathbf{f}$ is a forcing term, and $\mathbf{u}$ and p represent the velocity and pressure, respectively. This problem has also been considered in [17] to test *exact* Anderson acceleration. We choose the domain $\Omega = [0,1] \times [0,1] \times [0,1]$ with no-slip boundary conditions on the four sides and the bottom and a moving-lid on the top imposed by the Dirichlet boundary condition $\mathbf{u}(x, y, 1) = [1, 0, 0]^T$ for the velocity. There is no external force applied. The Reynolds numbers considered are $Re = 500$ and 1000. The first Hopf bifurcation appears to occur when $Re \approx 2000$. The equation is discretized using (P_3, P_2^{disc})-Scott-Vogelius finite elements on a barycenter refined tetrahedral mesh that provides 477 thousand and 1.4 million total degrees of freedom for the two Reynolds numbers, respectively. The Picard iteration constructs a sequence of approximate solutions by solving the linear Oseen problem [7]

$$\begin{aligned} -\nu\Delta\mathbf{u}^{(k+1)} + (\mathbf{u}^{(k)}\cdot\nabla)\mathbf{u}^{(k+1)} + \nabla p^{(k+1)} &= \mathbf{f}, \\ \nabla\cdot\mathbf{u}^{(k+1)} &= 0, \\ \mathbf{u}^{(k+1)}|_{\partial\Omega} &= \mathbf{g}. \end{aligned}$$

This relation defines $u^{(k+1)} = g(u^{(k)})$, where $u^{(k)}$ is the vectorized velocity $\mathbf{u}^{(k)}$.

As shown in [18], with the initial approximation $u^{(0)} = 0$, the fixed-point iteration $u^{(k+1)} = g(u^{(k)})$ does not converge when $Re \geq 400$. In addition, we let $u^{(0)}$ be the approximate fixed-point solution u^* (found by Anderson acceleration with $m = 10$), run the fixed-point iteration, and find that the nonlinear residual increases slowly with $Re = 500$ to

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$\mathcal{O}(10^{-2})$ and more quickly with $Re = 1000$ to $\mathcal{O}(10^{-1})$. These observations show that g is non-contractive near the fixed-point solution with these relatively high Reynolds numbers.

The Oseen problem requires a solution to a linear system of the form

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix},$$

which is equivalent, by the classical augmented Lagrangian-based approach [4], to

$$\begin{bmatrix} A + \gamma B^T W^{-1} B & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix}. \tag{5.2}$$

Here, we choose $\gamma = 0.1$, and we let the action of the preconditioner M^{-1} be the action of the inverse of the coefficient matrix in (5.2) with the Schur complement replaced with the pressure mass matrix. The right-hand side preconditioned GMRES(50) method is used to solve such a linear system at each Picard iteration step.

TABLE 5.3
Performance of exact and inexact Anderson acceleration for solving a 3-D steady-state Navier-Stokes equation.

	AA progress			total GMRES iterations		
parameters	$\max \kappa_2(U_k^{(k)})$	exact	inexact	exact	inexact	improvement
$Re = 500$, $m = 10$	7.14×10^2	27	27	1778	963	45.8%
$Re = 1000$, $m = 15$	2.63×10^3	46	46	2533	1430	43.5%

The initial approximation is $\mathbf{u}^{(0)} = \mathbf{0}$. The exact Anderson acceleration uses a relative tolerance of 10^{-8} for all GMRES solves, whereas the inexact variant sets the GMRES tolerance to $\max\left\{10^{-8}, \min\{10^{-3}, 10^{-3}\left\|g(u^{(k-1)}) - u^{(k-1)}\right\|_{\ell_2}\}\right\}$, where $\|u\|_{\ell_2}$ is the approximation of the L_2-norm of the velocity. Anderson acceleration stops when $\left\|g(u^{(k)}) - u^{(k)}\right\|_{\ell_2} \leq 10^{-5}$.

Table 5.3 shows that for both experiments, the exact and inexact Anderson acceleration take the same number of steps to converge, but the inexact variant needs fewer GMRES iterations. We also see from Figure 5.3 that the nonlinear residuals $\left\|g(u^{(k)}) - u^{(k)}\right\|_{\ell_2}$ of the two methods are fairly close throughout the computation, though the inexact variant takes a small number of GMRES iterations in the early stage. The pattern is consistent with the previous two numerical experiments. In fact, the nearly overlapping nonlinear residual curves suggest that there might be additional room to further relax the accuracy of GMRES for inexact Anderson acceleration to keep the convergence comparable with the exact method.

Finally, we note that the highest condition number of $U_k^{(k)}$ in Example 1 is higher than those in Examples 2 and 3, and we consequently use a smaller $\tau = 10^{-5}$ for Example 1 and a larger $\tau = 10^{-3}$ for Examples 2 and 3. With these parameters, inexact Anderson acceleration overall tracks the exact method fairly well. Using a larger τ for Example 1 would lead to a slower convergence of the inexact Anderson acceleration, whereas using a smaller τ for Examples 2 and 3 would result in more inner iterations without speeding up the convergence of the inexact method. This pattern seems consistent with our remark after Theorem 4.1.

