

# On the inheritance of the Discrete Picard Condition

Silvia Gazzola, Paolo Novati  
Department of Mathematics - University of Padova

February 4, 2015

## Abstract

When projection methods are employed to regularize linear discrete ill-posed problems, one implicitly assumes that the discrete Picard condition is somehow inherited by the projected problems. In this paper we prove that, when considering various Krylov subspace methods, the discrete Picard condition still holds for the projected uncorrupted systems. By exploiting the inheritance of the discrete Picard condition, some estimates on the behavior of the projected problems are also derived. Numerical examples are provided in order to illustrate the accuracy of the derived estimates.

## 1 Introduction

Let  $K : L^2(\Omega) \rightarrow L^2(\Omega)$ , be a linear operator defined by

$$(Kf)(s) := \int_{\Omega} k(s,t)f(t)dt$$

where  $\Omega \subset \mathbb{R}^q$  is compact and Jordan-measurable, and  $k : \Omega \times \Omega \rightarrow \mathbb{R}$  is such that

$$\|k\|^2 := \int_{\Omega} \int_{\Omega} |k(s,t)|^2 dsdt < \infty. \quad (1)$$

In this framework it is known that  $K$  is a compact operator that can be written in terms of its Singular Value Expansion (SVE) as

$$K = \sum_{i=1}^{\infty} \mu_i u_i \langle v_i, \cdot \rangle,$$

where  $\langle \cdot, \cdot \rangle$  is the scalar product on the Hilbert space  $L^2(\Omega)$  and the  $\mu_i$ 's are the singular values of  $K$ . The orthonormal sets  $\{u_i\}_i$  and  $\{v_i\}_i$  are such that

$$k(s,t) = \sum_{i=1}^{\infty} \mu_i u_i(s)v_i(t).$$

Since

$$\|k\|^2 = \sum_{i=1}^{\infty} \mu_i^2,$$

by (1) the singular values decay as  $i^{-\alpha}$ ,  $\alpha > 1/2$ .

In this paper we consider the numerical solution of the linear equation

$$Kf = g, \quad (2)$$

which is ill-posed in the sense of Hadamard (see [11, 14] for a background). The degree of ill-posedness is characterized by the decay rate of the singular values. Equation (2) admits a solution  $f \in L^2(\Omega)$  if and only if the right-hand side  $g$  satisfies the so called *Picard Condition*, that is,

$$\sum_{i=1}^{\infty} \left( \frac{\langle u_i, g \rangle}{\mu_i} \right)^2 < \infty, \quad (3)$$

(cf., for instance, [5, §2.2]). Condition (3) implies that, as  $i \rightarrow \infty$ , the absolute value of the so-called Fourier coefficients  $\langle u_i, g \rangle$  decays to zero as  $i^{-\alpha} \mu_i$ ,  $\alpha > 1/2$ .

After suitable discretization, depending on a parameter  $N$ , the solution of the linear equation (2) is approximated by the solution of a certain linear system (*discrete ill-posed problem*)

$$A^{(N)}x^{(N)} = b^{(N)}, \quad (4)$$

where the matrix  $A^{(N)}$  inherits the spectral properties of  $K$ , and  $x^{(N)}$  represents the function  $f$ . Without loss of generality we may assume that  $A^{(N)} \in \mathbb{R}^{N \times N}$ ,  $b^{(N)} \in \mathbb{R}^N$ . Let us consider the singular value decomposition (SVD) of  $A^{(N)}$ , given by the factorization

$$A^{(N)} = U^{(N)}\Sigma^{(N)}V^{(N)T}, \quad U^{(N)} \in \mathbb{R}^{N \times N}, \Sigma^{(N)} \in \mathbb{R}^{N \times N}, V^{(N)} \in \mathbb{R}^{N \times N},$$

where  $U^{(N)T}U^{(N)} = U^{(N)}U^{(N)T} = I_N$  (i.e., the identity matrix of order  $N$ ),  $V^{(N)T}V^{(N)} = V^{(N)}V^{(N)T} = I_N$ , and  $\Sigma^{(N)} = \text{diag}(\sigma_1^{(N)}, \dots, \sigma_N^{(N)})$ . We assume that, independently of  $N$ , the  $\sigma_i^{(N)}$ 's decay and cluster to zero with the same rate of the  $\mu_i$ 's, with no evident gap between two consecutive ones to indicate the numerical rank for  $A^{(N)}$ . If (2) is discretized by the Galerkin method, and  $N$  is sufficiently large, this is a meaningful assumption (see the analysis in [12]). Since the solution of (4) can be written as

$$x^{(N)} = \sum_{i=1}^N \frac{u_i^{(N)T} b^{(N)}}{\sigma_i^{(N)}} v_i,$$

in order to compute a meaningful solution of (4) the basic assumption is that there exists a constant  $C$  such that

$$\sup_N \sum_{i=1}^N \left( \frac{u_i^{(N)T} b^{(N)}}{\sigma_i^{(N)}} \right)^2 \leq C < \infty, \quad (5)$$

Alike the infinite dimensional problem (2), the above relation, which is called *Discrete Picard Condition* (DPC, see [13]), implies that the Fourier coefficients  $|u_i^{(N)T} b^{(N)}|$  decay to zero as  $i^{-\alpha} \sigma_i^{(N)}$ ,  $\alpha > 1/2$ , in order to ensure the convergence of the series. In this situation, independently of  $N$ , we have

$$\|x^{(N)}\|_2 \leq C < \infty.$$

In practice, when dealing with discrete models like (4), the DPC is not guaranteed to hold since  $b^{(N)}$  is usually affected by some error, i.e.,  $b^{(N)} = b_{ex}^{(N)} + e^{(N)}$ , where the vector  $e^{(N)}$  is assumed unknown. In the following, we always refer to the solution of  $A^{(N)}x^{(N)} = b_{ex}^{(N)}$  as exact solution of (4). In particular, when  $e^{(N)}$  is Gaussian white noise, typically  $|u_i^{(N)T}b^{(N)}|$  decays until  $\sigma_i^{(N)} > \|e^{(N)}\|_2$ , and then stagnates around  $\|e^{(N)}\|_2$  (cf. [14, Chapter 4]). Still in [13], the author shows that the DPC plays an important role in determining how well the Tikhonov or TSVD regularized solutions can approximate the desired exact solution of (4). For this reason, when regularizing perturbed problems of the form (4), one usually assumes that the corresponding unperturbed problems satisfy the DPC, i.e., condition (5) applied to  $b_{ex}^{(N)}$ . Moreover, one can devise a parameter choice strategy based on the so-called Picard plot (i.e., a plot of the quantities  $|u_i^{(N)T}b^{(N)}|$  and  $\sigma_i^{(N)}$  versus  $i, i = 1, \dots, N$ ); the parameter selected by a visual inspection of the Picard plot often agrees with the one selected by other popular parameter choice strategies (such as the L-curve and the GCV methods, cf. again the discussion in [13]).

When dealing with large-scale problems, direct approaches to regularization (such as the TSVD) are often unfeasible, because of their high computational cost; in these cases, just iterative or hybrid approaches to regularization are possible (cf. the discussion in [1]). Among the class of iterative regularization methods, a core role is played by Krylov subspace methods, which allow to compute approximations of the solution of (4) by solving projected subproblems of the form

$$W_k^{(N)T}A^{(N)}Z_k^{(N)}y = W_k^{(N)T}b^{(N)} \quad \text{or} \quad \min_y \left\| W_{k+1}^{(N)T}b^{(N)} - W_{k+1}^{(N)T}A^{(N)}Z_k^{(N)}y \right\|, \quad (6)$$

where  $W_k^{(N)}$  and  $Z_k^{(N)}$  are matrices whose columns span suitable Krylov subspaces of dimension  $k$ . Here and in the following,  $\|\cdot\|$  denotes the Euclidean vector norm. When Krylov subspace methods are employed to solve system (4) with a corrupted right-hand-side vector  $b^{(N)}$ , during the first iterations the approximate solutions typically converge to the exact one (i.e., the solution of (4) with  $b^{(N)} = b_{ex}^{(N)}$ ); then, as soon as the approximate solutions are affected by the high-frequency noise components perturbing  $b^{(N)}$ , they start to diverge. This behavior is known as semiconvergence phenomenon. Because of semiconvergence, a reliable stopping criterion is essential to perform iterative regularization. Alternatively, to overcome semiconvergence, one can resort to hybrid methods, in which the regularization of the projected problem is considered. Hybrid methods were originally introduced in [19] for the Lanczos-bidiagonalization case; hybrid methods based on the Arnoldi algorithm were considered in [4, 7].

During the last two decades, many Krylov subspace methods have been theoretically proved to be regularization methods in the classical sense, i.e., it has been proved that the sequence of the approximate solutions tends to the exact solution of (4) when  $\|e^{(N)}\| \rightarrow 0$  and a suitable stopping criterion is considered (see for instance [10, Chapter 3] and [3]). More recently, the regularization and convergence properties of methods based on the Arnoldi and the Lanczos bidiagonalization algorithms have been analyzed from the point of view of the spectral properties of the projected matrices  $W_k^{(N)T}A^{(N)}Z_k^{(N)}$ . In particular, the authors of [7, 18] experimentally show that the

behavior of the regularized solution obtained by truncating the Arnoldi process is similar to the one obtained by TSVD. Still in [7, 18], the rate of convergence of methods based on the Arnoldi algorithm is shown to be related to the decay rate of the singular values of  $A^{(N)}$ . Regarding hybrid methods, new parameter choice strategies have been devised in [6, 16, 18] under the assumption that the original uncorrupted problem (4) satisfies the DPC. Finally, estimates on the behavior of the GMRES residual for the exact and the corrupted problems are given in [6] under the assumption that the DPC still holds for the projected subproblems (6); this fact has been confirmed by many numerical experiments performed on the most common test problems from [15].

