The Electronic Journal of Linear Algebra.
A publication of the International Linear Algebra Society.
Volume 6, pp. 72-94, August 2000.
ISSN 1081-3810. http://math.technion.ac.il/iic/ela

ELA

A BI-CG TYPE ITERATIVE METHOD FOR DRAZIN-INVERSE SOLUTION OF SINGULAR INCONSISTENT NONSYMMETRIC LINEAR SYSTEMS OF ARBITRARY INDEX*

AVRAM SIDI† AND VLADIMIR KLUZNER‡

Abstract. Consider the linear system $Ax = b$, where b is a vector in $\mathbb{C}^N$, $A \in \mathbb{C}^{N\times N}$ is a singular matrix, and $\mathrm{ind}\,(A) = a$ is arbitrary. Here $\mathrm{ind}\,(\cdot)$ denotes the index of a matrix. The Drazin-inverse solution of this system is defined to be the vector $A^{\mathrm{D}}b$, where the matrix A^{D} is the Drazin inverse of A. The Drazin-inverse solution of singular linear systems has been considered recently by the first author within the context of extrapolation methods, when $\mathrm{ind}\,(A)$ is arbitrary. It has also been considered within the context of Krylov subspace methods, when A is real symmetric (hence $\mathrm{ind}\,(A) = 1$ necessarily). In addition, semi-iterative methods have been developed for the cases in which $\mathrm{ind}\,(A) = 1$ and $\mathrm{ind}\,(A) > 1$, assuming that the spectrum of A is real nonnegative. The purpose of the present work is to develop a Bi-CG type Krylov subspace method suitable for the general case in which A is not necessarily real symmetric, its index is arbitrary, and its spectrum is not necessarily real. The method that is developed can be implemented via a 4-term recursion relation independently of the size of $\mathrm{ind}\,(A)$ and produces $A^{\mathrm{D}}b$ in at most $N - a$ steps. A detailed error analysis for this method is provided and the results are illustrated with suitable numerical examples.

Key words. Singular linear systems, Drazin-inverse solution, Krylov subspace methods, Lanczos method, Bi-Conjugate Gradient algorithm.

AMS subject classifications. 15A06, 15A09, 65F10, 65F50

1. Introduction.

Consider the linear system

$$Ax = b \ , \tag{1.1}$$

where $A \in \mathbb{C}^{N\times N}$ is singular and $\mathrm{ind}\,(A) = a$ is arbitrary. Here $\mathrm{ind}\,(\cdot)$ denotes the index of a matrix, namely, the size of the largest Jordan block corresponding to its zero eigenvalue. The purpose of this paper is to develop a Krylov subspace method of the Bi-Conjugate Gradient (Bi-CG) type for computing the Drazin-inverse solution of (1.1), namely, the vector $A^{\mathrm{D}}b$, where A^{D} denotes the Drazin inverse of A. For the Drazin inverse and its properties, see, e.g., Ben-Israel and Greville [3] or Campbell and Meyer [5].

We recall the following definition of the Drazin inverse of A: Let

$$A = P \left[\begin{array}{c|c} J_0 & O \\ \hline O & J_1 \end{array}\right] P^{-1} \ ,$$

*Received by the editors on 8 August 1999. Accepted for publication on 12 July 2000. Handling Editor: Daniel Hershkowitz.

†Computer Science Department, Technion - Israel Institute of Technology, Haifa, 32000, ISRAEL (asidi@cs.technion.ac.il)

‡Mathematics Department, Technion - Israel Institute of Technology, Haifa, 32000, ISRAEL (mavladim@math.technion.ac.il). The work of this author was performed in partial fulfillment of the requirements for the M.Sc. degree in Applied Mathematics, Technion - Israel Institute of Technology.

where P is nonsingular, J_0 contains all the Jordan blocks of A corresponding to the zero eigenvalues, and J_1 containing all the remaining Jordan blocks. Then,

$$A^{\mathrm{D}} = P \left[\begin{array}{c|c} O & O \\ \hline O & J_1^{-1} \end{array} \right] P^{-1} \ .$$

We do not put any restrictions on the matrix A. Thus, A is not necessarily Hermitian or Hermitian positive semidefinite. It can have any type of spectrum, in addition to having an arbitrary index, as assumed above. Neither do we put any restrictions on the linear system (1.1). This system may be consistent or inconsistent. (Recall that in the case $a > 1$, $A^{\mathrm{D}}b$ is not necessarily a solution of $Ax = b$ even when the latter is consistent.) We are only required to know a, the index of A.

The subject of Krylov subspace methods for computing $A^{\mathrm{D}}b$ has been treated in a few papers. First, the method of Conjugate Gradients (CG) may be applied when A is Hermitian positive semidefinite and (1.1) is consistent, see Kaaschieter [11]. It is shown in Sidi [15] that the method of Arnoldi [1], the method of Generalized Conjugate Residuals (GCR) and the method of Lanczos [12] as well, can be applied to non-Hermitian but consistent systems when the index a is unity, and error bounds are given. In addition, [15] provides a complete convergence theory for these methods and others in the presence of initial iterations via the Richardson iterative method.

The treatment of the singular inconsistent systems by Krylov subspace methods has proved to be much harder, however. This is the case even for the simplest cases in which $a = 1$. To date we are aware of the CG type methods of Calvetti, Reichel and Zhang [4] that apply to Hermitian systems only. A recent work by Fischer, Hanke, and Hochbruck [8] provides a class of methods similar to that of [4] and that applies to the same problems.

A unified framework for the development of new Krylov subspace methods for singular inconsistent systems has recently been proposed in Sidi [17]. It turns out that the method we develop in the present work falls in this framework. We would like to note that the recursive algorithm that we develop here and the accompanying error analysis are completely new, however.

Finally, we mention the vector extrapolation methods developed in Sidi [16] for treating the most general case of singular non-Hermitian inconsistent systems with arbitrary index. This paper also contains a detailed convergence analysis for the methods developed in it.

In the next section we give some technical preliminaries in which we describe some of the basic requirements from the approximations x_n to $A^{\mathrm{D}}b$. In Sections 3 and 4 we develop a Lanczos type method and derive a Bi-Conjugate Gradient (Bi-CG) type algorithm for it that is motivated by the Bi-CG algorithm of Fletcher [9] that implements the method of Lanczos [12] for nonsingular systems. This algorithm, which we denote the DBi-CG algorithm, involves 4-term recursion relations independently of the size of $\mathrm{ind}\,(A)$. It is thus very economical both computationally and storage-wise. In Section 5 we give a detailed error analysis for the method developed. Finally, in Section 6 we give some numerical experiments with the new algorithm.

In view of the above, the present work seems to be the first to present a Krylov subspace method for singular non-Hermitian inconsistent linear systems with arbitrary index together with a recursive algorithm of finite length and an error analysis.

2. Theoretical preliminaries. Let x_0 be an arbitrary initial vector and $r_0 = b - Ax_0$ be the corresponding residual vector. Then, beginning with x_0, we generate a sequence of vectors $x_n,\ n = 1, 2, \ldots,$ that will hopefully approximate $A^{\mathrm{D}}b$ via the iteration

$$x_n = x_0 + q_{n-1}(A)r_0 = p_n(A)x_0 + q_{n-1}(A)b \ , \tag{2.1}$$

where $q_{n-1}(\lambda)$ is a polynomial of degree at most $n-1$ and $p_n(\lambda)$ is a polynomial of degree at most n given by $p_n(\lambda) = 1 - \lambda q_{n-1}(\lambda)$. We call $p_n(\lambda)$ the nth residual polynomial since $r_n = b - Ax_n = p_n(A)r_0$. Note that

$$p_n(0) = 1 \ . \tag{2.2}$$

As is shown in Eiermann, Marek, and Niethammer [7], necessary and sufficient conditions for the convergence of the sequence $\{x_n\}$ are that

$$\lim_{n\to\infty} p_n^{(i)}(0) = 0 \ , \qquad i = 1, \ldots, a \ , \tag{2.3}$$

and

$$\lim_{n\to\infty} p_n^{(i)}(\lambda_j) = 0 \ , \qquad i = 0, \ldots, k_j - 1 \ ,$$

where λ_j are the nonzero eigenvalues of A and $k_j = \mathrm{ind}\,(A - \lambda_j I)$.

The conditions in (2.3) will, of course, be satisfied if

$$p_n^{(i)}(0) = 0 \ , \qquad i = 1, \ldots, a \ , \text{ for all } n = 0, 1, \ldots \ , \tag{2.4}$$

Our purpose is to design a recursive method to construct the vectors $x_n,\ n = 1, 2, \ldots,$ beginning with an arbitrary vector x_0 that will be of the form described above, but with the corresponding polynomials $p_n(\lambda)$ satisfying (2.4) instead of (2.3), in addition to (2.2).

Before going on it will be convenient to introduce some notation that has been used before. We shall denote by Π_n the set of all polynomials of degree at most n. We shall also define

$$\Pi_n^0 = \{p \in \Pi_n \ : \ p(0) = 1, \ p^{(i)}(0) = 0, \ i = 1, \ldots, a\} \ . \tag{2.5}$$

That is to say, Π_n^0 is the collection of all polynomials of degree at most n that satisfy (2.2) and (2.4). Thus, the polynomials $p_n(\lambda)$ that we will be considering in the present work are all in Π_n^0. Note that $p_n(\lambda) = 1$ is the only member of Π_n^0 for $n = 0, 1, \ldots, a$.

Finally, we will work with the standard Euclidean inner product $(x, y) \equiv x^*y$, for which $(\alpha x, \beta y) = \overline{\alpha}\beta(x, y)$ for any $\alpha, \beta \in \mathbb{C}$ and any $x, y \in \mathbb{C}^N$. Also, $(x, y) = \overline{(y, x)}$. We shall also write $x \perp y$ to mean $(x, y) = 0$. We shall denote by $\|\cdot\|$ both the vector l_2-norm and the matrix norm induced by it. That is, $\|x\| = \sqrt{(x, x)}$ and $\|A\| = \sigma_{max}$, the largest singular value of A.

