An efficient truncated SVD of large matrices based on the low-rank approximation for inverse geophysical problems
Sergey Solovyev, Sébastien Tordeux

To cite this version:
Sergey Solovyev, Sébastien Tordeux. An efficient truncated SVD of large matrices based on the low-rank approximation for inverse geophysical problems. Sibirskie Elektronnye Matematicheskie Izvestiya, Russian Academy of Sciences, 2015, 12, pp.592-609. 10.17377/semi.2015.12.048. hal-01414769

HAL Id: hal-01414769
https://hal.archives-ouvertes.fr/hal-01414769
Submitted on 12 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DOI: http://dx.doi.org/10.17377/semi.2015.12.048
AN EFFICIENT TRUNCATED SVD OF LARGE MATRICES
BASED ON THE LOW-RANK APPROXIMATION FOR INVERSE
GEOPHYSICAL PROBLEMS

S.A. SOLOVYEV AND S. TORDEUX

Abstract. In this paper, we propose a new algorithm to compute a truncated singular value decomposition (T-SVD) of the Born matrix based on a low-rank arithmetic. This algorithm is tested in the context of acoustic media. Theoretical background to the low-rank SVD method is presented: the Born matrix of the acoustic problem can be approximated by a low-rank approximation derived thanks to a kernel independent multipole expansion. The new algorithm to compute T-SVD approximation consists of four steps, and they are described in detail. The largest singular values and their left and right singular vectors can be approximated numerically without performing any operation with the full matrix. The low-rank approximation is computed due to a dynamic panel strategy of cross approximation (CA) technique.

At the end of the paper, we present a numerical experiment to illustrate the efficiency and precision of the algorithm proposed.

Keywords: Born matrix, SVD algorithm, cross approximation (CA), low-rank approximation, high-performance computing, parallel computations.

1. Introduction

Computing the exact SVD or the T-SVD in full-arithmetic is a pretty expensive task. A standard way to compute the SVD consists of the two steps: first, the matrix is reduced to a bidiagonal form, then the SVD of the bidiagonal matrix is
computed. Usually, the first step is done by the Householder reflections and costs \((4/3)n^2(3m-n)\) arithmetic operations (FLOPS). The second step can be performed by an iterative refinement algorithm with stopping criterion equal to the machine precision, it takes about \(O(n)\) iterations, each costs \(O(n)\) FLOPS [3]. Another version of the second step is a QR algorithm for the computation of eigenvalues and takes \(O(n^2)\) FLOPS \([4,10]\). There are various modifications of them, using a divide-and-conquer method, preconditioned and Jacobi plane rotation methods. These ones are implemented in LAPACK routines. To compare with own developed algorithm we use the LAPACK functionality from Intel MKL. However, it is almost impossible to compute such a decomposition for large-scale problems using a robust arithmetic.

The objective of this paper is designing an efficient algorithm to compute an approximation of the T-SVD based on a low-rank arithmetic.

More precisely, in this paper we aim at looking for matrices \(\overline{U}_k = \{\overline{u}_1, \ldots, \overline{u}_k\} \in \mathbb{C}^{l \times k}, \overline{V}_k = \{\overline{v}_1, \ldots, \overline{v}_k\} \in \mathbb{C}^{l \times k}\) and \(\overline{D}_k = diag(\overline{d}_k)_{k \times k} \in \mathbb{R}^{k \times k}\) which approximate \(U_k, V_k, D_k\) of the T-SVD (28) in the following sense:

I. The difference between approximate singular values and exact singular values

\[d_i - \tilde{d}_i < \eta_1, \quad 1 < i \leq k \quad \text{and} \quad d_i - \delta_i < \eta_2, \quad k < i < n \]

II. The angles between the approximated and exact left and right singular spaces are smaller than a small parameter \(\eta_2\)

\[\angle(U_k, \overline{U}_k) < \eta_2 \quad \text{and} \quad \angle(V_k, \overline{V}_k) < \eta_2\]

where \(\angle(A, B) = \arccos(\sigma)\) with the smallest singular value \(\sigma\) of \(A^*B\).

The major application of the algorithm proposed is the geophysical inverse problem. It consists in determining the physical characteristics of a propagation medium by interpreting the measured data by receivers for different sources. One of the main features of this problem is the huge number of receivers, sources and parameters of the model which is under study. Among all different techniques that exist, the Full Waveform Inversion (FWI) is one of the most costly. It consists in an iterative Newton-type procedure which requires the computation of the so-called Born Matrix associated with the fullwave equation.