From a theoretical point of view, the investigation of the inheritance of the DPC by the projected subproblems (6) is still an open issue, which represents the goal of this paper. Denoting by  $y_k^{(N)}$  the solution of (6), we say that the DPC is inherited if there exists a constant  $C'$  such that

$$\sup_{k,N} \|y_k^{(N)}\| \leq C' < \infty,$$

To establish the inheritance of the DPC when the exact ( $b^{(N)} = b_{ex}^{(N)}$ ) problems (6) are solved, we use a backward induction argument, i.e., we prove that if the  $(k+1)$ th projected problem satisfies the DPC, so does the  $k$ th one. Thanks to this result, we can derive further estimates on the behavior of the residuals associated to the projection methods (6), and we can give an alternative justification of the typical semiconvergent behavior of the iterative methods applied to the corrupted ( $b^{(N)} = b_{ex}^{(N)} + e^{(N)}$ ) problem (4). Unless strictly necessary, from now on we avoid the use of the superscript  $(N)$ , assuming it implicit by the context.

This paper is organized as follows: in Section 2 we provide a brief overview on the considered Krylov subspace methods, along with their hybrid versions. In Section 3, after giving some more details about the DPC, we prove that the DPC is inherited when employing Arnoldi based methods (Section 3.1) and Lanczos bidiagonalization based methods (Section 3.2); we also analyze the behavior of the residuals for both the uncorrupted and the corrupted projected problems of the form (6). Our derivations are supported by many numerical tests, whose most meaningful results are displayed along Section 3. Finally, in Section 4 we present some concluding remarks.

**Remarks about the numerical tests.** All the considered test problems belong to the package *Regularization Tools* [15]: in particular we display the results for the test problems `baart`, `wing` (whose coefficient matrices are nonsymmetric), `deriv2`, and `shaw` (whose coefficient matrices are symmetric); the coefficient matrix of the system (4) has size  $120 \times 120$  for all the test problems. All the computations have been performed using Matlab 7.10 with 16 significant digits. Unless otherwise stated, we implement the Arnoldi algorithm by Householder reflections [22, Chapter 6], and we implement the Lanczos bidiagonalization algorithm with reorthogonalization following the indications in [21].

## 2 Regularization by Krylov Subspace Methods

Krylov subspaces methods are projection-type methods such that, at the  $k$ th iteration, an approximation  $x_k$  of the solution of (4) is formed by imposing that

$$x_k \in \mathcal{K}_k^{(1)} \quad \text{and} \quad r_k = b - Ax_k \perp \mathcal{K}_k^{(2)}, \quad (7)$$

where  $\mathcal{K}_k^{(i)} = \mathcal{K}_k(C_i, d_i) = \text{span}\{d_i, C_i d_i, \dots, C_i^{k-1} d_i\}$ ,  $C_i \in \mathbb{R}^{N \times N}$ ,  $d_i \in \mathbb{R}^N$ ,  $i = 1, 2$ . Usually the matrices  $C_i$  and the vectors  $d_i$  depend on  $A$  and  $b$ . Formulation (7) intrinsically assumes that the initial guess  $x_0$  is the zero vector. Moreover, in the following we always assume that  $\dim(\mathcal{K}_k^{(i)}) = k$ ,  $i = 1, 2$ .

### 2.1 The Arnoldi algorithm and the GMRES method

The Arnoldi algorithm [22, Chapter 6] underlies many of the most used Krylov methods for linear systems, and it is employed to compute an orthonormal basis for the space  $\mathcal{K}_k(A, b)$ ; more precisely, it leads to the decompositions

$$\begin{aligned} AW_k &= W_k H_k + h_{k+1,k} w_{k+1} e_k^T \\ &= W_{k+1} \bar{H}_k, \end{aligned} \quad (8)$$

where  $W_{k+1} = [W_k \ w_{k+1}] \in \mathbb{R}^{N \times (k+1)}$  has orthonormal columns that span the Krylov subspace  $\mathcal{K}_{k+1}(A, b)$ ,  $e_k$  is the  $k$ th element of the canonical basis of  $\mathbb{R}^k$  (here and in the following the dimension should be clear from the context), and  $w_1 = W_{k+1} e_1 = b / \|b\|_2$ . In the above relations, both  $H_k \in \mathbb{R}^{k \times k}$  and  $\bar{H}_k \in \mathbb{R}^{(k+1) \times k}$  are upper Hessenberg matrices;  $H_k$  is obtained by discarding the last row of  $\bar{H}_k$ . The Arnoldi algorithm terminates as soon as  $h_{k+1,k} = 0$ , which means that an invariant subspace of  $A$  has been computed.

Taking  $\mathcal{K}_k^{(i)} = \mathcal{K}_k(A, b)$ ,  $i = 1, 2$  in (7), the Arnoldi algorithm can be used to construct approximations of the solution of (4) in the following way (Full Orthogonalization Method, FOM)

$$x_k = W_k y_k, \quad \text{where} \quad y_k = (H_k)^{-1} c_k, \quad c_k = \|b\| e_1 \in \mathbb{R}^k. \quad (9)$$

Besides FOM, the most highly-regarded Krylov subspace method based on the Arnoldi algorithm is the GMRES [22, Chapter 6], which is a projection method having  $\mathcal{K}_k^{(1)} = \mathcal{K}_k(A, b)$  and  $\mathcal{K}_k^{(2)} = A\mathcal{K}_k(A, b)$ . At the  $k$ th iteration, the GMRES prescribes to take as approximate solution of (4) the vector

$$x_k = W_k y_k, \quad \text{where} \quad y_k = \arg \min_{y \in \mathbb{R}^k} \left\| \underbrace{\|b\| e_1}_{c_k \in \mathbb{R}^{k+1}} - \bar{H}_k y \right\|. \quad (10)$$

We note that the above projected problem is derived by taking into account relation (8). In the following, we employ the matrix relation

$$AW_k = W_{k+2} \bar{H}_k^0, \quad \text{where} \quad \bar{H}_k^0 = \begin{bmatrix} \bar{H}_k \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^{(k+2) \times k}, \quad (11)$$

which is basically equivalent to (8). Indeed, the solution of the least squares problem

$$\min_{y \in \mathbb{R}^k} \left\| \underbrace{\|b\|e_1}_{c_{k+1} \in \mathbb{R}^{k+2}} - \bar{H}_k^0 y \right\|, \quad (12)$$

solves (10), and viceversa.

Hybrid methods formulated with respect to the Arnoldi algorithm and Tikhonov regularization compute an approximate solution of the form  $x_{k,\lambda} = W_k y_{k,\lambda}$ , where

$$y_{k,\lambda} = \arg \min_{y \in \mathbb{R}^k} \left\{ \|c_k - \bar{H}_k y\|^2 + \lambda^2 \|y\|^2 \right\}; \quad (13)$$

a suitable regularization parameter  $\lambda$  has to be set at each iteration (cf. [7]).

## 2.2 The Lanczos bidiagonalization algorithm and the LSQR method

The Lanczos (Golub-Kahan) bidiagonalization algorithm [8] is employed to compute two matrices  $W_k, Z_k \in \mathbb{R}^{N \times k}$  having orthonormal columns, and being such that  $\mathcal{R}(W_k) = \mathcal{K}_k(A^T A, A^T b)$  and  $\mathcal{R}(Z_k) = \mathcal{K}_k(AA^T, b)$ . At the  $k$ th step, the Lanczos bidiagonalization algorithm can be expressed in matrix form by the following relations

$$\begin{aligned} A^T Z_k &= W_k B_k^T, \\ AW_k &= Z_{k+1} \bar{B}_k, \end{aligned}$$

where both  $B_k \in \mathbb{R}^{k \times k}$  and  $\bar{B}_k \in \mathbb{R}^{(k+1) \times k}$  are lower bidiagonal;  $B_k$  is obtained by deleting the last row of  $\bar{B}_k$ ; if we denote the diagonal elements of  $\bar{B}_k$  by  $\zeta_i$ , and its sub-diagonal elements by  $\nu_{i+1}$ ,  $i = 1, \dots, k$ , the Lanczos bidiagonalization algorithm terminates as soon as  $\zeta_i = 0$  or  $\nu_{i+1} = 0$ . The most highly-regarded Krylov subspace method based on the Lanczos bidiagonalization algorithm is the LSQR [21, 20], which is a projection method having  $\mathcal{K}_k^{(1)} = \mathcal{K}_k(A^T A, A^T b)$  and  $\mathcal{K}_k^{(2)} = A\mathcal{K}_k(A^T A, A^T b)$ . At the  $k$ th iteration, the LSQR prescribes to take as approximate solution of (4) the vector

$$x_k = W_k y_k, \quad \text{where} \quad y_k = \arg \min_{y \in \mathbb{R}^k} \left\| \underbrace{\|b\|e_1}_{c_k \in \mathbb{R}^{k+1}} - \bar{B}_k y \right\|. \quad (14)$$

Hybrid methods formulated with respect to the Lanczos bidiagonalization algorithm and Tikhonov regularization compute an approximate solution of the form  $x_{k,\lambda} = W_k y_{k,\lambda}$ , where

$$y_{k,\lambda} = \arg \min_{y \in \mathbb{R}^k} \left\{ \|c_k - \bar{B}_k y\|^2 + \lambda^2 \|y\|^2 \right\}; \quad (15)$$

a suitable regularization parameter  $\lambda$  has to be set at each iteration (cf. [17] and the references therein).

### 3 Inheritance of the Discrete Picard Condition

As stated in the Introduction, the DPC holds for (4) if

$$|u_i^T b| = O(i^{-\alpha} \sigma_i), \quad \alpha > 1/2, \quad i \rightarrow N \rightarrow \infty.$$

In the following, we theoretically prove the inheritance of the DPC by backward induction. We assume that both the Arnoldi and the Lanczos bidiagonalization algorithms do not break down until step  $N$ . This is not a restrictive assumption: indeed, if  $N^*$  is the smallest integer such that  $h_{N^*+1, N^*} = 0$  (in the Arnoldi case),  $\zeta_{N^*} = 0$  or  $\nu_{N^*+1} = 0$  (in the Lanczos bidiagonalization case), all the following derivations are still valid, provided that  $N$  is replaced by  $N^*$ . We use the notation  $[z]_i$  to denote the  $i$ th component of the vector  $z$ .