3. A Bi-CG type method and a preliminary algorithm. Let us pick x_0 arbitrarily. By the fact that $p_n(\lambda) = 1$ for $n = 1, \ldots, a$ (and hence $q_{n-1}(\lambda) = 0$), we have

$$x_a = \ldots = x_1 = x_0 \quad \text{and} \quad r_a = \ldots = r_1 = r_0 \ . \tag{3.1}$$

By (2.1) and by the fact that $p_{a+1} \in \Pi^0_{a+1}$, x_{a+1} must be of the form

$$x_{a+1} = x_0 + \rho A^a r_0 \ , \tag{3.2}$$

so that

$$r_{a+1} = b - Ax_{a+1} = r_0 - \rho A^{a+1} r_0 \ . \tag{3.3}$$

Now we need to determine ρ by imposing some suitable criterion. Let us pick a vector $\check{x}_0$ simultaneously with x_0 and define $\check{r}_0 = b - A^* \check{x}_0$. Similarly , let us also set

$$\check{x}_a = \ldots = \check{x}_1 = \check{x}_0 \quad \text{and} \quad \check{r}_a = \ldots = \check{r}_1 = \check{r}_0 \ . \tag{3.4}$$

Then our criterion will be that

$$r_{a+1} \perp (A^*)^{a+1} \check{r}_0 \ , \tag{3.5}$$

from which we can determine ρ as

$$\rho = \frac{((A^*)^{a+1}\check{r}_0 \ , \ r_0)}{((A^*)^{a+1}\check{r}_0 \ , \ A^{a+1} r_0)} \ , \tag{3.6}$$

assuming that the denominator does not vanish. As in Bi-CG, we now develop an algorithm that computes simultaneously two sets of vectors $\{x_n\}$ and $\{\check{x}_n\}$. Analogously to x_{a+1} and r_{a+1}, let us define $\check{x}_{a+1}$ and $\check{r}_{a+1}$ by

$$\check{x}_{a+1} = \check{x}_0 + \check{\rho}(A^*)^a \check{r}_0 \ , \qquad \check{r}_{a+1} = b - A^* \check{x}_{a+1} = \check{r}_0 - \check{\rho}(A^*)^{a+1} \check{r}_0 \ , \tag{3.7}$$

subject to the criterion

$$\check{r}_{a+1} \perp A^{a+1} r_0 \ . \tag{3.8}$$

As a result,

$$\check{\rho} = \frac{(A^{a+1} r_0 \ , \ \check{r}_0)}{(A^{a+1} r_0 \ , \ (A^*)^{a+1} \check{r}_0)} \ . \tag{3.9}$$

We now aim at obtaining the x_n and $\check{x}_n$ by 4-term recursion relations of the form

$$\begin{aligned} x_{n+1} &= x_n + \omega_n A(x_n - x_{n-1}) + \mu_n (x_n - x_{n-1}) + \nu_n (x_{n-1} - x_{n-2}) \ , \\ \check{x}_{n+1} &= \check{x}_n + \check{\omega}_n A(\check{x}_n - \check{x}_{n-1}) + \check{\mu}_n (\check{x}_n - \check{x}_{n-1}) + \check{\nu}_n (\check{x}_{n-1} - \check{x}_{n-2}) \ , \end{aligned} \tag{3.10}$$

for all values of the index a. We already know that the x_n can be obtained as in (3.10) in the semi-iterative method of Climent, Neumann, and Sidi [6] for all a, and hope that this will be the case here too. Recalling that $r_{n+1} = b - Ax_{n+1}$, $\check{r}_{n+1} = b - A^*\check{x}_{n+1}$, from (3.10) we obtain

$$\begin{aligned} r_{n+1} &= r_n + \omega_n A(r_n - r_{n-1}) + \mu_n(r_n - r_{n-1}) + \nu_n(r_{n-1} - r_{n-2}) \ , \\ \check{r}_{n+1} &= \check{r}_n + \check{\omega}_n A^*(\check{r}_n - \check{r}_{n-1}) + \check{\mu}_n(\check{r}_n - \check{r}_{n-1}) + \check{\nu}_n(\check{r}_{n-1} - \check{r}_{n-2}) \ . \end{aligned} \tag{3.11}$$

Defining

$$z_n = r_n - r_{n-1} \quad \text{and} \quad \check{z}_n = \check{r}_n - \check{r}_{n-1} \tag{3.12}$$

for all n, we realize that (3.11) will become 3-term recursion relations for the z_n and $\check{z}_n$, namely,

$$\begin{aligned} z_{n+1} &= \omega_n A z_n + \mu_n z_n + \nu_n z_{n-1} \ , \\ \check{z}_{n+1} &= \check{\omega}_n A^* \check{z}_n + \check{\mu}_n \check{z}_n + \check{\nu}_n \check{z}_{n-1} \ . \end{aligned} \tag{3.13}$$

We also have two initial conditions, namely,

$$\begin{gathered} z_a = 0 \ , \qquad \check{z}_a = 0 \ , \qquad \text{and} \\ z_{a+1} = r_{a+1} - r_a = -\rho A^{a+1} r_0 \ , \qquad \check{z}_{a+1} = \check{r}_{a+1} - \check{r}_a = -\check{\rho}(A^*)^{a+1}\check{r}_0 \ . \end{gathered} \tag{3.14}$$

At this point we recall that sets of vectors $\{z_{a+1}, z_{a+2}, \ldots\}$ and $\{\check{z}_{a+1}, \check{z}_{a+2}, \ldots\}$ that satisfy recursion relations of the form given in (3.13) can be obtained by applying the biorthogonalization process of Lanczos [12] to the Krylov subspaces $\{z_{a+1}, Az_{a+1}, A^2 z_{a+1}, \ldots\}$ and $\{\check{z}_{a+1}, A^*\check{z}_{a+1}, (A^*)^2\check{z}_{a+1}, \ldots\}$, where by biorthogonality of the vectors z_i and $\check{z}_j$ we mean that

$$(z_i \ , \ \check{z}_j) = 0 \quad \text{if} \quad i \neq j \ . \tag{3.15}$$

We also recall that this process is very economical in the sense that it requires a fixed amount of memory, and the computation of each z_n and $\check{z}_n$ requires a fixed number of arithmetic operations independent of n.

Let us now determine the z_n and $\check{z}_n$. We start with z_{a+2}. By the fact that $z_a = 0$ and $\check{z}_a = 0$ we have

$$\begin{aligned} z_{a+2} &= \omega_{a+1} A z_{a+1} + \mu_{a+1} z_{a+1} \ , \\ \check{z}_{a+2} &= \check{\omega}_{a+1} A^* \check{z}_{a+1} + \check{\mu}_{a+1} \check{z}_{a+1} \ . \end{aligned} \tag{3.16}$$

Forming the inner products with the vectors $\check{z}_{a+1}$ and z_{a+1}, and invoking (3.15), we obtain one equation for ω_{a+1} and μ_{a+1} and another for $\check{\omega}_{a+1}$ and $\check{\mu}_{a+1}$:

$$\begin{aligned} (\check{z}_{a+1} \ , \ Az_{a+1})\omega_{a+1} + (\check{z}_{a+1} \ , \ z_{a+1})\mu_{a+1} = 0 \ , \\ (z_{a+1} \ , \ A^*\check{z}_{a+1})\check{\omega}_{a+1} + (z_{a+1} \ , \ \check{z}_{a+1})\check{\mu}_{a+1} = 0 \ , \end{aligned} \tag{3.17}$$

Obviously, we need to complement each of these equations with an additional condition so that we will be able to determine ω_{a+1} and μ_{a+1} on the one hand and $\check{\omega}_{a+1}$ and $\check{\mu}_{a+1}$ on the other. The additional conditions that we choose to impose are

$$\check{r}_{a+2} \perp Az_{a+1} \quad \text{and} \quad r_{a+2} \perp A^*\check{z}_{a+1} \ . \tag{3.18}$$

The way we invoke (3.18) is as follows: using (3.12) on the left-hand side of (3.16), we first have:

$$
\begin{aligned}
r_{a+2} &= r_{a+1} + \omega_{a+1} A z_{a+1} + \mu_{a+1} z_{a+1} \ , \\
\check{r}_{a+2} &= \check{r}_{a+1} + \check{\omega}_{a+1} A^* \check{z}_{a+1} + \check{\mu}_{a+1} \check{z}_{a+1} \ .
\end{aligned} \tag{3.19}
$$

We next take the inner product of the first equality in (3.19) with $A^*\check{z}_{a+1}$ and of the second with Az_{a+1} and invoke (3.18). The additional conditions then are

$$
\begin{aligned}
&(A^*\check{z}_{a+1} \ , \ r_{a+1}) + (A^*\check{z}_{a+1} \ , \ Az_{a+1})\omega_{a+1} + (A^*\check{z}_{a+1} \ , \ z_{a+1})\mu_{a+1} = 0 \ , \\
&(Az_{a+1} \ , \ \check{r}_{a+1}) + (Az_{a+1} \ , \ A^*\check{z}_{a+1})\check{\omega}_{a+1} + (Az_{a+1} \ , \ \check{z}_{a+1})\check{\mu}_{a+1} = 0 \ .
\end{aligned} \tag{3.20}
$$

For $n = a+2,\ a+3, \ldots$, we can determine $\omega_n,\ \mu_n,\ \nu_n,\ \check{\omega}_n,\ \check{\mu}_n,\ \check{\nu}_n$ in (3.13) by imposing (3.15), i.e.

$$
(\check{z}_i \ , \ z_{n+1}) = 0 \quad \text{and} \quad (z_i \ , \ \check{z}_{n+1}) = 0 \ , \quad i = n-1,\ n \ ,
$$

and the additional conditions

$$
\check{r}_{n+1} \perp Az_n \quad \text{and} \quad r_{n+1} \perp A^*\check{z}_n \ . \tag{3.21}
$$

The resulting equations are

$$
\begin{aligned}
&(\check{z}_{n-1} \ , \ Az_n)\omega_n + (\check{z}_{n-1} \ , \ z_{n-1})\nu_n = 0 \ , \\
&(\check{z}_n \ , \ Az_n)\omega_n + (\check{z}_n \ , \ z_n)\mu_n = 0 \ , \\
&(A^*\check{z}_n \ , \ r_n) + (A^*\check{z}_n \ , \ Az_n)\omega_n + (A^*\check{z}_n \ , \ z_n)\mu_n + (A^*\check{z}_n \ , \ z_{n-1})\nu_n = 0 \ ,
\end{aligned} \tag{3.22}
$$

$$
\begin{aligned}
&(z_{n-1} \ , \ A^*\check{z}_n)\check{\omega}_n + (z_{n-1} \ , \ \check{z}_{n-1})\check{\nu}_n = 0 \ , \\
&(z_n \ , \ A^*\check{z}_n)\check{\omega}_n + (z_n \ , \ \check{z}_n)\check{\mu}_n = 0 \ , \\
&(Az_n \ , \ \check{r}_n) + (Az_n \ , \ A^*\check{z}_n)\check{\omega}_n + (Az_n \ , \ \check{z}_n)\check{\mu}_n + (Az_n \ , \ \check{z}_{n-1})\check{\mu}_n = 0 \ .
\end{aligned} \tag{3.23}
$$

After picking x_0 and $\check{x}_0$, the way we determine the vectors $x_n,\ \check{x}_n,\ n = 1, 2, \ldots$, is now clear. First, we have $x_a = x_0$ and $\check{x}_a = \check{x}_0$. Next, we compute x_{a+1} and $\check{x}_{a+1}$ as in (3.2) with (3.6) and (3.7) with (3.9), respectively. With the help of these we now have z_{a+1} and $\check{z}_{a+1}$. Invoking the biorthogonality requirement in (3.15) and adding the extra criterion in (3.18), we compute z_{a+2} and $\check{z}_{a+2}$ in (3.16) through (3.17) and (3.20). This also enables us to compute x_{a+2} and $\check{x}_{a+2}$. We now continue with $n = a+2, a+3, \ldots$, and compute z_{n+1} and $\check{z}_{n+1}$ as in (3.13) with the help of the equations in (3.22) and (3.23), which are a result of the biorthogonality property in (3.15) and the additional criterion in (3.21).