6. Conclusion. In this paper, we developed a one-step convergence analysis of inexact Anderson acceleration where the optimization is performed in the vector 2-norm by the linear least-squares method for computing a fixed point solution of $x = g(x)$, where g is non-contractive. Existing results for inexact Anderson acceleration for contractive mappings [26] are not applicable in this case. Our main result (Theorem 3.1) shows that if each $g(x_k)$ is

ETNA
Kent State University and
Johann Radon Institute (RICAM)

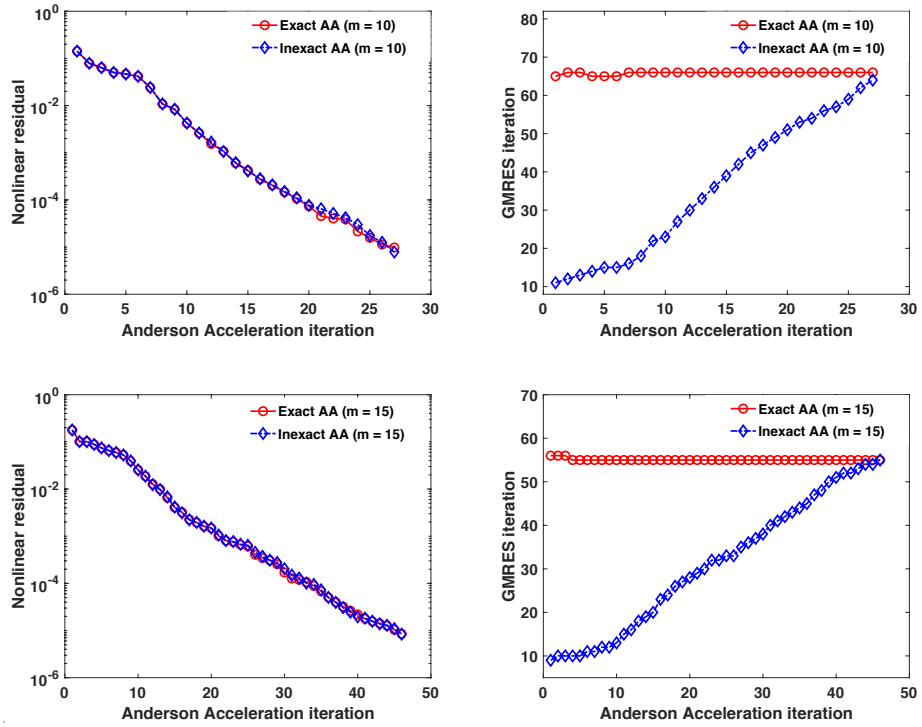

FIG. 5.3. *Performance of exact and inexact Anderson acceleration for solving a 3-D steady-state Navier-Stokes equation.*

evaluated approximately with an error proportional to the nonlinear residual norm $\|w_k\| = \|g(x_k) - x_k\|$, then the inexact algorithm may still converge as rapidly as the exact method if the optimization at each step has a sufficient gain and the linear least-squares problem is not ill-conditioned. Our insight is obtained from properties of orthogonal and oblique projectors and their perturbations arising from approximate evaluations of $g(x_k)$.

Our numerical examples cover a few well-known nonlinear partial differential equations with carefully chosen parameters for which the fixed-point iterations under consideration are non-contractive. We show consistently that inexact Anderson acceleration would save computational cost in the early stage of this algorithm when the residual norm is large. With a reasonably small tolerance for the stopping criterion of Anderson acceleration, numerical tests suggest that the inexact algorithm typically could save about 40%–50% of the total cost for evaluating $g(x_k)$ while maintaining the convergence of the exact method.

Acknowledgement. I am grateful to my colleague, Leo Rebholz, for the problem motivation and the Anderson acceleration code for solving the nonlinear Helmholtz and the Navier-Stokes equations. I also appreciate the two anonymous reviewers whose careful reading and suggestions helped me improve the manuscript.

REFERENCES

[1] D. G. M. ANDERSON, *Iterative procedures for nonlinear integral equations*, J. Assoc. Comput. Mach., 12 (1965), pp. 547–560.