#### 3.1 Methods based on the Arnoldi algorithm

Let us consider the SVD of the matrix  $\bar{H}_k^0$  in (12), given by

$$\bar{H}_k^0 = U_k^0 \Sigma_k V_k^T, \quad U_k^0 \in \mathbb{R}^{(k+2) \times k}, \quad \Sigma_k \in \mathbb{R}^{k \times k}, \quad V_k \in \mathbb{R}^{k \times k}, \quad (16)$$

where  $(U_k^0)^T U_k^0 = I_k$ ,  $(V_k)^T V_k = V_k (V_k)^T = I_k$ , and  $\Sigma_k = \text{diag}(\sigma_1^{(k)}, \dots, \sigma_k^{(k)})$ . We remark that the SVD of  $\bar{H}_k$  is closely linked to (16): indeed,

$$\bar{H}_k = U_k \Sigma_k V_k^T, \quad \text{where } U_k \in \mathbb{R}^{(k+1) \times k} \quad \text{and} \quad U_k^0 = \begin{bmatrix} U_k \\ \mathbf{0} \end{bmatrix}. \quad (17)$$

In the following, we will extensively exploit an update formula for the SVD given in [2]. Since  $\bar{H}_k^0$  is obtained by deleting the last column of  $\bar{H}_{k+1}$ , i.e.,

$$\bar{H}_{k+1} = [\bar{H}_k^0, \bar{h}_{k+1}] \in \mathbb{R}^{(k+2) \times (k+1)},$$

we can state that

$$U_k^0 = U_{k+1} X_{k+1}, \quad (18)$$

where  $U_{k+1}$  is the matrix of the left singular values of  $\bar{H}_{k+1}$ . The  $i$ th column  $x_i^{(k+1)}$  of  $X_{k+1} \in \mathbb{R}^{(k+1) \times k}$  is computed by taking

$$x_i^{(k+1)} = \frac{1}{\left\| \begin{pmatrix} (D_i^{(k+1)})^{-1} \\ w^{(k+1)} \end{pmatrix} \right\|} \begin{pmatrix} (D_i^{(k+1)})^{-1} \\ w^{(k+1)} \end{pmatrix} \in \mathbb{R}^{k+1}, \quad i = 1, \dots, k, \quad (19)$$

where

$$D_i^{(k+1)} = (\Sigma_{k+1})^2 - (\sigma_i^{(k)})^2 I_{k+1} \in \mathbb{R}^{(k+1) \times (k+1)}$$

and  $\Sigma_{k+1}$  is the diagonal matrix of the singular values of  $\bar{H}_{k+1}$ . Therefore,  $D_i^{(k+1)}$  is diagonal, and

$$w^{(k+1)} = \frac{1}{\|\bar{h}_{k+1}\|} U_{k+1}^T \bar{h}_{k+1} \in \mathbb{R}^{k+1}. \quad (20)$$

Without loss of generality, in the following we assume that the component  $[w^{(k+1)}]_j$ ,  $j = 1, \dots, k+1$ , of  $w^{(k+1)}$  (20) is different from zero, and that the  $\sigma_j^{(k+1)}$ 's,  $k = 1, \dots, N-1$ , are distinct: this implies that the updated singular values  $\sigma_j^{(k)}$ 's are distinct. If this is not the case, then we have to perform deflation and we basically act on the nonzero components of  $w^{(k+1)}$  as we describe below (see again [2] for the details). The singular values  $\sigma_j^{(k+1)}$ ,  $j = 1, \dots, k+1$ , and  $\sigma_j^{(k)}$ ,  $j = 1, \dots, k$ , are linked by the so-called (strict) interlacing property [9, §8.6.1]

$$\sigma_1^{(k+1)} > \sigma_1^{(k)} > \dots > \sigma_i^{(k+1)} > \sigma_i^{(k)} > \sigma_{i+1}^{(k+1)} > \dots > \sigma_k^{(k)} > \sigma_{k+1}^{(k+1)}. \quad (21)$$

Since the singular values of  $H_N$  and  $A$  obviously coincide (i.e.,  $\sigma_j^{(N)} = \sigma_j$ ), thanks to (21) one has

$$\sigma_j - \sigma_j^{(k+1)} < \sigma_j - \sigma_j^{(k)},$$

which shows that the convergence  $\sigma_j^{(k)} \rightarrow \sigma_j$  for  $k \rightarrow N$  is monotone. The next result requires that the convergence is very fast; even if a theoretical analysis of this behavior is still missing, experimentally it is clear. Figure 1 compares the behavior of the singular values  $\mu_i$ 's (of the continuous operator),  $\sigma_i$ 's,  $\sigma_i^{(50)}$ 's,  $\sigma_i^{(30)}$ 's, and  $\sigma_i^{(10)}$ 's for a couple of test problems.

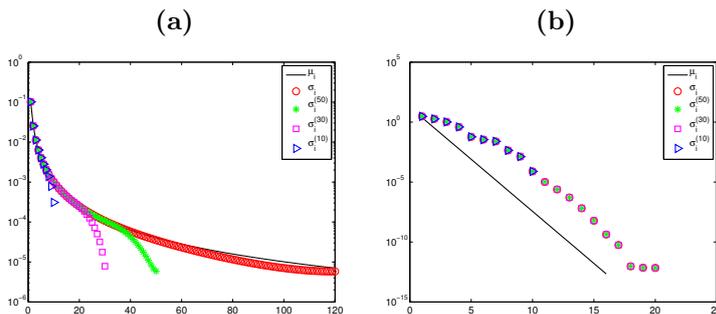


Figure 1: Behavior of the continuous singular values, the singular values of  $A$ , and the singular values of the matrices  $\bar{H}_{50}$ ,  $\bar{H}_{30}$ , and  $\bar{H}_{10}$ . The considered test problems are `deriv2` (frame (a)) and `shaw` (frame (b)). Just the values above  $10^{-14}$  are displayed.

**Theorem 1.** Let (4) satisfy the DPC. Assume that the rate of convergence of  $\{\sigma_i^{(k)}\}_k$  is such that

$$\sum_{j=1}^{k+1} [x_i^{(k+1)}]_j O(j^{-\alpha} \sigma_j^{(k+1)}) = O(i^{-\alpha} \sigma_i^{(k+1)}) \quad (22)$$

for a given  $\alpha > 1/2$ . Then the projected least square problem (10) satisfies the DPC for  $k = 1, \dots, N$ .

*Proof.* Our goal is to prove that  $\|y_k\| \leq C' < \infty$  independently of  $N$ , and for  $k \leq N$ . We proceed by backward induction. The DPC obviously holds for  $k = N$ , since  $x_N = W_N y_N$  solves (4). Next, let us assume that (12) satisfies the DPC for

$2 \leq k+1 \leq N$ , and let us prove that (12) satisfies the DPC for  $k$ . Thanks to relation (18), we can express the norm of the solution of (12) as

$$\begin{aligned} \|y_k\| &= \|(\bar{H}_k^0)^\dagger c_{k+1}\| \\ &= \|\Sigma_k^{-1} (U_k^0)^T c_{k+1}\| \\ &= \|\Sigma_k^{-1} X_{k+1}^T U_{k+1}^T c_{k+1}\| \end{aligned}$$

Thanks to the DPC, there exists  $\alpha > 1/2$  such that

$$\left| [U_{k+1}^T c_{k+1}]_j \right| = O(j^{-\alpha} \sigma_j^{(k+1)}), \quad j = 1, \dots, k+1,$$

and hence

$$\begin{aligned} |[\Sigma_k^{-1} X_{k+1}^T U_{k+1}^T c_{k+1}]_i| &= \frac{1}{\sigma_i^{(k)}} \sum_{j=1}^{k+1} [x_i^{(k+1)}]_j [U_{k+1}^T c_{k+1}]_j \\ &= O(i^{-\alpha}), \quad i = 1, \dots, k, \end{aligned}$$

where we have exploited (22). This concludes the proof.  $\square$

**Remark 2.** The hypothesis (22) substantially means that each row of  $X_{k+1}^T$  is a discretization of the Delta function, or, in other words, that  $X_{k+1}$  must be close to the identity matrix. Despite this may appear as a very strong hypothesis, in practice we can easily see that it is true thanks to the fast convergence of  $\{\sigma_i^{(k)}\}_k$ . Indeed, defining the quantities

$$\varepsilon_{ij}^{(k+1)} := (\sigma_j^{(k+1)})^2 - (\sigma_i^{(k)})^2, \quad j = 1, \dots, k+1, \quad (23)$$

we can write

$$\begin{aligned} [x_i^{(k+1)}]_i^2 &= \left[ \sum_{j=1}^{k+1} \frac{[w^{(k+1)}]_j^2}{\varepsilon_{ij}^{(k+1)2}} \right]^{-1} \frac{[w_i^{(k+1)}]_i^2}{\varepsilon_{ii}^{(k+1)2}} \\ &= \left[ 1 + \frac{\varepsilon_{ii}^{(k+1)2}}{[w^{(k+1)}]_i^2} \sum_{\substack{j=1 \\ j \neq i}}^{k+1} \frac{[w^{(k+1)}]_j^2}{\varepsilon_{ij}^{(k+1)2}} \right]^{-1}. \end{aligned}$$

In this way,  $\left| [x_i^{(k+1)}]_i \right| \simeq 1$  provided that  $\varepsilon_{ii}^{(k+1)2} / \varepsilon_{ij}^{(k+1)2} \simeq 0$  for  $i \neq j$ , which is true if the convergence of  $\{\sigma_i^{(k)}\}_k$  is fast, and if  $\sigma_i^{(k)}$  and  $\sigma_j^{(k+1)}$  are well-separated for  $i \neq j$ . As a clear consequence of  $\|x_i^{(k+1)}\| = 1$  (see (19)),  $\left| [x_i^{(k+1)}]_j \right| \simeq 0$  for  $i \neq j$ . This allows to write

$$x_i^{(k+1)} = e_i - \delta_i^{(k+1)} \quad (24)$$

where

$$0 \leq [\delta_i^{(k+1)}]_i \leq \frac{\varepsilon_{ii}^{(k+1)2}}{[w^{(k+1)}]_i^2} \sum_{\substack{j=1 \\ j \neq i}}^{k+1} \frac{[w^{(k+1)}]_j^2}{\varepsilon_{ij}^{(k+1)2}} \quad (25)$$

and

$$\sum_{\substack{j=1 \\ j \neq i}}^{k+1} \left[ \delta_i^{(k+1)} \right]_j^2 \leq 2 \left[ \delta_i^{(k+1)} \right]_i. \quad (26)$$