Needless to say, all the above will be true as long as the process can be continued, i.e., as long as the $\omega_n,\ \mu_n,\ \nu_n,\ \check{\omega}_n,\ \check{\mu}_n,\ \check{\nu}_n$ can be determined uniquely. Examining the equations in (3.17), (3.20), (3.22), and (3.23), we see that a necessary condition for the process not to fail is that $(\check{z}_i \ , \ z_i) \neq 0$ for $i = a+1, a+2, \ldots$, in addition to (3.15). We shall say more on the following Algorithm DBi-CG at the end of Section 4.

ELA

We have shown above a way of constructing vectors x_n that can be obtained by a 4-term recursion relation of the form given in (3.10). From the way this construction takes place it is clear that the amount of computing for each vector x_n is fixed, i.e., does not increase with n. In addition, the vectors that need to be stored in the memory are fixed in number too. We have, however, not shown yet that these vectors are of the form described in the previous section. Neither do we know anything about their other properties. Finally, we have no knowledge about their quality as approximations to $A^{\mathrm{D}}b$. The algebraic properties of the x_n are explored completely in Theorem 3.2 that forms one of the most important developments of this work, after the method that we have developed for computing the x_n.

The following lemma will be very useful in the proof of this theorem.

LEMMA 3.1. *Let R_i and S_i, $i = 0, 1, 2, \ldots,$ be matrices in $\mathbb{C}^{N\times N}$ and let $u_i = R_i w$, $\breve{u}_i = R_i^* \breve{w}$, $v_i = S_i w$ and $\breve{v}_i = S_i^* \breve{w}$ for all i, where $w,\ \breve{w} \in \mathbb{C}^N$. Assume that $S_i R_j = R_j S_i$ for all i and j, and $S_i T = T S_i$ and $R_i T = T R_i$ for all i, where T is some matrix in $\mathbb{C}^{N\times N}$. Then $(u_m\ ,\ T^*\breve{v}_n) = (Tv_n\ ,\ \breve{u}_m)$.*

Proof. From the fact that the R_i commute with the S_j and that T commutes with all the R_i and S_j, we have

$$\begin{aligned}(u_m\ ,\ T^*\breve{v}_n) &= (R_m w\ ,\ T^* S_n^* \breve{w}) = (w\ ,\ R_m^* T^* S_n^* \breve{w}) \\ &= (w\ ,\ S_n^* T^* R_m^* \breve{w}) = (T S_n w\ ,\ R_m^* \breve{w}) = (T v_n\ ,\ \breve{u}_m)\ . \qquad \square\end{aligned}$$

THEOREM 3.2. *Assuming that we have generated the vectors x_n, $n = 0, 1, \ldots,$ successfully as described above, we have the following:*

(i) $\mathrm{span}\{A^{a+1}r_0, A^{a+2}r_0, \ldots, A^n r_0\} = \mathrm{span}\{z_{a+1}, z_{a+2}, \ldots, z_n\}$
$= \mathrm{span}\{z_{a+1}, Az_{a+1}, \ldots, A^{n-a-1}z_{a+1}\}$.

(ii) $\mathrm{span}\{(A^*)^{a+1}\breve{r}_0, (A^*)^{a+2}\breve{r}_0, \ldots, (A^*)^n\breve{r}_0\} = \mathrm{span}\{\breve{z}_{a+1}, \breve{z}_{a+2}, \ldots, \breve{z}_n\}$
$= \mathrm{span}\{\breve{z}_{a+1}, A^*\breve{z}_{a+1}, \ldots, (A^*)^{n-a-1}\breve{z}_{a+1}\}$.

(iii) $(\breve{z}_i\ ,\ z_n) = 0$ *and* $(z_i\ ,\ \breve{z}_n) = 0$, $\quad i = a+1, \ldots, n-1$.

(iv) $r_n \perp \mathrm{span}\{\breve{z}_{a+1}, \ldots, \breve{z}_n\}$ *and* $\breve{r}_n \perp \mathrm{span}\{z_{a+1}, \ldots, z_n\}$.

(v) $r_n = p_n(A)r_0$ *and* $\breve{r}_n = [p_n(A)]^*\breve{r}_0$, $\quad p_n(\lambda) \in \Pi_n^0$.

(vi) $z_n = s_n(A)r_0$ *and* $\breve{z}_n = [s_n(A)]^*\breve{r}_0$, $\quad$ *where* $\ s_n(\lambda) = p_n(\lambda) - p_{n-1}(\lambda)$.

Proof. Part (vi) of the theorem is an immediate consequence of (3.12) and part (v). Thus we need prove only parts $(i) - (v)$. We will do this by using induction on all five parts simultaneously.

We start with $n = a+1$. For the proof of parts (i) and (ii) it is enough to observe that $z_{a+1} = -\rho A^{a+1} r_0$ and $\breve{z}_{a+1} = -\breve{\rho}(A^*)^{a+1}\breve{r}_0$, which follow from (3.1), (3.3), (3.4), (3.7) and (3.12). There is nothing to prove for part (iii) when $n = a+1$. Part (iv) is true as ρ and $\breve{\rho}$ are determined by actually imposing (3.5) and (3.8), which means that $r_{a+1} \perp \breve{z}_{a+1}$ and $\breve{r}_{a+1} \perp z_{a+1}$ in view of parts (i) and (ii). As for part (v), we note that $r_{a+1} = p_{a+1}(A)r_0$, where $p_{a+1}(\lambda) = 1 - \rho\lambda^{a+1}$. We also note that $(A^{a+1}r_0\ ,\ \breve{r}_0) = (r_0\ ,\ (A^*)^{a+1}\breve{r}_0) = \overline{((A^*)^{a+1}\breve{r}_0\ ,\ r_0)}$. Substituting this in (3.9) and comparing the result with (3.6), we obtain $\breve{\rho} = \overline{\rho}$. This implies that $\breve{r}_{a+1} = [p_{a+1}(A)]^*\breve{r}_0$. This proves part (v).

To complete the induction basis we next need to verify that the assertions above are true also for $n = a+2$. The reason for this is that the z_k and $\breve{z}_k$ with the smallest k that can be computed by using the recursion relations in (3.13), are z_{a+2} and $\check{z}_{a+2}$ as is clear from the initial conditions given in (3.14). The truth of parts (i) and (ii) is obvious from (3.16). The validity of part (iii) is immediate as z_{a+2} and $\check{z}_{a+2}$ are constructed such that $z_{a+2} \perp \check{z}_{a+1}$ and $\check{z}_{a+2} \perp z_{a+1}$ explicitly. As for part (iv), we proceed as follows: from the fact that $r_{a+2} = r_{a+1} + z_{a+2}$ and $r_{a+1} \perp \check{z}_{a+1}$ and $z_{a+2} \perp \check{z}_{a+1}$ it follows that $r_{a+2} \perp \check{z}_{a+1}$. Also $r_{a+2} \perp A^*\check{z}_{a+1}$. But since $z_{a+2} \in \text{span}\{\check{z}_{a+1},\ A^*\check{z}_{a+1}\}$, we see that $r_{a+2} \perp \check{z}_{a+2}$ as well. Thus $r_{a+2} \perp \text{span}\{\check{z}_{a+1},\ \check{z}_{a+2}\}$. An analogous argument leads to the conclusion that $\check{r}_{a+2} \perp \text{span}\{z_{a+1},\ z_{a+2}\}$. Finally, for the proof of part (v), we start by taking the complex conjugate of the equations in (3.17) and (3.20) that involve only ω_{a+1} and μ_{a+1}. This gives the system

$$(Az_{a+1}\ ,\ \check{z}_{a+1})\overline{\omega_{a+1}} + (z_{a+1}\ ,\ \check{z}_{a+1})\overline{\mu_{a+1}} = 0\ ,$$
$$(r_{a+1}\ ,\ A^*\check{z}_{a+1}) + (Az_{a+1}\ ,\ A^*\check{z}_{a+1})\overline{\omega_{a+1}} + (z_{a+1}\ ,\ A^*\check{z}_{a+1})\overline{\mu_{a+1}} = 0\ .$$

This system can also be written as

$$(z_{a+1}\ ,\ A^*\check{z}_{a+1})\overline{\omega_{a+1}} + (z_{a+1}\ ,\ \check{z}_{a+1})\overline{\mu_{a+1}} = 0\ ,$$
$$(r_{a+1}\ ,\ A^*\check{z}_{a+1}) + (Az_{a+1}\ ,\ A^*\check{z}_{a+1})\overline{\omega_{a+1}} + (Az_{a+1}\ ,\ \check{z}_{a+1})\overline{\mu_{a+1}} = 0\ .$$

Finally, by the fact that $r_{a+1},\ \check{r}_{a+1},\ z_{a+1}$ and $\check{z}_{a+1}$ are as in parts (v) and (vi), and by Lemma 3.1, we have that $(r_{a+1}\ ,\ A^*\check{z}_{a+1}) = (Az_{a+1}\ ,\ \check{r}_{a+1})$ in the second equation above. Thus, $\overline{\omega_{a+1}}$ and $\overline{\mu_{a+1}}$ satisfy the same equations as $\check{\omega}_{a+1}$ and $\check{\mu}_{a+1}$. Consequently, $\check{\omega}_{a+1} = \overline{\omega_{a+1}}$ and $\check{\mu}_{a+1} = \overline{\mu_{a+1}}$. This, together with the fact that part (v) holds for r_{a+1} and $\check{r}_{a+1}$, implies that part (v) holds for r_{a+2} and $\check{r}_{a+2}$ as well. Part (vi) again follows from part (v).