[2] ———, *Comments on "Anderson acceleration, mixing and extrapolation"*, Numer. Algorithms, 80 (2019), pp. 135–234.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[3] A. H. BAKER, E. R. JESSUP, AND T. MANTEUFFEL, *A technique for accelerating the convergence of restarted GMRES*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 962–984.
[4] M. BENZI AND M. A. OLSHANSKII, *An augmented Lagrangian-based approach to the Oseen problem*, SIAM J. Sci. Comput., 28 (2006), pp. 2095–2113.
[5] E. CĂTINAŞ, *The inexact, inexact perturbed, and quasi-Newton methods are equivalent models*, Math. Comp., 74 (2005), pp. 291–301.
[6] R. S. DEMBO, S. C. EISENSTAT, AND T. STEIHAUG, *Inexact Newton methods*, SIAM J. Numer. Anal., 19 (1982), pp. 400–408.
[7] H. C. ELMAN, D. J. SILVESTER, AND A. J. WATHEN, *Finite Elements and Fast Iterative Solvers*, 2nd ed., Oxford University Press, Oxford, 2014.
[8] M. EMBREE, *The tortoise and the hare restart GMRES*, SIAM Rev., 45 (2003), pp. 259–266.
[9] C. EVANS, S. POLLOCK, L. REBHOLZ, AND M. XIAO, *A proof that Anderson acceleration increases the convergence rate in linearly converging fixed point methods (but not in quadratically converging ones)*, SIAM J. Numer. Anal., 58 (2020), pp. 788–810.
[10] V. EYERT, *A comparative study on methods for convergence acceleration of iterative vector sequences*, J. Comput. Phys., 124 (1996), pp. 271–285.
[11] H.-R. FANG AND Y. SAAD, *Two classes of multisecant methods for nonlinear acceleration*, Numer. Linear Algebra Appl., 16 (2009), pp. 197–221.
[12] D. R. FOKKEMA, G. L. G. SLEIJPEN, AND H. A. VAN DER VORST, *Jacobi-Davidson style QR and QZ algorithms for the reduction of matrix pencils*, SIAM J. Sci. Comput., 20 (1998), pp. 94–125.
[13] C. T. KELLEY, *Numerical methods for nonlinear equations*, Acta Numer., 27 (2018), pp. 207–287.
[14] J. LIESEN AND P. TICHÝ, *Convergence analysis of Krylov subspace methods*, GAMM Mitt. Ges. Angew. Math. Mech., 27 (2004), pp. 153–173 (2005).
[15] K. LIPNIKOV, D. SVYATSKIY, AND Y. VASSILEVSKI, *Anderson acceleration for nonlinear finite volume scheme for advection-diffusion problems*, SIAM J. Sci. Comput., 35 (2013), pp. A1120–A1136.
[16] R. B. MORGAN, *GMRES with deflated restarting*, SIAM J. Sci. Comput., 24 (2002), pp. 20–37.
[17] S. POLLOCK AND L. G. REBHOLZ, *Anderson acceleration for contractive and noncontractive operators*, IMA J. Numer. Anal., 41 (2021), pp. 2841–2872.
[18] S. POLLOCK, L. G. REBHOLZ, AND M. XIAO, *Anderson-accelerated convergence of Picard iterations for incompressible Navier-Stokes equations*, SIAM J. Numer. Anal., 57 (2019), pp. 615–637.
[19] P. PULAY, *Convergence acceleration of iterative sequences. The case of SCF iteration*, Chem. Phys. Lett., 73 (1980), pp. 393–398.
[20] ———, *Improved SCF convergence*, J. Comput. Chem., 3 (1982), pp. 556–560.
[21] T. ROHWEDDER AND R. SCHNEIDER, *An analysis for the DIIS acceleration method used in quantum chemistry calculations*, J. Math. Chem., 49 (2011), pp. 1889–1914.
[22] V. SIMONCINI AND D. B. SZYLD, *On the occurrence of superlinear convergence of exact and inexact Krylov subspace methods*, SIAM Rev., 47 (2005), pp. 247–272.
[23] ———, *Recent computational developments in Krylov subspace methods for linear systems*, Numer. Linear Algebra Appl., 14 (2007), pp. 1–59.
[24] G. L. G. SLEIJPEN AND H. A. VAN DER VORST, *A Jacobi-Davidson iteration method for linear eigenvalue problems*, SIAM Rev., 42 (2000), pp. 267–293.
[25] D. B. SZYLD AND F. XUE, *Local convergence analysis of several inexact Newton-type algorithms for general nonlinear eigenvalue problems*, Numer. Math., 123 (2013), pp. 333–362.
[26] A. TOTH, J. A. ELLIS, T. EVANS, S. HAMILTON, C. T. KELLEY, R. PAWLOWSKI, AND S. SLATTERY, *Local improvement results for Anderson acceleration with inaccurate function evaluations*, SIAM J. Sci. Comput., 39 (2017), pp. S47–S65.
[27] A. TOTH AND C. T. KELLEY, *Convergence analysis for Anderson acceleration*, SIAM J. Numer. Anal., 53 (2015), pp. 805–819.
[28] H. F. WALKER AND P. NI, *Anderson acceleration for fixed-point iterations*, SIAM J. Numer. Anal., 49 (2011), pp. 1715–1735.
[29] T. WASHIO AND C. W. OOSTERLEE, *Krylov subspace acceleration for nonlinear multigrid schemes*, Electron. Trans. Numer. Anal., 6 (1997), pp. 271–290.
http://etna.ricam.oeaw.ac.at/vol.6.1997/pp271-290.dir/pp271-290.pdf