To support the above derivations, in Figure 2 we show the behavior of the singular values of the matrices  $\bar{H}_k$  (8) when two consecutive steps of the Arnoldi algorithm applied to the test problem **baart** are performed. We also display the structure of the matrix  $X_{k+1}$  appearing in (18). Referring to (21), for this test problem it is evident that  $\sigma_i^{(k)} \simeq \sigma_i^{(k+1)}$ ,  $i = 1, \dots, k$ . Referring to (24), we can clearly see that, as far as the singular values are well-separated, the  $k$  columns of the matrices  $X_{k+1}$  are essentially the first  $k$  columns of the identity matrix  $I_{k+1}$ : for instance, this is the case for  $X_6$ . However, as soon as some clustering of the singular values happens (typically, the clustered singular values are the smallest ones), the columns of  $X_{k+1}$  corresponding to the clustered singular values are not anymore comparable to the columns of  $I_{k+1}$ . This behavior is particularly evident when we consider the  $k = 19$  case, where we have a cluster starting from the 11th singular value. Correspondingly, the  $i$ th element of the  $i$ th column of  $X_{k+1}$ ,  $11 \leq i \leq 19$ , is of the same order as its neighbors. In Figure 3 we give some examples of Picard plots for the projected problems (10) associated to the **baart** test problem. In this setting, by Picard plot we mean a graph of the Fourier coefficients  $|(u_i^{(k)})^T c_k|$  and of the singular values  $\sigma_i^{(k)}$  versus  $i$  (where  $i = 1, \dots, k$  and  $k = 1, \dots, N$ ). Looking at the Picard plot is perhaps the most immediate way to assess if the considered problem satisfies the DPC (i.e., the DPC is satisfied if the numerically nonzero Fourier coefficients lie approximately below the numerically nonzero singular values). Figures 4 and 5 are analogous to Figure 2, except that the test problems **shaw** and **deriv2** are taken into account, respectively. For the **shaw** test problem we know that the coefficient matrix has  $\sigma_i \simeq \exp(-2i)$ ,  $i = 1, \dots, N$  (see Figure 1) and, therefore, it is severely ill-posed (cf. [14, Chapter 1]). Looking at the plots in the first row of Figure 4 we can state that, typically,  $\sigma_i^{(k)} \simeq \sigma_i^{(k+1)}$ ,  $i = 1, \dots, k$ . In this case (and, more in general, whenever an analytical expression of the decay of the singular values is available), we can derive more precise estimates in Theorem 1 and Remark 2. For instance, assuming that  $\sigma_i^{(k+1)} \simeq \exp(-2i)$ ,  $i = 1, \dots, k+1$ , we can consider the first-order approximation

$$\underbrace{\exp(-4i)}_{(\sigma_i^{(k+1)})^2} - \underbrace{\exp(-4(i + \tilde{\varepsilon}_i))}_{(\sigma_i^{(k)})^2} \simeq 4 \exp(-4i) \tilde{\varepsilon}_i, \quad \tilde{\varepsilon}_i > 0, \tilde{\varepsilon}_i \simeq 0,$$

which gives a more precise estimate for the quantities in (23). We again see that, as soon as clustered singular values appear,  $X_{k+1}$  is not anymore comparable to an identity matrix. If we focus on frames **(b')** and **(c'')**, the 18th and the 19th columns of  $X_{20}$  are not comparable to the 18th and 19th columns of  $I_{20}$ ; correspondingly, looking at frame **(b)**,  $\sigma_{18}^{(20)}$  and  $\sigma_{19}^{(20)}$  are clustered. We remark that  $\sigma_{18}^{(20)}$  and  $\sigma_{19}^{(20)}$  are the only clustered singular values of  $\bar{H}_{20}$ ; indeed, also the first singular values seem clustered, but we have to take into account that they are displayed in logarithmic scale. Analogous remarks hold for the **deriv2** test problem: in this case, we know that the coefficient matrix has  $\sigma_i \simeq (i\pi)^{-2}$ ,  $i = 1, \dots, N$  (see again Figure 1) and, therefore, it is moderately ill-posed

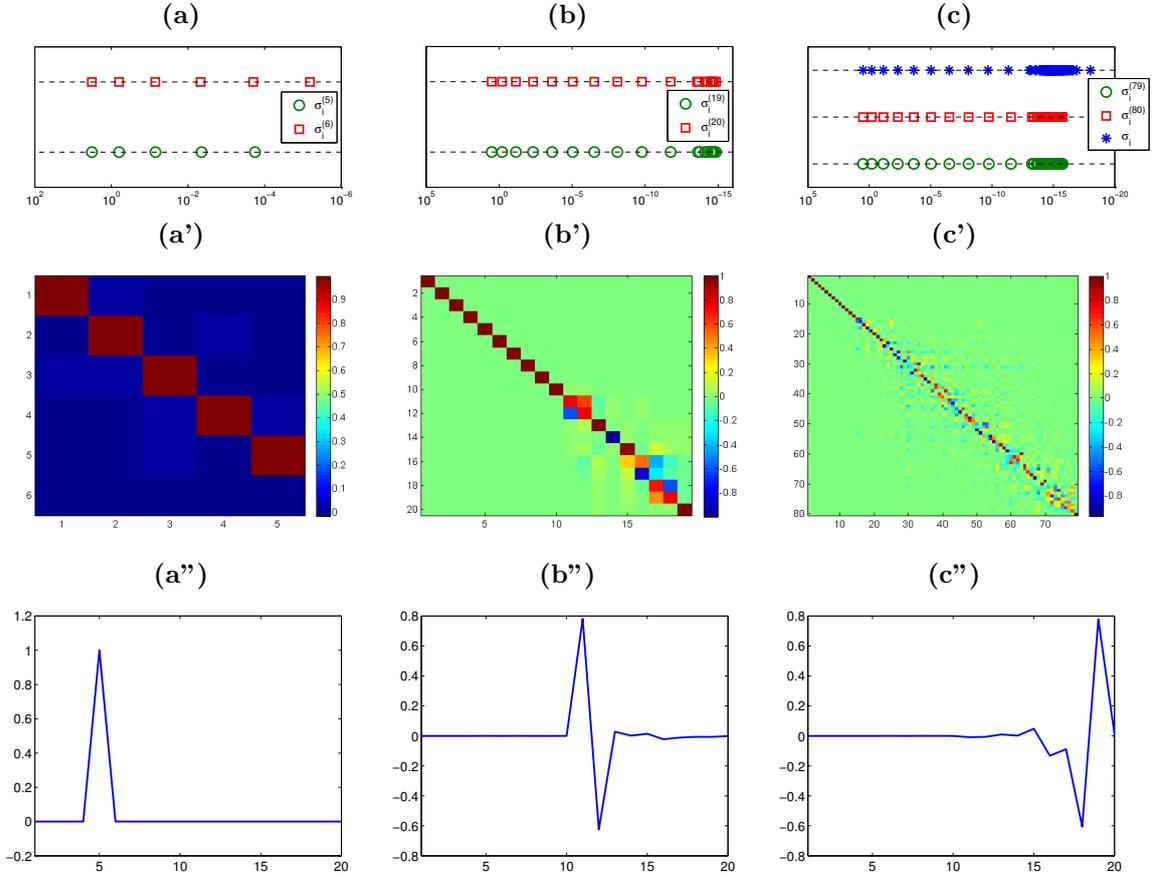


Figure 2: Test problem baart. (a) singular values of  $\bar{H}_6$  and  $\bar{H}_5$ ; (b) singular values of  $\bar{H}_{20}$  and  $\bar{H}_{19}$ ; (c) singular values of  $\bar{H}_{80}$ ,  $\bar{H}_{79}$  and  $A$ ; (a'') matrix  $X_6$ ; (b'') matrix  $X_{20}$ ; (c'') matrix  $X_{80}$ ; (a'') column  $x_5^{(20)}$ ; (b'') column  $x_{11}^{(20)}$ ; (c'') column  $x_{18}^{(20)}$ .

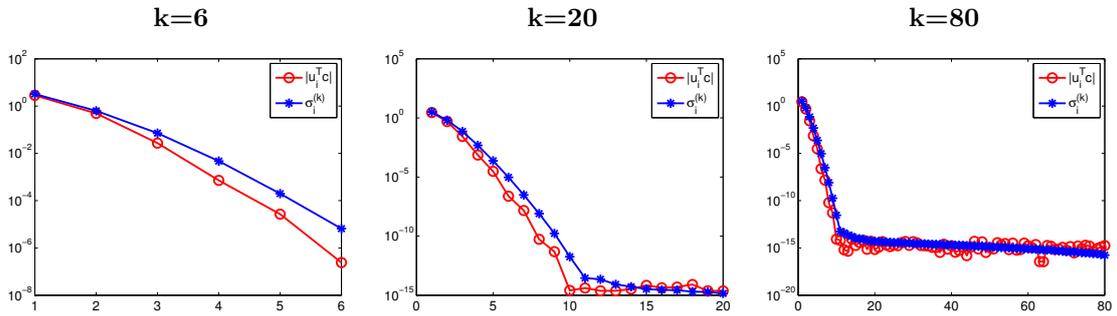


Figure 3: Test problem baart: Picard plots for the  $k$ th projected problems computed by the Arnoldi algorithm.