Let us now assume that the assertions of the theorem are true for $n \geq a+2$ and show that they are true for $n+1$ as well. The truth of parts (i) and (ii) is obvious from (3.13). For part (iii) we need to show that $(z_{n+1}\ ,\ \check{z}_i) = 0$ and $(\check{z}_{n+1}\ ,\ z_i) = 0$ for $i = a+1,\ldots,n$. We already have that these hold for $i = n$ and $i = n-1$ by the way z_{n+1} and $\check{z}_{n+1}$ are constructed. By the induction hypothesis it is clear that $(\check{z}_i\ ,\ z_{n+1}) = \omega_n(\check{z}_i\ ,\ Az_n)$ and $(z_i\ ,\ \check{z}_{n+1}) = \check{\omega}_n(z_i\ ,\ A^*\check{z}_n)$ for $i = a+1,\ldots,n-2$. But $(\check{z}_i\ ,\ Az_n) = (A^*\check{z}_i\ ,\ z_n) = 0$ and $(z_i\ ,\ A^*\check{z}_n) = (Az_i\ ,\ \check{z}_n) = 0$ for $i = a+1,\ldots,n-2$, by $A^*\check{z}_i \in \text{span}\{\check{z}_{a+1},\ldots,\check{z}_{n-1}\}$ and $Az_i \in \text{span}\{z_{a+1},\ldots,z_{n-1}\}$, and by the induction hypothesis. This completes the proof of part (iii).

For part (iv) we need to show that $(r_{n+1}\ ,\ \check{z}_i) = 0$ and $(\check{r}_{n+1}\ ,\ z_i) = 0$, $i = a+1,\ldots,n+1$. From the fact that $r_{n+1} = z_{n+1} + r_n$, from $(z_{n+1}\ ,\ \check{z}_i) = 0$, $i = a+1,\ldots,n$, which we have just shown, and from the induction hypothesis that $(r_n\ ,\ \check{z}_i) = 0, i = a{+}1,\ldots,n$, we already have that $(r_{n+1}\ ,\ \check{z}_i) = 0, i = a{+}1,\ldots,n$. We also have $(r_{n+1}\ ,\ A^*\check{z}_n) = 0$, which, by the fact that $A^*\check{z}_n \in \text{span}\{\check{z}_{a+1},\ldots,\check{z}_{n+1}\}$, implies that $(r_{n+1}\ ,\ \check{z}_{n+1}) = 0$. An analogous argument applies to $\check{r}_{n+1}$. This completes the proof of part (iv).

For the proof of part (v) we start by taking the complex conjugate of the equations in (3.22). After also invoking $(u\ ,\ Av) = (A^*u\ ,\ v)$ in appropriate places we obtain

$$(Az_n\ ,\ \check{z}_{n-1})\overline{\omega_n} + (z_{n-1}\ ,\ \check{z}_{n-1})\overline{\nu_n} = 0\ ,$$

$$(z_n\ ,\ A^*\check{z}_n)\overline{\omega_n} + (z_n\ ,\ \check{z}_n)\overline{\mu_n} = 0\ ,$$
$$(r_n\ ,\ A^*\check{z}_n) + (Az_n\ ,\ A^*\check{z}_n)\overline{\omega_n} + (Az_n\ ,\ \check{z}_n)\overline{\mu_n} + (z_{n-1}\ ,\ A^*\check{z}_n)\overline{\nu_n} = 0\ .$$

Now by Lemma 3.1 we have $(Az_n\ ,\ \check{z}_{n-1}) = (z_{n-1}\ ,\ A^*\check{z}_n)$, $(r_n\ ,\ A^*\check{z}_n) = (Az_n\ ,\ \check{r}_n)$, and $(z_{n-1}\ ,\ A^*\check{z}_n) = (Az_n\ ,\ \check{z}_{n-1})$. When we substitute these in the equations above, we realize that these equations become identical to the equations in (3.23), that involve $\check{\omega}_n$, $\check{\mu}_n$ and $\check{\nu}_n$. This, of course, means that $\check{\omega}_n = \overline{\omega_n}$, $\check{\mu}_n = \overline{\mu_n}$ and $\check{\nu}_n = \overline{\nu_n}$. This along with (3.11) and the induction hypothesis proves that part (v) holds for r_{n+1} and $\check{r}_{n+1}$.

This completes the proof of the theorem. □

REMARK 3.3. From (3.1), (3.2), and (3.12) it is clear that x_n is as in (2.1). In other words, $x_n \in x_0 + \mathcal{K}_{n-a}$ with $\mathcal{K}_{n-a} = \text{span}\{A^a r_0, A^{a+1}r_0, \ldots, A^{n-1}r_0\}$. From this and from part (iv) of the theorem it is seen that the method we have developed is characterized by the additional condition that $r_n \perp \mathcal{L}_{n-a} = \text{span}\{(A^*)^{a+1}\check{r}_0, (A^*)^{a+2}\check{r}_0, \ldots, (A^*)^n\check{r}_0\}$. Here $\mathcal{K}_{n-a}$ and $\mathcal{L}_{n-a}$ serve as right and left subspaces of a suitable projection method for $Ax = b$ with A singular and $\text{ind}\,(A) = a$. This condition can also be expressed in the equivalent form $A^a r_n \perp \text{span}\{A^*\check{r}_0, (A^*)^2\check{r}_0, \ldots, (A^*)^{n-a}\check{r}_0\}$. Precisely this puts DBi-CG in the unified framework of Sidi [17].

In summary, we have devised a Lanczos type method for the Drazin-inverse solution in which $x_n \in x_0 + \mathcal{K}_{n-a}$ such that $r_n = b - Ax_n \perp \mathcal{L}_{n-a}$ with $\check{r}_0 = b - A^*\check{x}_0$ and $\check{x}_0$ arbitrary. We have also provided a recursive algorithm for it whose length is fixed and independent of a. In the next section we refine this algorithm in an appropriate fashion.

4. The DBi-CG algorithm. In this section we will make use of the developments in the previous section, including Theorem 3.2, to devise a Bi-CG type algorithm for the vectors x_n. We recall that the x_n and $\check{x}_n$ satisfy 4-term recursion relation

$$\begin{aligned} x_{n+1} &= x_n + \omega_n A(x_n - x_{n-1}) + \mu_n(x_n - x_{n-1}) + \nu_n(x_{n-1} - x_{n-2})\ , \\ \check{x}_{n+1} &= \check{x}_n + \overline{\omega_n}A^*(\check{x}_n - \check{x}_{n-1}) + \overline{\mu_n}(\check{x}_n - \check{x}_{n-1}) + \overline{\nu_n}(\check{x}_{n-1} - \check{x}_{n-2}) \end{aligned} \tag{4.1}$$

with appropriate initial conditions on $x_a = x_0,\ x_{a+1},\ x_{a+2}$ and $\check{x}_a = \check{x}_0,\ \check{x}_{a+1},\ \check{x}_{a+2}$. Here we have already invoked Theorem 3.2 to replace $\check{\omega}_n$, $\check{\mu}_n$, $\check{\nu}_n$ by $\overline{\omega_n}$, $\overline{\mu_n}$, $\overline{\nu_n}$, respectively. Defining

$$\begin{aligned} d_n &= A(x_n - x_{n-1}) + \frac{\mu_n}{\omega_n}(x_n - x_{n-1}) + \frac{\nu_n}{\omega_n}(x_{n-1} - x_{n-2})\ , \\ \check{d}_n &= A^*(\check{x}_n - \check{x}_{n-1}) + \frac{\overline{\mu_n}}{\overline{\omega_n}}(\check{x}_n - \check{x}_{n-1}) + \frac{\overline{\nu_n}}{\overline{\omega_n}}(\check{x}_{n-1} - \check{x}_{n-2})\ , \end{aligned} \tag{4.2}$$

we obtain from (4.1) that the iterates x_n and $\check{x}_n$ of our method satisfy

$$\begin{aligned} x_{n+1} - x_n &= \omega_n d_n\ , \\ \check{x}_{n+1} - \check{x}_n &= \overline{\omega_n}\check{d}_n\ , \end{aligned} \tag{4.3}$$

and consequently,

$$\begin{aligned} r_{n+1} &= b - Ax_{n+1} = r_n - \omega_n A d_n \ , \\ \check{r}_{n+1} &= b - A^*\check{x}_{n+1} = \check{r}_n - \overline{\omega_n} A^* \check{d}_n \ . \end{aligned}$$

Defining

$$v_n = Ad_n \quad \text{and} \quad \check{v}_n = A^*\check{d}_n \ , \tag{4.4}$$

and recalling (3.12), we obtain

$$z_{n+1} = -\omega_n v_n \quad \text{and} \quad \check{z}_{n+1} = -\overline{\omega_n}\check{v}_n \ . \tag{4.5}$$

Substituting (4.3) in (4.2) and letting

$$\delta_n = \frac{\mu_n}{\omega_n} \quad \text{and} \quad \gamma_n = \frac{\nu_n \omega_{n-2}}{\omega_n \omega_{n-1}} \ ,$$

we realize that

$$\begin{aligned} d_n &= \omega_{n-1}(Ad_{n-1} + \delta_n d_{n-1} + \gamma_n d_{n-2}) \ , \\ \check{d}_n &= \overline{\omega_{n-1}}(A^*\check{d}_{n-1} + \overline{\delta_n}\check{d}_{n-1} + \overline{\gamma_n}\check{d}_{n-2}) \ . \end{aligned} \tag{4.6}$$

Let us now compute the coefficients δ_n, γ_n and ω_n: from (3.22) we obtain

$$\begin{aligned} \delta_n = \frac{\mu_n}{\omega_n} &= -\frac{(\check{z}_n \ , \ Az_n)}{(\check{z}_n \ , \ z_n)} = -\frac{(A^*\check{z}_n \ , \ z_n)}{(\check{z}_n \ , \ z_n)} \\ &= -\frac{(-\overline{\omega_{n-1}}A^*\check{v}_{n-1} \ , \ -\omega_{n-1}v_{n-1})}{(-\overline{\omega_{n-1}}\check{v}_{n-1} \ , \ -\omega_{n-1}v_{n-1})} \\ &= -\frac{(A^*\check{v}_{n-1} \ , \ v_{n-1})}{(\check{v}_{n-1} \ , \ v_{n-1})} \ , \quad n \geq a+1 \ , \end{aligned}$$

$$\begin{aligned} \gamma_n = \frac{\nu_n\omega_{n-2}}{\omega_n\omega_{n-1}} &= -\frac{(\check{z}_{n-1} \ , \ Az_n)\omega_{n-2}}{(\check{z}_{n-1} \ , \ z_{n-1})\omega_{n-1}} \\ &= -\frac{(-\overline{\omega_{n-2}}\check{v}_{n-2} \ , \ -\omega_{n-1}Av_{n-1})\omega_{n-2}}{(-\overline{\omega_{n-2}}\check{v}_{n-2} \ , \ -\omega_{n-2}v_{n-2})\omega_{n-1}} \\ &= -\frac{(\check{v}_{n-2} \ , \ Av_{n-1})}{(\check{v}_{n-2} \ , \ v_{n-2})} \ , \quad n \geq a+2 \ . \end{aligned}$$

From Theorem 3.2 part (iv) we have $r_{n+1} \perp \check{z}_{n+1}$. Thus, $r_{n+1} \perp \check{v}_n$. Recalling that $r_{n+1} = r_n + z_{n+1} = r_n - \omega_n v_n$, we obtain

$$(\check{v}_n \ , \ r_n) - (\check{v}_n \ , \ v_n)\omega_n = 0 \ .$$

Hence

$$\omega_n = \frac{(\check{v}_n \ , \ r_n)}{(\check{v}_n \ , \ v_n)} \ .$$

ELA

The only thing we have to do now is to define a recursion for v_n and $\check{v}_n$. Substituting (4.6) in (4.4) we obtain:

$$v_n = \omega_{n-1}(Av_{n-1} + \delta_n v_{n-1} + \gamma_n v_{n-2}) \ ,$$
$$\check{v}_n = \overline{\omega_{n-1}}(A^* \check{v}_{n-1} + \overline{\delta_n} \check{v}_{n-1} + \overline{\gamma_n} \check{v}_{n-2}) \ .$$

Combining all the above we can now give the following algorithm that we denote the DBi-CG. Here are the steps of this algorithm:

ALGORITHM 4.1.

Step 1. *Pick* x_0 *and* $\check{x}_0$ *arbitrarily, compute* $r_0 = b - Ax_0, \ \check{r}_0 = b - A^*\check{x}_0$, *and set* $x_a = x_0, \ r_a = r_0, \ v_{a-1} = A^a r_0, \ \check{v}_{a-1} = (A^*)^a \check{r}_0, \ \omega_{a-1} = 1$ *;* $n = a$ *;*

Step 2. *while* $\|residual\| \ > tolerance$ *do*
begin

$$\text{if } n \geq a+1 \text{ then } \delta_n := -\frac{(A^*\check{v}_{n-1} \ , \ v_{n-1})}{(\check{v}_{n-1} \ , \ v_{n-1})} \ , \qquad \text{else } \ \delta_n := 0 \ ;$$

$$\text{if } n \geq a+2 \text{ then } \gamma_n := -\frac{(\check{v}_{n-2} \ , \ Av_{n-1})}{(\check{v}_{n-2} \ , \ v_{n-2})} \ , \qquad \text{else } \ \gamma_n := 0 \ ;$$

$$d_n := \omega_{n-1}(v_{n-1} + \delta_n d_{n-1} + \gamma_n d_{n-2}) \ ;$$

$$v_n := \omega_{n-1}(Av_{n-1} + \delta_n v_{n-1} + \gamma_n v_{n-2}) \ ;$$

$$\check{v}_n := \overline{\omega_{n-1}}(A^*\check{v}_{n-1} + \overline{\delta_n}\check{v}_{n-1} + \overline{\gamma_n}\check{v}_{n-2}) \ ;$$

$$\omega_n := \frac{(\check{v}_n \ , \ r_n)}{(\check{v}_n \ , \ v_n)} \ ;$$

$$r_{n+1} := r_n - \omega_n v_n \ ;$$

$$x_{n+1} := x_n + \omega_n d_n \ ;$$

$$n := n+1$$

end;

It is obvious that we do not need $\check{x}_0$, but we need to pick $\check{r}_0$. In addition, we can pick $\check{r}_0$ arbitrarily.

From the steps of Algorithm DBi-CG it is clear that the process will continue as long as $(\check{v}_n \ , \ v_n) \neq 0$ and $(\check{v}_n \ , \ r_n) \neq 0$, which together guarantee that ω_n is well defined and $\omega_n \neq 0$. That $\omega_n \neq 0$ as long as x_n is not the solution is obvious as we need $r_{n+1} \neq r_n$ and $x_{n+1} \neq x_n$ in this case. It is seen that problems will arise when $(\check{v}_n\ ,\ v_n) \approx 0$ and/or $(\check{v}_n\ ,\ r_n) \approx 0$, which are known as "breakdowns" in the literature of the Lanczos method and Bi-CG. A few approaches to overcome these problems, among those the look-ahead strategies, have been suggested in the past and references to them can be found in , e.g., [2]. As mentioned in [2], sometimes breakdowns can be satisfactorily avoided by a restart at the iteration step immediately before the breakdown step. We shall not pursue this subject here any further.

5. Error analysis for DBi-CG. In this section we will present error analysis for the vectors x_n obtained by the DBi-CG method for two different cases: (i) A Hermitian and (ii) A non-Hermitian.

Let $\mathcal{R}(B)$ and $\mathcal{N}(B)$ stand, respectively, for the range and null space of a matrix B. Then it is known that $\mathbb{C}^N = \mathcal{R}(A^a) \oplus \mathcal{N}(A^a)$ and that every vector $z \in \mathbb{C}^N$ can be written in the form $z = \hat{z} + \tilde{z}$, where $\hat{z} \in \mathcal{R}(A^a)$ and $\tilde{z} \in \mathcal{N}(A^a)$ and they are unique. Furthermore, the Drazin-inverse solution of $Ax = b$, namely, the vector $A^{\mathrm{D}}b$ is in $\mathcal{R}(A^a)$.

We first recall Theorem 4.3 of Sidi [17] concerning the finite termination of Krylov subspace methods in general and DBi-CG in particular, for the Drazin-inverse solution of singular systems $Ax = b$.

THEOREM 5.1. *Let $x_0 = \hat{x}_0 + \tilde{x}_0$, where $\hat{x}_0 \in \mathcal{R}(A^a)$ and $\tilde{x}_0 \in \mathcal{N}(A^a)$. Then, for some integer $n_0 \leq \dim \mathcal{R}(A^a) + a \leq N$, we have $x_{n_0} = A^{\mathrm{D}}b + \tilde{x}_0$ and equivalently $A^a r_{n_0} = 0$.*

5.1. Error analysis for DBi-CG applied to singular inconsistent Hermitian linear systems. We shall first look at the case in which the matrix A is Hermitian semidefinite or indefinite, for which $\operatorname{ind}(A) = 1$ necessarily. Since $A^* = A$ now, the DBi-CG algorithm assumes the following simple form, provided we also pick $\check{x}_0 = x_0$.

ALGORITHM 5.2.

Step 1. *Pick x_0 arbitrarily, compute $r_0 = b - Ax_0$, and set*
$x_1 = x_0,\ r_1 = r_0,\ v_0 = Ar_0,\ \omega_0 = 1$;
$n = 1$;

Step 2. *while* $\|residual\| > tolerance$ *do*
begin

$$\textit{if } n \geq 2 \textit{ then } \delta_n := -\frac{(Av_{n-1}\ ,\ v_{n-1})}{(v_{n-1}\ ,\ v_{n-1})}\ , \qquad \textit{else}\ \ \delta_n := 0\ ;$$

$$\textit{if } n \geq 3 \textit{ then } \gamma_n := -\frac{(v_{n-2}\ ,\ Av_{n-1})}{(v_{n-2}\ ,\ v_{n-2})}\ , \qquad \textit{else}\ \ \gamma_n := 0\ ;$$

$$d_n := \omega_{n-1}(v_{n-1} + \delta_n d_{n-1} + \gamma_n d_{n-2})\ ;$$

$$v_n := \omega_{n-1}(Av_{n-1} + \delta_n v_{n-1} + \gamma_n v_{n-2})\ ;$$

$$\omega_n := \frac{(v_n\ ,\ r_n)}{(v_n\ ,\ v_n)}\ ;$$

$$r_{n+1} := r_n - \omega_n v_n\ ;$$

$$x_{n+1} := x_n + \omega_n d_n\ ;$$

$$n := n+1$$

end;

Obviously, there are no breakdowns in this algorithm as $(v_n, v_n) > 0$ for all n except when $v_n = 0$.

We also note that this algorithm is identical to the Conjugate Residual type algorithm of Calvetti, Reichel, and Zhang [4] up to scaling.

The following is a direct consequence of Theorem 3.2.

THEOREM 5.3.

1. $(v_i\ ,\ v_j) = 0\ , \qquad i \neq j$
2. $\mathrm{span}\{d_1, \ldots, d_n\} = \mathrm{span}\{Ar_0, \ldots, A^n r_0\}\ , \qquad n \geq 1$
3. $(A^i r_0\ ,\ r_j) = 0\ , \qquad 2 \leq i \leq j$
4. $(v_i\ ,\ r_j) = 0\ , \qquad 1 \leq i \leq j-1$

The error analysis for the case of Hermitian inconsistent linear systems has been given in the paper by Calvetti, Reichel, and Zhang [4]. We shall state it in the following lemma.

LEMMA 5.4. *Denote the semi-norm*

$$\|z\|' = (Az\ ,\ Az)^{1/2} = \|Az\|\ , \qquad z \in \mathbb{C}^N\ .$$

Let $\mathcal{R}(A)$ denote the range of A and let $\mathcal{N}(A)$ denote the null space of A. Let $x_0 = \hat{x}_0 + \tilde{x}_0$, where $\hat{x}_0 \in \mathcal{R}(A)$ and $\tilde{x}_0 \in \mathcal{N}(A)$. Then the iterate x_n determined by the above algorithm satisfies

$$\|x_n - (\hat{s} + \tilde{x}_0)\|' \leq \|\hat{s} - \hat{x}_0\|' \min_{u \in \Pi_{n-2}} \max_{\lambda \in \sigma(A)\backslash\{0\}} |1 - \lambda^2 u(\lambda)|\ ,$$

where $\hat{s}$ denotes the Drazin-inverse solution of (1.1) and $\sigma(A)$ is the spectrum of A.