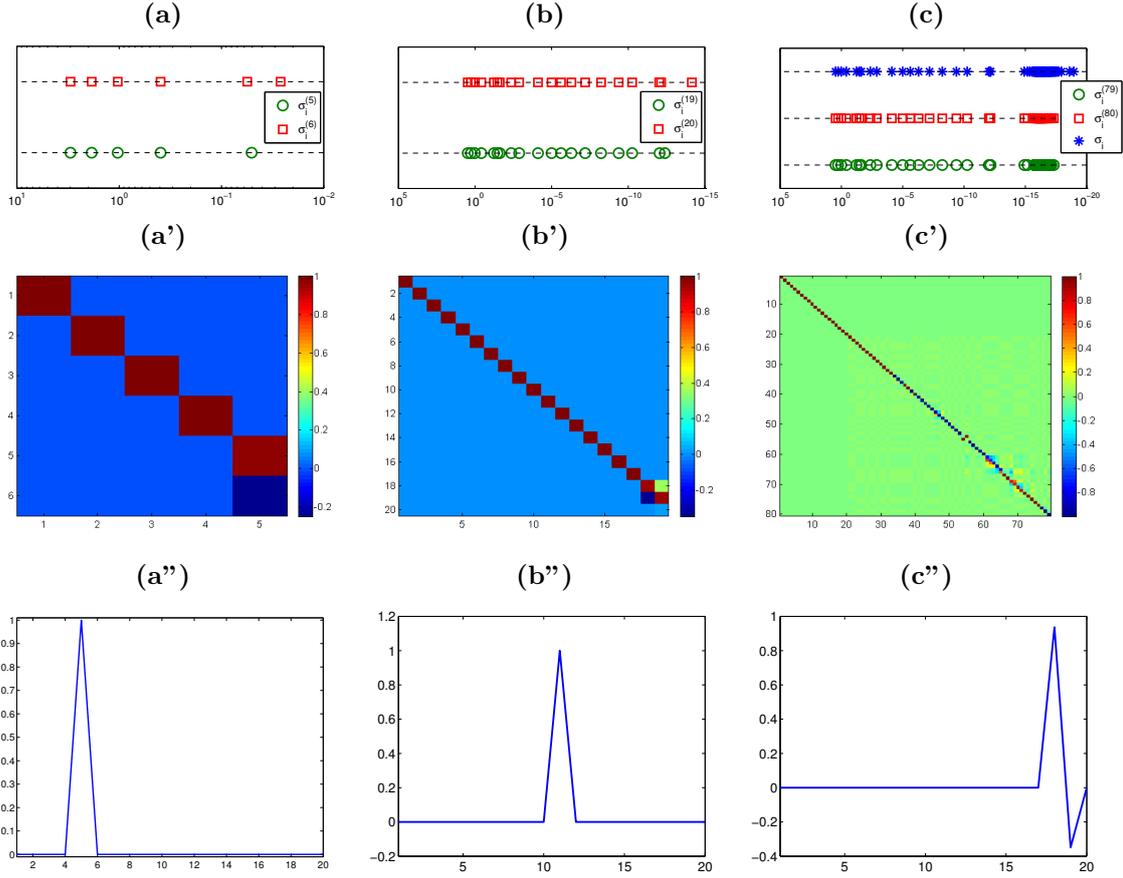


Figure 4: Test problem *shaw*. **(a)** singular values of  $\bar{H}_6$  and  $\bar{H}_5$ ; **(b)** singular values of  $\bar{H}_{20}$  and  $\bar{H}_{19}$ ; **(c)** singular values of  $\bar{H}_{80}$ ,  $\bar{H}_{79}$  and  $A$ ; **(a'')** matrix  $X_6$ ; **(b'')** matrix  $X_{20}$ ; **(c'')** matrix  $X_{80}$ ; **(a'')** column  $x_5^{(20)}$ ; **(b'')** column  $x_{11}^{(20)}$ ; **(c'')** column  $x_{18}^{(20)}$ .

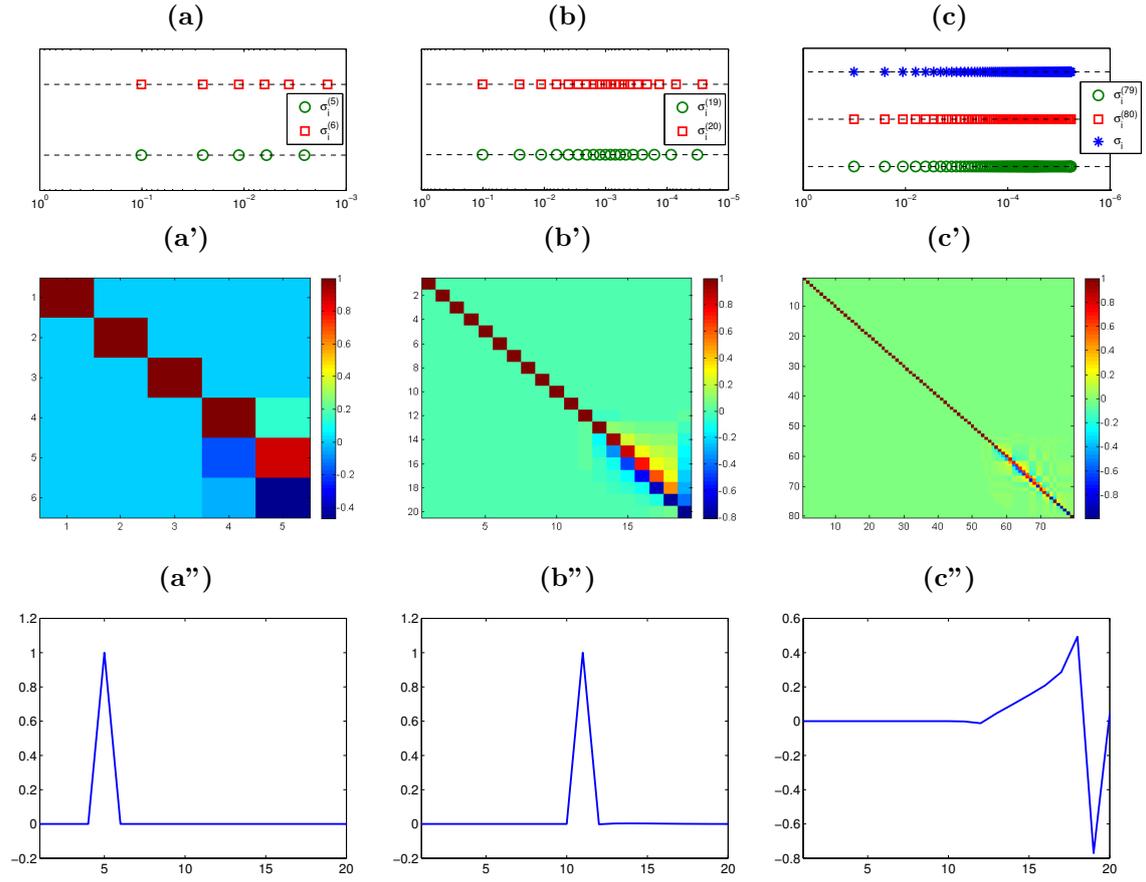


Figure 5: Test problem `deriv2`. **(a)** singular values of  $\bar{H}_6$  and  $\bar{H}_5$ ; **(b)** singular values of  $\bar{H}_{20}$  and  $\bar{H}_{19}$ ; **(c)** singular values of  $\bar{H}_{80}$ ,  $\bar{H}_{79}$  and  $A$ ; **(a')** matrix  $X_6$ ; **(b')** matrix  $X_{20}$ ; **(c')** matrix  $X_{80}$ ; **(a'')** column  $x_5^{(20)}$ ; **(b'')** column  $x_{11}^{(20)}$ ; **(c'')** column  $x_{18}^{(20)}$ .

(cf. again [14, Chapter 1]). Looking at the plots in the first row of Figure 5, we can still state that, typically,  $\sigma_i^{(k)} \simeq \sigma_i^{(k+1)}$ ,  $i = 1, \dots, k$ . Also in this case, assuming that  $\sigma_i^{(k+1)} \simeq (i\pi)^{-2}$ ,  $i = 1, \dots, k+1$ , we can consider the first-order approximation

$$\underbrace{(i\pi)^{-4}}_{(\sigma_i^{(k+1)})^2} - \underbrace{((i + \tilde{\varepsilon}_i)\pi)^{-4}}_{(\sigma_i^{(k)})^2} \simeq \frac{4}{\pi^4} i^{-5} \tilde{\varepsilon}_i, \quad \tilde{\varepsilon}_i > 0, \quad \tilde{\varepsilon}_i \simeq 0,$$

which is still analogous to (23).

In [6, Corollary 3], the authors give an estimate on the norm of the GMRES residual by assuming that the FOM inherits the DPC; more precisely, they prove that

$$\|r_k\| = O(k^{3/2}\sigma_k). \quad (27)$$

In this setting we can rigorously prove that the FOM inherits the DPC by employing essentially the same arguments as Theorem 1. More precisely, in the transition from the  $(k+1)$ th to the  $k$ th FOM iteration, we delete the  $(k+1)$ th column of  $H_{k+1}$  so to obtain  $\bar{H}_k$ , and we delete the  $(k+1)$ th row of  $\bar{H}_k$  so to obtain  $H_k$ . Just in this setting, we will denote the SVDs of the involved matrices in the following way:

$$H_{k+1} = U_{k+1}\Sigma_{k+1}V_{k+1}^T, \quad \bar{H}_k = \bar{U}_k\bar{\Sigma}_k\bar{V}_k^T, \quad H_k = U_k\Sigma_kV_k^T.$$

To estimate the SVD of  $H_k$ , one should twice apply an update formula that is analogous to (18), and that is still derived in [2]. In particular: recalling that  $\bar{H}_k = H_{k+1}[I_k, \mathbf{0}]^T$ , going from  $H_{k+1}$  to  $\bar{H}_k$  we get

$$\bar{U}_k = U_{k+1}\bar{X}_{k+1};$$

recalling that  $H_k = [I_k, \mathbf{0}]\bar{H}_k$ , going from  $\bar{H}_{k+1}$  to  $H_k$  we get

$$U_k = [I_k, \mathbf{0}]\bar{U}_k\bar{\Sigma}_kX_k\Sigma_k^{-1}.$$

At this point, the vector  $y_k$  in (9) is such that

$$\|y_k\| = \|\Sigma_k^{-1}U_k^T c_k\| = \left\| \Sigma_k^{-2}X_k^T\bar{\Sigma}_k\bar{U}_k^T \underbrace{\begin{bmatrix} I_k \\ \mathbf{0} \end{bmatrix}}_{c_{k+1}} c_k \right\| = \|\Sigma_k^{-2}X_k^T\bar{\Sigma}_k\bar{X}_{k+1}^T U_{k+1}^T c_{k+1}\|,$$

Let us assume that

$$[X_k^T \bar{z}]_j = [\bar{z}]_j \quad \forall \bar{z} \in \mathbb{R}^k, \quad \text{and} \quad [\bar{X}_{k+1}^T \bar{\bar{z}}]_j = [\bar{\bar{z}}]_j \quad \forall \bar{\bar{z}} \in \mathbb{R}^{k+1};$$

these assumptions are analogous to (22), and can be justified as in Remark 2. Since there exists  $\alpha > 1/2$  such that  $\left| [U_{k+1}^T c_{k+1}]_j \right| = O(j^{-\alpha}\sigma_j^{(k+1)})$ , we get

$$\left| [\Sigma_k^{-2}X_{k+1}^T\bar{\Sigma}_k\bar{X}_{k+1}^T U_{k+1}^T c_{k+1}]_j \right| = \frac{1}{(\sigma_j^{(k)})^2} \bar{\sigma}_j^{(k)} O(j^{-\alpha}\sigma_j^{(k+1)}).$$

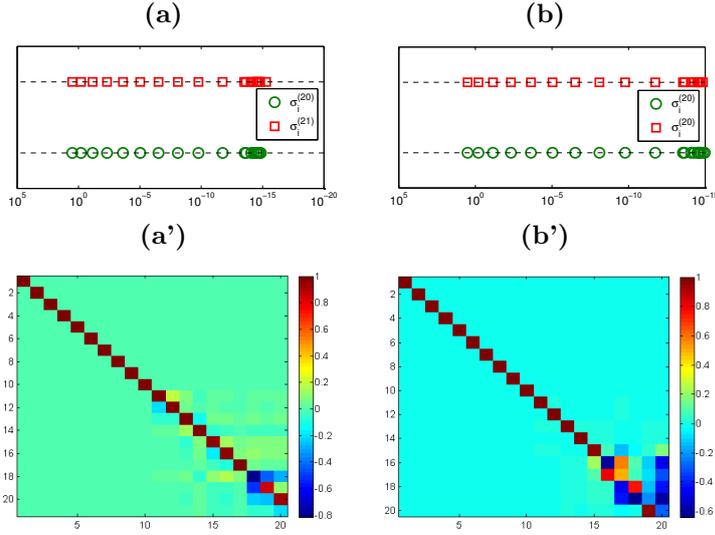


Figure 6: Test problem **baart**. (a) singular values of  $H_{21}$  and  $\bar{H}_{20}$ ; (b) singular values of  $\bar{H}_{20}$  and  $H_{20}$ ; (a') matrix  $\bar{X}_{21}$ ; (b') matrix  $X_{20}$ .

Thanks to the interlacing property of the singular values (which is twice applied: to  $H_{k+1}$  and  $\bar{H}_k$ , and to  $\bar{H}_k$  and  $H_k$ ), we can conclude that  $\|y_k\| < \infty$ . In Figure 6 we display the interlacing property of the singular values, and the matrices  $\bar{X}_{k+1}$  and  $X_k$  for  $k = 20$  and the test problem **baart**.

Alternatively, under stricter assumptions, in the following proposition we give a new estimate on the norm of the GMRES residual without relying on the behavior of the FOM solution.

**Proposition 3.** *Assume that (4) satisfies the DPC, and let  $r_k = b - Ax_k$  be the  $k$ th residual of GMRES applied to (4). Assume that the condition (24) holds. Then there exists  $\alpha > 1/2$  such that*

$$\|r_k\| = O\left((k+1)^{-\alpha} \sigma_{k+1}^{(k+1)}\right) + O\left(\left\|\delta_{k+1}^{(k+2)}\right\|\right) \quad (28)$$

where the components of  $\delta_{k+1}^{(k+2)}$  are analogous to (25), (26).

*Proof.* Thanks to relations (8) and (10),

$$\begin{aligned} r_k &= W_{k+1} (c_k - \bar{H}_k y_k) = W_{k+1} (I_{k+1} - \bar{H}_k \bar{H}_k^\dagger) c_k \\ &= W_{k+1} (I_{k+1} - U_k U_k^T) c_k, \end{aligned} \quad (29)$$

where  $U_k$  is the matrix appearing in (17). If we consider the complete SVD of  $\bar{H}_k$ , given by

$$\bar{H}_k = \hat{U}_k \hat{\Sigma}_k V_k^T, \quad \hat{U}_k \in \mathbb{R}^{(k+1) \times (k+1)}, \quad \hat{\Sigma}_{k+1} \in \mathbb{R}^{(k+1) \times k},$$

we have that  $\hat{U}_k = [U_k, u_{k+1}^{(k)}]$ . Therefore, equality (29) can be rewritten as  $r_k = W_{k+1} u_{k+1}^{(k)} (u_{k+1}^{(k)})^T c_k$ . Thanks to the orthonormality of the columns of  $W_{k+1}$

and  $u_{k+1}^{(k)}$ , we immediately get

$$\|r_k\| = \left| (u_{k+1}^{(k)})^T c_k \right|. \quad (30)$$

Let us consider the reduced and complete SVDs of the matrices  $\bar{H}_k^0$  (defined in (11)) and  $\bar{H}_{k+1}$ , given by

$$\bar{H}_k^0 = U_k^0 \Sigma_k V_k^T = \hat{U}_k^0 \hat{\Sigma}_k^0 V_k^T \quad \text{and} \quad \bar{H}_{k+1} = U_{k+1} \Sigma_{k+1} V_{k+1}^T = \hat{U}_{k+1} \hat{\Sigma}_{k+1} V_{k+1}^T,$$

respectively. We remark that, in the above equalities,

$$U_k^0 = \begin{bmatrix} U_k \\ \mathbf{0} \end{bmatrix}, \quad U_k^0 = \hat{U}_k^0 \begin{bmatrix} I_k \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad U_{k+1} = \hat{U}_{k+1} \begin{bmatrix} I_{k+1} \\ \mathbf{0} \end{bmatrix}. \quad (31)$$

Analogously to (18), let us consider  $X_{k+1} = U_{k+1}^T U_k^0 \in \mathbb{R}^{(k+1) \times k}$ , and  $\hat{X}_{k+1} := \hat{U}_{k+1}^T \hat{U}_k^0 \in \mathbb{R}^{(k+2) \times (k+2)}$ . We can immediately state that  $\hat{X}_{k+1}$  is orthogonal and, directly from (31),

$$X_{k+1} = [I_{k+1}, \mathbf{0}] \hat{X}_{k+1} \begin{bmatrix} I_k \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}.$$

At this point, we can express  $u_{k+1}^{(k)}$  in the following way:

$$u_{k+1}^{(k)} = [I_{k+1}, \mathbf{0}] \hat{U}_k^0 e_{k+1} = [I_{k+1}, \mathbf{0}] \hat{U}_{k+1} \hat{X}_{k+1} e_{k+1} = [I_{k+1}, \mathbf{0}] \hat{U}_{k+1} \hat{x}_{k+1}^{(k+1)}, \quad (32)$$

where  $\hat{x}_{k+1}^{(k+1)}$  is the  $(k+1)$ th column of  $\hat{X}_{k+1}$ . Thanks to (24) and the orthogonality of  $\hat{X}_{k+1}$ , we can write

$$\left| \hat{x}_i^{(k+1)} \right| = e_i - \delta_i^{(k+2)} \in \mathbb{R}^{k+2}, \quad \text{for } 1 \leq i \leq k+1, \quad (33)$$

where the components of  $\delta_i^{(k+2)}$  are defined as in (25) and (26). Directly from (30), thanks to relations (32) and (33), the Cauchy-Schwarz inequality, and Theorem 1, we can conclude that

$$\begin{aligned} \|r_k\| &= \left| (u_{k+1}^{(k)})^T c_k \right| = \left| (\hat{x}_{k+1}^{(k+1)})^T \hat{U}_{k+1}^T \underbrace{\begin{bmatrix} I_k \\ \mathbf{0} \end{bmatrix}}_{c_{k+1}} c_k \right| \\ &\leq \left| e_{k+1}^T \hat{U}_{k+1}^T c_{k+1} \right| + \left| (\delta_{k+1}^{(k+2)})^T \hat{U}_{k+1}^T c_{k+1} \right| \\ &\leq \left| (\hat{u}_{k+1}^{(k+1)})^T c_{k+1} \right| + \|\delta_{k+1}^{(k+2)}\| \|b\| \\ &= O\left((k+1)^{-\alpha} \sigma_{k+1}^{(k+1)}\right) + O\left(\|\delta_i^{(k+2)}\|\right), \quad \alpha > 1/2, \end{aligned}$$

where  $\hat{u}_{k+1}^{(k+1)}$  is the  $(k+1)$ th column of  $\hat{U}_{k+1}$ . □

We remark that the assumptions of Proposition 3 hold for all the considered test problems (see Figures 2, 4, and 5). Moreover, the new estimate (28) is more accurate than (27). In Figure 7 we consider the ideally exact systems (4), we perform 30 iterations of the Arnoldi algorithm, and we compare the values of  $\|r_k\|$  with the estimates (27) and (28); in the latter we take  $\alpha = 1/2$ , and  $\delta_{k+1}^{(k+2)} = \mathbf{0}$ .