We also bring here the convergence rate for the case in which the spectrum of A is nonnegative:

$$\sigma(A) \subset \{0\} \cup [c-d\ ,\ c+d], \qquad 0 < d < c\ . \tag{5.1}$$

LEMMA 5.5. *Denote*

$$\kappa = \exp[-\cosh^{-1}(c/d)] = \frac{c - \sqrt{c^2 - d^2}}{d} < 1\ . \tag{5.2}$$

Then

$$\min_{p \in \Pi_n^0} \max_{\lambda \in [c-d, c+d]} |p(\lambda)| \leq 2(\kappa^{-1} - \kappa) n \kappa^n [1 + o(1)] \quad as \quad n \to \infty\ .$$

Proof. The proof is given in the paper by Hanke and Hochbruck [10]. □

From these lemmas and by the fact that

$$\min_{u\in\Pi_{n-2}} \max_{\lambda\in\sigma(A)\setminus\{0\}} |1-\lambda^2 u(\lambda)| \le \min_{p\in\Pi_n^0} \max_{\lambda\in[c-d,c+d]} |p(\lambda)|$$

and by the assumption that A is Hermitian, there holds

$$\|x_n - (\hat{s}+\tilde{x}_0)\|' \le 2\|\hat{s}-\hat{x}_0\|'(\kappa^{-1}-\kappa)n\kappa^n[1+o(1)] \quad \text{as} \quad n\to\infty \ .$$

5.2. Error analysis for DBi-CG applied to singular non-Hermitian linear systems. We now consider the DBi-CG algorithm applied to the consistent or inconsistent linear system $Ax = b$, where $A \in \mathbb{C}^{N\times N}$ is a non-Hermitian singular matrix of index a. The error analysis we are about to present is inspired by the works of Saad [13] and [14].

With the vectors r_0 and $\check{r}_0$ as before, define the $N\times(n-a)$ matrices V_{n-a} and W_{n-a} by

$$V_{n-a} = [A^a r_0 \mid A^{a+1} r_0 \mid \ldots \mid A^{n-1} r_0]$$

and

$$W_{n-a} = [(A^*)^{a+1}\check{r}_0 \mid (A^*)^{a+2}\check{r}_0 \mid \ldots \mid (A^*)^n \check{r}_0] \ .$$

As before, the span of the columns of V_{n-a} (respectively W_{n-a}) will be denoted by $\mathcal{K}_{n-a}$ (respectively $\mathcal{L}_{n-a}$).

In the sequel we assume that the matrices V_{n-a} and W_{n-a} have full (column) rank and satisfy

$$\det\left(W_{n-a}^* V_{n-a}\right) \ne 0 \ , \tag{5.3}$$

and

$$\det\left(W_{n-a}^* A V_{n-a}\right) \ne 0 \ . \tag{5.4}$$

Now the method we developed in Section 3 is a projection method onto $\mathcal{K}_{n-a}$ and orthogonal to $\mathcal{L}_{n-a}$: it obtains an approximate solution x_n to the Drazin-inverse solution of the singular system $Ax = b$ that belongs to $x_0 + \mathcal{K}_{n-a}$ and satisfies the orthogonality relations

$$r_n = b - Ax_n \perp \mathcal{L}_{n-a} \ . \tag{5.5}$$

It is more convenient to write x_n in the form

$$x_n = x_0 + u_n \ , \qquad u_n \in \mathcal{K}_{n-a} \ .$$

The condition in (5.5) then implies

$$r_0 - Au_n \perp \mathcal{L}_{n-a} \ . \tag{5.6}$$

Thus, u_n is an approximation to the Drazin-inverse solution of the singular system

$$Au = r_0 \ . \tag{5.7}$$

Writing

$$u_n = V_{n-a}y \ , \tag{5.8}$$

it is immediate that y must satisfy the $(n-a)\times(n-a)$ linear system

$$W_{n-a}^*(r_0 - AV_{n-a}y) = 0 \ ,$$

which, by (5.4), has a unique solution for y that is given by

$$y = (W_{n-a}^*AV_{n-a})^{-1}W_{n-a}^*r_0 \ .$$

Noting that $r_0 = \hat{r}_0 + \tilde{r}_0$, where $\hat{r}_0 \in \mathcal{R}(A^a)$ and $\tilde{r}_0 \in \mathcal{N}(A^a)$ and they are unique, and that

$$\begin{aligned} W_{n-a}^*r_0 &= \left((A^*)^a[A^*\check{r}_0 \mid (A^*)^2\check{r}_0 \mid \ldots \mid (A^*)^{n-a}\check{r}_0]\right)^* r_0 \\ &= [A^*\check{r}_0 \mid (A^*)^2\check{r}_0 \mid \ldots \mid (A^*)^{n-a}\check{r}_0]^* A^a r_0 \\ &= [A^*\check{r}_0 \mid (A^*)^2\check{r}_0 \mid \ldots \mid (A^*)^{n-a}\check{r}_0]^* A^a \hat{r}_0 \\ &= \left((A^*)^a[A^*\check{r}_0 \mid (A^*)^2\check{r}_0 \mid \ldots \mid (A^*)^{n-a}\check{r}_0]\right)^* \hat{r}_0 = W_{n-a}^*\hat{r}_0 \ , \end{aligned}$$

we realize that y is given by

$$y = (W_{n-a}^*AV_{n-a})^{-1}W_{n-a}^*\hat{r}_0 \ . \tag{5.9}$$

Let $x_0 = \hat{x}_0 + \tilde{x}_0$, where $\hat{x}_0 \in \mathcal{R}(A^a)$ and $\tilde{x}_0 \in \mathcal{N}(A^a)$, and let $\hat{u}$ be the Drazin-inverse solution to (5.7), i.e., $\hat{u} = A^{\mathrm{D}}r_0$. Denote also $\hat{s} = A^{\mathrm{D}}b$. Thus, $\hat{u} = \hat{s} - \hat{x}_0$.

Let P_n be the orthogonal projector onto the subspace $\mathcal{K}_{n-a}$. In the following lemma we study the error $x_n - (\hat{s} + \tilde{x}_0)$ in terms of the distance

$$\varepsilon_n = \|(I - P_n)\hat{u}\| \ ; \quad \hat{u} = \hat{s} - \hat{x}_0 \ , \tag{5.10}$$

where $\|\cdot\|$ denotes the Euclidean norm.

LEMMA 5.6. *The distance $\|(I-P_n)\hat{u}\|$ between $\hat{u}$ and the Krylov subspace $\mathcal{K}_{n-a}$ satisfies*

$$\|(I - P_n)\hat{u}\| = \min_{p\in\Pi_n^0} \|p(A)(\hat{x}_0 - \hat{s})\| \ . \tag{5.11}$$

Proof. It is known that

$$\begin{aligned} \|(I - P_n)\hat{u}\| &= \min_{u\in\mathcal{K}_{n-a}} \|\hat{u} - u\| \\ &= \min_{q\in\Pi_{n-1-a}} \|\hat{u} - q(A)A^a r_0\| \\ &= \min_{q\in\Pi_{n-1-a}} \|\hat{u} - q(A)A^{a+1}\hat{u}\| \\ &= \min_{q\in\Pi_{n-1-a}} \|\left(I - A^{a+1}q(A)\right)\hat{u}\| = \min_{p\in\Pi_n^0} \|p(A)(\hat{x}_0 - \hat{s})\| \ . \end{aligned}$$

ELA

Here we have used the facts that $A^a r_0 = A^a \hat{r}_0$ and $A^{a+1}\hat{u} = A^a \hat{r}_0$. □

We now turn to the error analysis of our method. We will need an interpretation of the method in terms of operator equations. Let us define the operator Q_n onto $\hat{r}_0 + \mathcal{K}_{n-a}$ and orthogonal to $\mathcal{L}_{n-a}$ by

$$Q_n x \in \hat{r}_0 + \mathcal{K}_{n-a} \text{ and } x - Q_n x \perp \mathcal{L}_{n-a} \ .$$

LEMMA 5.7. *The matrix representation of Q_n in the standard basis in $\mathbb{C}^N$ is*

$$Q_n x = \hat{r}_0 + V_{n-a}\left(W_{n-a}^* V_{n-a}\right)^{-1} W_{n-a}^*(x - \hat{r}_0) \ .$$

Proof. We have by definition of Q_n

$$Q_n x = \hat{r}_0 + V_{n-a} y \quad \text{and} \quad W_{n-a}^*(x - Q_n x) = 0 \ . \tag{5.12}$$

Hence

$$W_{n-a}^* x = W_{n-a}^* Q_n x = W_{n-a}^* \hat{r}_0 + W_{n-a}^* V_{n-a} y \ ,$$

from which we obtain

$$y = \left(W_{n-a}^* V_{n-a}\right)^{-1} W_{n-a}^*(x - \hat{r}_0) \ . \tag{5.13}$$

Substituting (5.13) in (5.12), the result follows. □

LEMMA 5.8. *Q_n is a projector onto $\hat{r}_0 + \mathcal{K}_{n-a}$.*

Proof.

$$\begin{aligned}
Q_n^2 x = Q_n(Q_n x) &= \hat{r}_0 + V_{n-a}\left(W_{n-a}^* V_{n-a}\right)^{-1} W_{n-a}^*(Q_n x - \hat{r}_0)\\
&= \hat{r}_0 + V_{n-a}\left(W_{n-a}^* V_{n-a}\right)^{-1} W_{n-a}^*\left(V_{n-a}\left(W_{n-a}^* V_{n-a}\right)^{-1} W_{n-a}^*(x - \hat{r}_0)\right)\\
&= \hat{r}_0 + V_{n-a}\left(W_{n-a}^* V_{n-a}\right)^{-1}\left(W_{n-a}^* V_{n-a}\right)\left(W_{n-a}^* V_{n-a}\right)^{-1} W_{n-a}^*(x - \hat{r}_0)\\
&= \hat{r}_0 + V_{n-a}\left(W_{n-a}^* V_{n-a}\right)^{-1} W_{n-a}^*(x - \hat{r}_0) = Q_n x \ . \qquad \square
\end{aligned}$$

Note that the vector $Q_n x$ is uniquely defined only when (5.3) holds.

Let us now define the operator A_n by

$$A_n = Q_n A P_n \ .$$

We then have the following result.

LEMMA 5.9. *The problem*

$$u \in \mathcal{K}_{n-a} \quad \textit{and} \quad \hat{r}_0 - A_n u = 0 \tag{5.14}$$

has as its unique solution the vector u_n defined following (5.5).