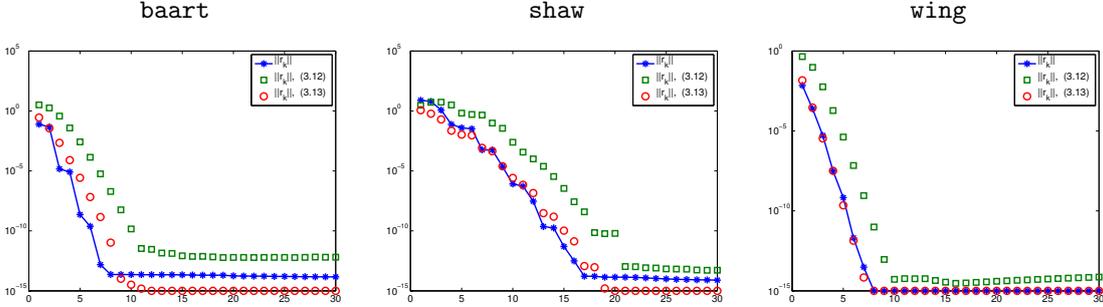


Figure 7: History of the GMRES residuals for three ideally exact test problems. The asterisks mark  $\|r_k\|$ , while the circles mark the estimates of Proposition 3.

When the right-hand side vector of problem (4) is perturbed, i.e., when  $b = b^{ex} + e$  ( $e$  Gaussian white noise), relation (28) is not valid anymore. The reason behind this is that, when  $\sigma_{k+1}^{(k+1)} < \|e\|$ ,  $(u_{k+1}^{(k+1)})^T c_k \simeq \|e\|$ : this typically happens when a few, say  $\bar{k}$ , iterations of the GMRES have been performed. Therefore, with derivations analogous to the ones of Proposition 3, in the perturbed case we can conclude that

$$\|r_k\| = O(\|e\|), \quad k \geq \bar{k}. \quad (34)$$

Estimate (34) agrees with the ones derived in [6]. In Figure 8 we illustrate the above estimate for some common test problems affected by additive noise  $e$  such that  $\|e\|/\|b^{ex}\| = 10^{-2}$  and  $\|e\|/\|b^{ex}\| = 10^{-3}$ ; we perform 60 iterations of the Arnoldi algorithm, which is simply implemented by standard Gram-Schmidt orthogonalization. As extensively explained in [6], (34) is particularly useful when a stopping criterion for problem (10) has to be set, or when a parameter selection strategy for problem (13) has to be considered.

### 3.2 Methods based on the Lanczos bidiagonalization algorithm

When analyzing the LSQR method, we can adopt a strategy similar to the one exploited to prove Theorem 1; namely, referring to (14), we can consider the relations between the SVD of the matrix  $\bar{B}_k^0$  given by

$$\bar{B}_k^0 = \begin{bmatrix} \bar{B}_k \\ \mathbf{0} \end{bmatrix} = U_k^0 \Sigma_k V_k^T \in \mathbb{R}^{(k+2) \times k},$$

and the SVD of the matrix  $\bar{B}_{k+1} = U_{k+1} \Sigma_{k+1} V_{k+1}^T$ . As in the GMRES case, we should assume that the columns of the matrix  $X_{k+1} = U_{k+1}^T U_k^0 \in \mathbb{R}^{(k+1) \times k}$  behave as described

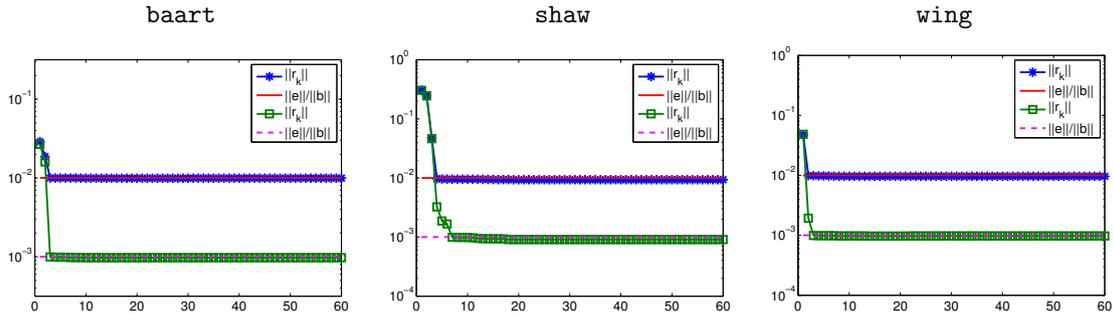


Figure 8: History of the GMRES relative residuals  $\|r_k\|/\|b\|$  for three test problems corrupted by different noises. The noise levels  $\|e\|/\|b^{ex}\|$  are  $10^{-2}$  and  $10^{-3}$ , and they are highlighted by two horizontal lines. The residuals relative to the more corrupted problems are marked by asterisks, the residuals relative to the less corrupted problems are marked by squares.

in (24). In Figure 9 we show some examples of the interlacing property of the singular values, and of the behavior of the matrix  $X_{k+1}$ , for two consecutive iterates of LSQR applied to the `baart` test problem.

However, in the LSQR case, we can prove the inheritance of the DPC in a much simpler way (see also [7]). Basically, we exploit the fact that the LSQR is mathematically equivalent to the CGLS method, and the property that the norms of the approximate solutions computed by CGLS are increasing (cf. [14, Chapter 6]).

**Theorem 4.** *If (4) satisfies the discrete Picard condition, then the projected least squares problems (14) satisfy the discrete Picard condition for  $k = 1, \dots, N$ .*

*Proof.* Since the solutions of (14) are such that  $\|x_k\| = \|y_k\| \leq \|y_{k+1}\| = \|x_{k+1}\|$ ,  $k = 1, \dots, N - 1$ , we can conclude that

$$\|y_1\| \leq \|y_2\| \leq \dots \leq \|y_N\| = \|x\| \leq C < \infty,$$

where  $x$  is the solution of (4). □

In Figure 10 we give some examples of Picard plots for the projected problems (14) associated to the `baart` test problem.

Concerning the behavior of the residual vector, we can again repeat the reasoning exploited in Proposition 3 and conclude that, if (4) satisfies the DPC, then the residual associated to the approximate solutions  $x_k$  in (14) are analogous to (28); if the right-hand side vector of (4) is corrupted, then  $\|r_k\| = O(\|e\|)$ . The latter estimate agrees with the one given in [16], where the authors essentially propose to recover  $\|e\|$  by evaluating the norm of the LSQR residual. Some examples about the behavior of the LSQR residuals are provided in Figure 11, where we display the residuals relative to both the exact and the corrupted problems (the noise level  $\|e\|/\|b^{ex}\|$  is  $10^{-2}$ ), and we perform 20 iterations of the Lanczos bidiagonalization algorithm. As in the Arnoldi case, these estimates can be employed when a stopping criterion for problem (14) has

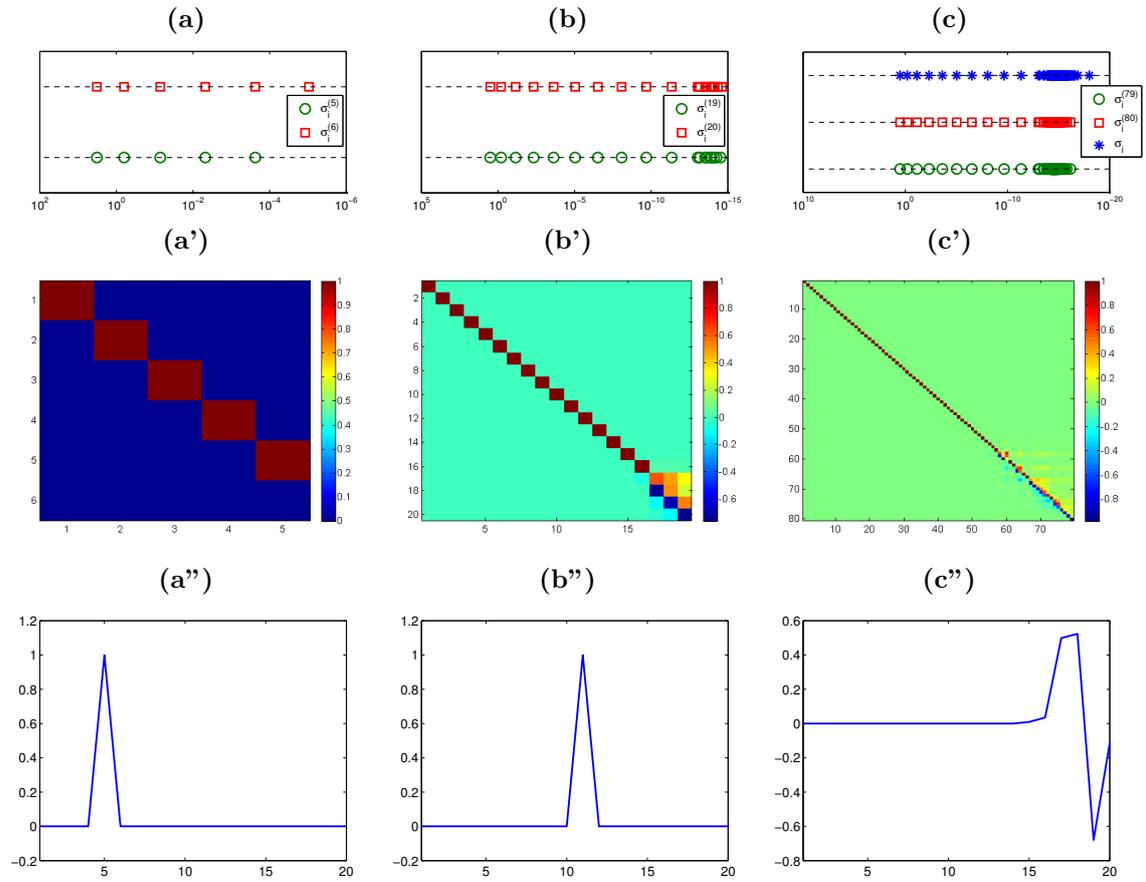


Figure 9: Test problem **baart**. (a) singular values of  $\bar{B}_6$  and  $\bar{B}_5$ ; (b) singular values of  $\bar{B}_{20}$  and  $\bar{B}_{19}$ ; (c) singular values of  $\bar{B}_{80}$ ,  $\bar{B}_{79}$  and  $A$ ; (a'') matrix  $X_6$ ; (b'') matrix  $X_{20}$ ; (c'') matrix  $X_{80}$ ; (a'') column  $x_5^{(20)}$ ; (b'') column  $x_{11}^{(20)}$ ; (c'') column  $x_{18}^{(20)}$ .

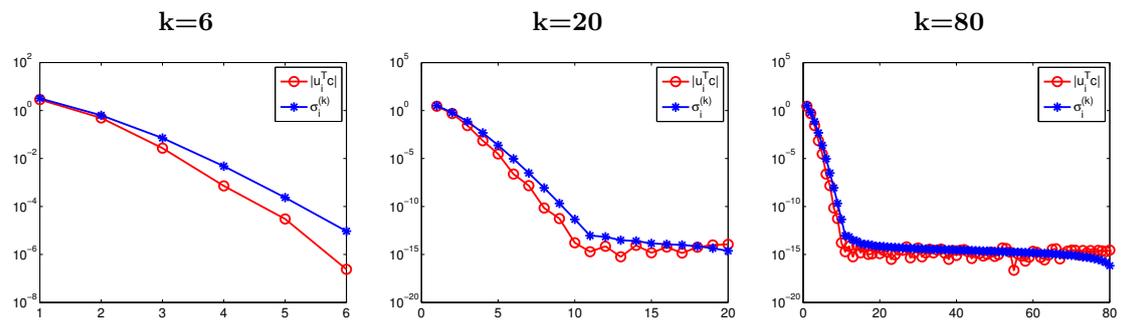


Figure 10: Test problem **baart**: Picard plots for the  $k$ th projected problems computed by the Lanczos bidiagonalization algorithm.

to be set, or when a parameter selection strategy for problem (15) has to be considered.

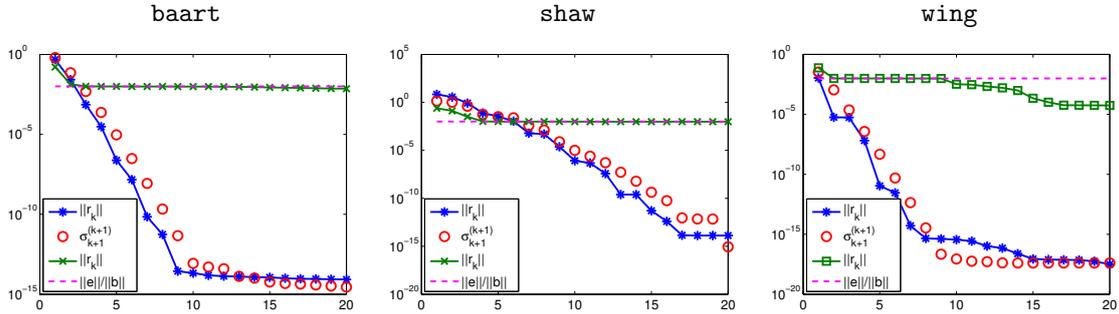


Figure 11: History of the LSQR residuals. The residuals of the ideally exact problems are marked by asterisks, the last singular value of the matrix  $\bar{B}_{k+1}$ ,  $k = 1, \dots, 20$  is marked by a circle, the relative residuals of the corrupted problems are marked by squares, the noise level is highlighted by an horizontal line.

In the LSQR case, some theoretical estimates have been derived to prove that the biggest singular values of  $A$  are well approximated by the biggest singular values of  $\bar{B}_k$  during the early iterations of the Lanczos bidiagonalization algorithm (cf. [1] and the references therein): thanks to this fact, we can formally justify the semi-convergence phenomenon. Indeed, when  $k$  is small, i.e., when  $\sigma_i^{(k)} > \|e\|$  for all  $i = 1, \dots, k$ , the Fourier coefficients are such that  $|(u_i^{(k)})^T c_k| \simeq i^{-\alpha} \sigma_i^{(k)}$ ,  $\alpha > 1/2$ , and the norm of the vector  $y_k$  is finite; in other words, as far as  $k$  is small, a sort of DPC holds. However, if at the  $\bar{k}$ th iteration  $\sigma_i^{(\bar{k})} \leq \|e\|$  for some  $i = 1, \dots, \bar{k}$ , then the corresponding Fourier coefficients are such that  $|(u_i^{(\bar{k})})^T c_{\bar{k}}| = O(\|e\|)$ . When computing  $\|y_{\bar{k}}\|$ , i.e., when considering the sum of the components  $((u_i^{(\bar{k})})^T c_{\bar{k}})^2 / (\sigma_i^{(\bar{k})})^2$ ,  $i = 1, \dots, \bar{k}$ , the contribution of the Fourier coefficients laying around  $\|e\|$  dominates, and the projected solution begins to diverge from the exact one. Furthermore, we can qualitatively explain why LSQR is a regularization method (in the classical sense recalled in the Introduction). Indeed, LSQR can deliver meaningful solutions whenever all the singular values of the projected matrices lie above the quantity  $\|e\|$ : as  $\|e\|$  approaches zero, an increasing number of iterations can be performed, and the corresponding projected solutions are limited because a sort of DPC is satisfied above  $\|e\|$ . Eventually, when  $\|e\|$  is numerically zero,  $N$  iterations of the LSQR can be performed, and the computed solution coincides with the exact one. The same derivations cannot be applied to the GMRES case, since no estimate about the approximation of the singular values has been proved, yet. From many numerical experiments (cf. for instance [7]), we learn that the biggest singular values of  $A$  are well approximated by the biggest singular values of  $\bar{H}_k$  during the first iterations of the Arnoldi algorithm. If we assume this behavior of the singular values, then we can extend the above considerations to the Arnoldi case.

## 4 Conclusions

In this paper we proved that the DPC is inherited when the GMRES and the LSQR methods are employed to solve linear problems satisfying the DPC: to do this we exploited some general SVD update formulas. For this reason we believe that the same reasoning can be applied to other Krylov subspace methods based on the Arnoldi and Lanczos bidiagonalization algorithms (including their range-restricted variants), and can be extended to other generic projection methods (for instance, the ones based on the nonsymmetric Lanczos algorithm, [22, Chapter 7]). Starting from the inheritance of the DPC, other properties of the GMRES and LSQR methods have been recovered: more precisely, we revisited some estimates on the behavior of the residuals, and we gave a further justification of the semiconvergence phenomenon. We believe that, thanks to the inheritance of the DPC, similar properties can be proved for a wider class of projection methods.

## References

- [1] S. Berisha and J. G. Nagy. Iterative image restoration. In R. Chellappa and S. Theodoridis, editors, *Academic Press Library in Signal Processing*, volume 4, chapter 7, pages 193–243. Elsevier, 2014.
- [2] J. R. Bunch and C. P. Nielsen. Updating the Singular Value Decomposition. *Numer. Math.*, 31:111–129, 1978.
- [3] D. Calvetti, B. Lewis, and L. Reichel. On the regularizing properties of the GMRES method. *Numer. Math.*, 91:605–625, 2002.
- [4] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari. Tikhonov regularization and the L-curve for large discrete ill-posed problems. *J. Comput. Appl. Math.*, 123:423–446, 2000.
- [5] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dordrecht, 2000.
- [6] S. Gazzola, P. Novati, and M. R. Russo. Embedded techniques for choosing the parameter in Tikhonov regularization. *Numer. Linear Algebra Appl.*, 21(6):796–812, 2014.
- [7] S. Gazzola, P. Novati, and M. R. Russo. On Krylov projection methods and Tikhonov regularization. *Electron. Trans. Numer. Anal.*, 2015. To Appear.
- [8] G. H. Golub and W. Kahan. Calculating the singular values and pseudo-inverse of a matrix. *SIAM J. Numer. Anal.*, 2:205–224, 1965.
- [9] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, third edition, 1996.
- [10] M. Hanke. *Conjugate Gradient Type Methods for Ill-Posed Problems*. Longman, Essex, UK, 1995.

- [11] M. Hanke and P. C. Hansen. Regularization methods for large-scale problems. *Surv. Math. Ind.*, 3:253–315, 1993.
- [12] P. C. Hansen. Computation of the singular value expansion. *Computing*, 40(3):185–199, 1988.
- [13] P. C. Hansen. The discrete Picard condition for discrete ill-posed problems. *BIT*, 30:658–672, 1990.
- [14] P. C. Hansen. *Rank-deficient and discrete ill-posed problems*. SIAM, Philadelphia, PA, 1998.
- [15] P. C. Hansen. Regularization Tools: A Matlab package for analysis and solution of discrete ill-posed problems, 2008.  
<http://www2.imm.dtu.dk/~pcha/Regutools/RTv4manual.pdf>.
- [16] I. Hnetynkova, M. Plesinger, and Z. Strakos. The regularizing effect of the Golub-Kahan iterative bidiagonalization and revealing the noise level in the data. *BIT*, 49:669–696, 2009.
- [17] M. E. Kilmer and D.P. O’Leary. Choosing regularization parameters in iterative methods for ill-posed problems. *SIAM J. Matrix Anal. Appl.*, 22(4):1204–1221, 2001.
- [18] P. Novati and M. R. Russo. A GCV based Arnoldi-Tikhonov regularization method. *BIT*, 54(2):501–521, 2014.
- [19] D. P. O’Leary and J. A. Simmons. A bidiagonalization-regularization procedure for large scale discretizations of ill-posed problems. *SIAM J. Sci. Stat. Comp.*, 2(4):474–489, 1981.
- [20] C. C. Paige and M. A. Saunders. Algorithm 583 LSQR: sparse linear equations and sparse least squares. *ACM Trans. Math. Software*, 8:195–209, 1982.
- [21] C. C. Paige and M. A. Saunders. LSQR: an algorithm for sparse linear equations and and sparse least squares. *ACM Trans. Math. Software*, 8:43–71, 1982.
- [22] Y. Saad. *Iterative Methods for Sparse Linear Systems, 2nd Ed.* SIAM, Philadelphia, PA, 2003.