Proof. It is sufficient to express the problem (5.14) in matrix notation. Since $u \in \mathcal{K}_{n-a}$, it can be written as

$$u = V_{n-a} y \ . \tag{5.15}$$

Furthermore, $P_n u = u$. Thus from (5.14) we obtain

$$\begin{aligned}\hat{r}_0 - A_n u &= \hat{r}_0 - Q_n A P_n u = \hat{r}_0 - Q_n A u \\ &= \hat{r}_0 - \left(\hat{r}_0 + V_{n-a}\left(W_{n-a}^* V_{n-a}\right)^{-1} W_{n-a}^*(Au - \hat{r}_0)\right) \\ &= V_{n-a}\left(W_{n-a}^* V_{n-a}\right)^{-1}\left(W_{n-a}^* \hat{r}_0 - W_{n-a}^* A V_{n-a} y\right) = 0 \ .\end{aligned}$$

The columns of matrix V_{n-a} are linearly independent , hence

$$W_{n-a}^* \hat{r}_0 - W_{n-a}^* A V_{n-a} y = 0 \ ,$$

which yields

$$y = \left(W_{n-a}^* A V_{n-a}\right)^{-1} W_{n-a}^* \hat{r}_0 \ . \tag{5.16}$$

This means that the problem (5.14) has a unique solution. The proof is completed by realizing that the solution given in (5.15) and (5.16) is identical to that given in (5.8)–(5.9). □

We shall refer to the problem (5.14) as the "approximate problem". What Lemma 5.9 shows is that the method described in Section 3.2 amounts to replacing the problem (5.7) by the "approximate problem". Our next task is to give a bound for the residual of $\hat{u}$ for problem (5.14). This is considered in the next lemma.

LEMMA 5.10. *Let ε be as in (5.10) and define*

$$\theta_n = \|Q_n A(I - P_n)\| \ . \tag{5.17}$$

Then the Drazin-inverse solution $\hat{u}$ for the problem (5.7) satisfies

$$\|\hat{r}_0 - A_n \hat{u}\| \leq \theta_n \varepsilon_n \ .$$

Proof. Since $\hat{r}_0$ belongs to $\hat{r}_0 + \mathcal{K}_{n-a}$ and since Q_n is a projection operator onto $\hat{r}_0 + \mathcal{K}_{n-a}$, one can see that $\hat{r}_0 = Q_n \hat{r}_0$. Using this fact and the fact that $A\hat{u} = \hat{r}_0$, we realize that

$$\begin{aligned}\hat{r}_0 - A_n \hat{u} &= Q_n \hat{r}_0 - Q_n A P_n \hat{u} = Q_n(\hat{r}_0 - A P_n \hat{u}) \\ &= Q_n(A\hat{u} - A P_n \hat{u}) = Q_n A(I - P_n)\hat{u} \\ &= Q_n A(I - P_n)(I - P_n)\hat{u} \ .\end{aligned}$$

The result now follows. □

LEMMA 5.11. *The operator $A_n|_{\mathcal{K}_{n-a}}$, namely, the restriction of A_n to the subspace $\mathcal{K}_{n-a}$, is invertible.*

Proof. We have to show that the equation

$$(A_n|_{\mathcal{K}_{n-a}})u = \hat{r}_0 + v \ , \quad v \in \mathcal{K}_{n-a} \tag{5.18}$$

has a unique solution $u \in \mathcal{K}_{n-a}$. Since $u, v \in \mathcal{K}_{n-a}$, we have

$$u = V_{n-a} z \ \text{ and } \ v = V_{n-a} y \ , \qquad y, z \in \mathbb{C}^{n-a} \ .$$

Invoking the definitions of A_n and Q_n, (5.18) becomes

$$\hat{r}_0 + V_{n-a}(W^*_{n-a}V_{n-a})^{-1}W^*_{n-a}(AP_nu - \hat{r}_0) = \hat{r}_0 + V_{n-a}y \ .$$

Recalling that the columns of V_{n-a} are linearly independent, we realize that the last equation is equivalent to

$$(W^*_{n-a}V_{n-a})^{-1}W^*_{n-a}(AP_nu - \hat{r}_0) = y \ . \tag{5.19}$$

Multiplying both sides of (5.19) by $W^*_{n-a}V_{n-a}$ and using the fact that $P_nu = u$ we obtain

$$W^*_{n-a}AV_{n-a}z = W^*_{n-a}\hat{r}_0 + W^*_{n-a}V_{n-a}y \ .$$

Since $W^*_{n-a}AV_{n-a}$ is invertible by (5.4), a unique solution for z exists. This completes the proof. □

THEOREM 5.12. *Let ε and θ_n be as defined in (5.10) and (5.17) respectively, and let $\kappa_n = \|(A_n|_{\mathcal{K}_{n-a}})^{-1}\|$. Then the error $x_n - (\hat{s} + \tilde{x}_0)$ satisfies*

$$\|x_n - (\hat{s} + \tilde{x}_0)\| \le (1 + \theta_n^2\kappa_n^2)^{1/2}\varepsilon_n \ . \tag{5.20}$$

Proof. By Lemma 5.9 and Lemma 5.11 we have that $u_n = (A_n|_{\mathcal{K}_{n-a}})^{-1}\hat{r}_0$. Recalling that $P_nu_n = u_n$ by the fact that $u_n \in \mathcal{K}_{n-a}$ we obtain

$$\begin{aligned} P_n(u_n - \hat{u}) &= u_n - P_n\hat{u} \\ &= (A_n|_{\mathcal{K}_{n-a}})^{-1}[\hat{r}_0 - (A_n|_{\mathcal{K}_{n-a}})P_n\hat{u}] \ . \end{aligned}$$

Next, since $P_n\hat{u} \in \mathcal{K}_{n-a}$ and $P_n^2 = P_n$, we have $(A_n|_{\mathcal{K}_{n-a}})P_n\hat{u} = A_nP_n\hat{u} = A_n\hat{u}$. Therefore,

$$P_n(u_n - \hat{u}) = (A_n|_{\mathcal{K}_{n-a}})^{-1}(\hat{r}_0 - A_n\hat{u}) \ .$$

Using Lemma 5.10, we realize that

$$\|P_n(u_n - \hat{u})\| \le \kappa_n\theta_n\varepsilon_n \ .$$

Writing

$$\hat{u} - u_n = \hat{u} - P_nu_n = (I - P_n)\hat{u} + P_n(\hat{u} - u_n)$$

and observing that the two vectors on the right-hand side of the above equation are orthogonal (by the fact that P_n and also $I - P_n$ are orthogonal projectors), we obtain

$$\|u_n - \hat{u}\|^2 = \|(I - P_n)\hat{u}\|^2 + \|P_n(\hat{u} - u_n)\|^2 \le \varepsilon_n^2 + \kappa_n^2\theta_n^2\varepsilon_n^2 \ .$$

Hence, recalling also that $x_n = x_0 + u_n$,

$$\begin{aligned} \|x_n - (\hat{s} + \tilde{x}_0)\| &= \|x_0 + u_n - \hat{s} - \tilde{x}_0\| \\ &= \|u_n - (\hat{s} - \hat{x}_0)\| = \|u_n - \hat{u}\| \le (1 + \kappa_n^2\theta_n^2)^{1/2}\varepsilon_n \ . \end{aligned}$$ □

Now θ_n can be bounded through $\theta_n \leq \|Q_n\|\|A\|$ since $\|I - P_n\| = 1$ by the fact that $I - P_n$ is an orthogonal projector. Therefore, $\theta_n \kappa_n \leq \|Q_n\|\|A\|\|(A_n|_{\mathcal{K}_{n-a}})^{-1}\|$. That is, $\theta_n \kappa_n$ plays the role of a condition number for the "approximate problem". Thus, Theorem 5.12 implies that, as long as the "approximate problem" is not too badly conditioned, the error $x_n - (\hat{s} + \tilde{x}_0)$ is of the order of $\varepsilon_n = \|(I - P_n)\hat{u}\|$.

As far as ε_n is concerned, we have the following result that can be proved in exactly the same way as Theorem 6.1 of Sidi [17].

THEOREM 5.13. *Choose Ω to be a closed domain that contains $\sigma(A) \setminus \{0\}$ but not $\lambda = 0$, such that its boundary is twice differentiable with respect to arclength. Denote by $\Phi(\lambda)$ the conformal mapping of the exterior of Ω onto the exterior of the unit disk $\{w : |w| \geq 1\}$. Then*

$$\varepsilon_n \leq K n^{a+2(\hat{k}-1)} \rho^n \ ,$$

where $K > 0$ is some constant independent of n, $\hat{k} = \max\{k_j : \ k_j = \mathrm{ind}\,(A - \lambda_j I),\ \lambda_j \in \sigma(A) \setminus \{0\}\}$, and and $\rho = 1/|\Phi(0)| < 1$.

This theorem shows that ε_n ultimately tends to 0 exponentially in n.

6. Numerical examples. We tested the numerical properties of the DBi-CG method for the examples taken from the paper by Climent, Neumann, and Sidi [6]. The algorithm developed in Section 4 was used to compute the eigenprojection $Z_A := I - AA^{\mathrm{D}}$ onto the eigenspace of A corresponding to the zero eigenvalue of three singular matrices whose index exceeds 1.

By Theorem 5.1 there exists a smallest integer $n_0 \leq \dim \hat{\mathcal{R}}(A^a) + a \leq N$, for which $x_{n_0} = \hat{s} + \tilde{x}_0$, provided $\det(W^*_{n_0-a} A V_{n_0-a}) \neq 0$. Thus, if we take $b = 0$, then $\hat{s} = A^{\mathrm{D}} b = 0$, and

$$x_{n_0} = \tilde{x}_0 = (I - AA^{\mathrm{D}}) x_0 \ .$$

Now, if we choose x_0 as the ith column of I, the vector $(I - AA^{\mathrm{D}})x_0$ is the ith column of the eigenprojection Z_A. Using DBi-CG and stopping when $\omega_n = 0$ or $\dfrac{\|x_{n+1} - x_n\|_\infty}{\|x_n\|_\infty} \leq 10^{-15}$, we have obtained exact eigenprojections Z_A in at most N iterations. (In all these examples N varies between $6 - 8$.)

We have also considered the solution of the Poisson equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) u(x,y) = f(x,y)$$

on the unit square

$$\Omega = [0,1] \times [0,1] = \{(x,y) : 0 \leq x \leq 1\ ,\ 0 \leq y \leq 1\}$$

with Neumann boundary conditions

$$\frac{\partial}{\partial n} u(x,y) = \varphi(x,y) \quad \text{on } \partial\Omega\ .$$

This problem was also considered by Hanke and Hochbruck [10] for testing the Chebyshev type semi-iterative method that was developed there.

First, we replace Ω by the uniform grid

$$\Omega_h = \{(x_j, y_k) = (jh, kh) : 0 \le j \le M \ , \ 0 \le k \le M\} \ ,$$

where $h = 1/M$. Thus we obtain a grid which contains $M+1$ rows and $M+1$ points in each row, hence a total of $(M+1)^2$ points. Next, we discretize the Laplace operator and the boundary conditions with central differences in the standard way. Finally, we take M to be an odd integer and arrange the grid points using the red-black ordering. As a result, we end up with the following $(M+1)^2 \times (M+1)^2$ nonsymmetric coefficient matrix A:

$$\left[\begin{array}{cccccccc|cccccccc}
4I & O & \cdots & \cdots & \cdots & \cdots & \cdots & O & T_1 & -2I & O & \cdots & \cdots & \cdots & \cdots & O \\
O & 4I & \ddots & & & & & \vdots & -I & T_2 & -I & O & & & & \vdots \\
\vdots & \ddots & \ddots & \ddots & & & & \vdots & O & -I & T_1 & -I & O & & & \vdots \\
\vdots & & \ddots & \ddots & \ddots & & & \vdots & \vdots & O & -I & T_2 & -I & O & & \vdots \\
\vdots & & & \ddots & \ddots & \ddots & & \vdots & \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & & & \ddots & \ddots & \ddots & \vdots & \vdots & & & \ddots & \ddots & \ddots & \ddots & O \\
\vdots & & & & & \ddots & 4I & O & \vdots & & & & O & -I & T_1 & -I \\
O & \cdots & \cdots & \cdots & \cdots & \cdots & O & 4I & O & \cdots & \cdots & \cdots & \cdots & O & -2I & T_2 \\
\hline
T_2 & -2I & O & \cdots & \cdots & \cdots & \cdots & O & 4I & O & \cdots & \cdots & \cdots & \cdots & \cdots & O \\
-I & T_1 & -I & O & & & & \vdots & O & 4I & \ddots & & & & & \vdots \\
O & -I & T_2 & -I & O & & & \vdots & \vdots & \ddots & \ddots & \ddots & & & & \vdots \\
\vdots & O & -I & T_1 & -I & O & & \vdots & \vdots & & \ddots & \ddots & \ddots & & & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & & & \ddots & \ddots & \ddots & & \vdots \\
\vdots & & & \ddots & \ddots & \ddots & \ddots & O & \vdots & & & & \ddots & \ddots & \ddots & \vdots \\
\vdots & & & & O & -I & T_2 & -I & \vdots & & & & & \ddots & 4I & O \\
O & \cdots & \cdots & \cdots & \cdots & O & -2I & T_1 & O & \cdots & \cdots & \cdots & \cdots & \cdots & O & 4I
\end{array}\right]$$

where

$$T_1 = \begin{bmatrix} -2 & 0 & \cdots & \cdots & 0 \\ -1 & -1 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 & 0 \\ 0 & \cdots & 0 & -1 & -1 \end{bmatrix}_{\frac{M+1}{2} \times \frac{M+1}{2}}$$

and

$$T_2 = \begin{bmatrix} -1 & -1 & 0 & \dots & 0 \\ 0 & -1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & -1 & -1 \\ 0 & \dots & \dots & 0 & -2 \end{bmatrix}_{\frac{M+1}{2}\times\frac{M+1}{2}}$$

and I and O are, respectively, the $\frac{M+1}{2} \times \frac{M+1}{2}$ identity and zero matrices. Note that A is singular with a one dimensional null space spanned by the vector $e = [1 \cdots 1]^T$. Even if the continuous problem has a solution, the discretized problem need not be consistent. In the sequel we consider only the Drazin-inverse solution of $Ax = b$ for arbitrary b, not necessarily related to f and φ.

We first construct a consistent system with known solution $\hat{s} \in \mathcal{R}(A)$ via $\hat{s} = Ay$, where $y = [0 \cdots\ 0\ 1]^T$. Then we perturb $A\hat{s}$, the right-hand side of $Ax = A\hat{s} = \hat{b}$, with a constant multiple of the null space vector e. In this way we end up with an inconsistent system with Drazin-inverse solution $\hat{s}$. For this example our perturbation amounts to one percent in norm, i.e., $\dfrac{\|\tilde{b}\|_2}{\|\hat{b}\|_2} = 0.01$. The initial vector x_0 is the zero vector.

In our numerical experiments we took $M = 63$. Therefore, the number of unknowns is 4096. With $M = 63$ the solution we are looking for is the vector $\hat{s}$, whose components are zeros except

$$\hat{s}_{2016} = -1\ ,\ \hat{s}_{2047} = -1\ ,\ \hat{s}_{2048} = -2\ ,\ \hat{s}_{4096} = 4\ .$$

Using DBi-CG and stopping when $\dfrac{\|x_{n+1} - x_n\|_\infty}{\|x_n\|_\infty} \le 2 \times 10^{-9}$, after 230 iterations we obtain the following results:

n	$x_{n,2016}$	$x_{n,2047}$	$x_{n,2048}$	$x_{n,4096}$
10	-1.010098531489	-1.0222184392605	-2.030941821045	3.9681647141065
50	-1.000219129362	-0.9994840140640	-1.999193023834	4.0008989000809
100	-1.000006579682	-0.9999900670946	-1.999983487480	4.0000185531542
150	-1.000001210051	-0.9999979782034	-1.999996693250	4.0000037089364
200	-1.000000176275	-0.9999997734417	-1.999999589236	4.0000004580209
230	-0.999999977161	-0.9999999353241	-1.999999922886	4.0000000816397

All our computations in this section have been performed in FORTRAN 77 using double-precision arithmetic.

7. Conclusions. In this work we have developed a Krylov subspace method of the Bi-Conjugate Gradient (Bi-CG) type for computing the Drazin-inverse solution of singular linear system $Ax = b$, $A \in \mathbb{C}^{N\times N}$. We have not put any restrictions on the matrix A. Thus, A is not necessarily Hermitian or Hermitian positive semidefinite, and can have any type of spectrum. In addition, it can have an arbitrary index.

Neither have we put any restrictions on the system $Ax = b$. This system may be consistent or inconsistent. We are only required to know a, the index of A. In this method the approximation to $A^{\mathrm{D}}b$, the Drazin-inverse solution of $Ax = b$, is of the form $x_n = x_0 + q_{n-1}(A)r_0$, where x_0 is the initial vector, $r_0 = b - Ax_0$, and $q_{n-1}(\lambda)$ is a polynomial of degree at most $n-1$, and the polynomial $p_n(\lambda) = 1 - \lambda q_{n-1}(\lambda) \in \Pi_n^0$, with Π_n^0 as defined in (2.5). It follows a posteriori that the coefficients of $q_{n-1}(\lambda)$ are determined by requiring that $A^a r_n$, where $r_n = b - Ax_n$, be orthogonal to the $(n-a)$-dimensional subspace $W = \text{span}\{t, A^*t, \ldots, (A^*)^{n-a-1}t\}$, where t is a suitable vector. After showing the relevance and theoretical validity of this approach, we have given a detailed error analysis for the above method in both Hermitian and non-Hermitian cases. Finally, we have presented some numerical experiments with the new algorithm.

The present work seems to be the first to present a Krylov subspace method for singular non-Hermitian consistent or inconsistent linear systems with arbitrary index. One important achievement of this work is the development of a recursive algorithm of fixed length independently of the size of the index of A.

REFERENCES

[1] W. E. Arnoldi. The principal of minimized iterations in the solution of the matrix eigenvalue problem. *Quart. Appl. Math.*, 9:17–29, 1951.

[2] R. Barrett et al. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods.* SIAM, Philadelphia, 1994.

[3] A. Ben-Israel and T. N. E. Greville. *Generalized Inverses: Theory and Applications.* Wiley, New York, 1974.

[4] D. Calvetti, L. Reichel, and Q. Zhang. Conjugate Gradient algorithms for symmetric inconsistent linear systems. In *Proceedings of the Cornelius Lanczos International Centenary Conference*, J. D. Brown et al., eds., pp.267–272. SIAM, Philadelphia, 1994.

[5] S. L. Campbell and C. D. Meyer, Jr. *Generalized Inverses of Linear Transformations.* Pitman, London, 1979.

[6] J.-J. Climent, M. Neumann, and A. Sidi. A semi-iterative method for real spectrum singular linear systems with an arbitrary index. *J. Comput. Appl. Math.*, 87:21–38, 1997.

[7] M. Eiermann, I. Marek, and W. Niethammer. On the solution of singular linear systems of algebraic equations by semi-iterative methods. *Numer. Math.*, 53:265–283, 1988.

[8] B. Fischer, M. Hanke, and M. Hochbruck. A note on conjugate gradient type methods for indefinite and/or inconsistent linear systems. *Numer. Algorithms*, 11:181–187, 1996.

[9] R. Fletcher. Conjugate Gradient methods for indefinite systems. In *Numerical Analysis – Dundee 1975*, G. A. Watson, ed., *Lecture Notes in Mathematics*, volume 506, Springer, New York, 1976, pp.73–89.

[10] M. Hanke and M. Hochbruck. A Chebyshev-like semi-iteration for inconsistent linear systems. *Electron. Trans. Numer. Anal.*, 1:89–103, 1993.

[11] E. F. Kaaschieter. Preconditioned conjugate gradients for solving singular systems. *J. Comput. Appl. Math.*, 24:265–275, 1988.

[12] C. Lanczos. Solution of systems of linear equations by minimized iterations. *J. Res. Nat. Bur. Standards*, 49:33–53, 1952.

[13] Y. Saad. Krylov subspace methods for solving large unsymmetric linear systems. *Math. Comp.*, 37:105–126, 1981.

[14] Y. Saad. The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems. *SIAM J. Numer. Anal.*, 19:485–506, 1982.

[15] A. Sidi. Application of vector extrapolation methods to consistent singular linear systems. *Appl. Numer. Math.*, 6:487–500, 1989–1990.

[16] A. Sidi. Development of iterative techniques and extrapolation methods for Drazin-inverse solution of consistent or inconsistent linear systems. *Linear Algebra Appl.*, 167:171–203, 1992.

[17] A. Sidi. A unified approach to Krylov subspace methods for Drazin-inverse solution of singular nonsymmetric linear systems. *Linear Algebra Appl.*, 298:99–113, 1999.