It is known that because of its poor conditioning this very large scale problem is difficult to solve: some (combinations of) the sources, receivers or model parameters are very important to take into account whereas some other (combinations) do not have so strong impact. The numerical method proposed in this paper needs computing the Singular Value Decomposition \([14,7]\) of the Born matrix to reduce the complexity of this problem by identifying the most important model parameters, sources and receivers.

This paper will be composed as follows: In Section 2, we briefly recall what is a low-rank approximation of a matrix and some algorithm to compute these approximations. Section 3 is the "core" of the paper. We describe an algorithm to compute a numerical approximation of the Truncated Singular Value decomposition based on a low-rank arithmetic. In section 4, some numerical results are presented. These results illustrate the efficiency and accuracy of the method. There are three appendices: in Appendix A, we describe in detail the Singular Value Decomposition and the Truncated Singular Value Decomposition. Appendix B is concerned with a theoretical result about the Born acoustic matrix. We recall that the Born matrix
associated with homogeneous acoustic media can be approximated by a low-rank approximation under suitable assumptions. Finally, Appendix C briefly recalls classical results about the Chebyshev tensorial interpolation.

2. The low rank approximation of a matrix

In what follows, we will make an intensive use of the low-rank approximation.

2.1. Definition of the $\varepsilon$-rank of a matrix and low-rank approximation. Let $A \in \mathbb{C}^{I \times J}$ be a matrix with $I$ rows and $J$ columns with $J \leq I$. The matrix $A$ has $\varepsilon$-rank $k$ if $k$ is the smallest integer such that there exists a matrix $A_k \in \mathbb{C}^{I \times J}$ with the rank $k$ satisfying

$$\frac{\|A - A_k\|}{\|A\|} < \varepsilon.$$  

(3)

When $\|\cdot\|$ is the euclidean matrix norm $\|\cdot\|_2$, the $\varepsilon$-rank of a matrix $A$ is explicitly given by the number of singular values which are larger than $\varepsilon$. Moreover, the matrix $A_k$ can be deduced from the T-SVD of the matrix $A$

$$A_k = U_k D_k V_k^*$$

(4)

with $U_k \in \mathbb{C}^{I \times k}$, $D_k \in \mathbb{R}^{k \times k}$, and $V_k \in \mathbb{C}^{J \times k}$.

In reality, the computation of the T-SVD of a large-size matrix is very expensive, many authors have proposed other algorithms to compute non-optimal (in the sense of the matrix euclidean norm) low-rank approximation of a matrix $A$. These methods consist in looking for a matrix $A_k$ as a product of the two matrices $B_k \in \mathbb{C}^{I \times k}$ and $C_k \in \mathbb{C}^{J \times k}$, which minimize $k$, satisfying

$$\frac{\|A - B_k C_k^T\|}{\|A\|} < \varepsilon.$$  

(5)

The norm $\|A\|_2$ is given by the largest singular value of the matrix $A$ and its computation is pretty expensive. So, we prefer to use $\|A\| = \max_{1 \leq i, j \leq k} |A_{i,j}|$ norm.

The two common techniques to obtain this factorization are the QR factorization with pivoting of the matrix and the Cross Approximation (CA) technique which is similar to the incomplete LU factorization with pivoting. To determine $\varepsilon$-rank of a matrix and obtain low-rank approximation, the rank-revealing QR modification with pivoting is used (RRQR-piv) [8]. The approximate number of floating-point operations for real values is $(2/3)n^2(3m - n), m \geq n$. The RRQR-piv algorithm is slow (however faster than the computation of the T-SVD of a matrix $A$) algorithm and is almost optimal in terms of the rank $k$, whereas the CA algorithm is rapid, but can give rise to non-optimal $k$.

Let us briefly describe the RRQR-piv algorithm and describe in detail various modifications of the CA approach.


- Initialization:

$$R_0 = A \in \mathbb{C}^{I \times J}, \quad n = 0$$

(6)

- While stopping-criteria ($R_n$) > $\varepsilon \|A\|

Step 1. Choose a pivot $(i_*, j_*)$ in $R_n$
Step 2. Define two vector columns $b_{n+1} \in \mathbb{C}^I$ and $c_{n+1} \in \mathbb{C}^J$

$$
\begin{align*}
(b_{n+1})_i &= (R_n)_{i,\cdot}, & (c_{n+1})_j &= \frac{(R_n)_{\cdot,j}}{(R_n)_{i,j}}.
\end{align*}
$$

Step 3. Increment $n$: $n = n+1$. Define the matrices $B_n \in \mathbb{C}^{I \times n}$ and $C_n \in \mathbb{C}^{J \times n}$ and

$$
B_n = [b_1, b_2, \ldots, b_n] \quad \text{and} \quad C_n = [c_1, c_2, \ldots, c_n].
$$

Step 3. Update

$$
R_n = A - B_n C_n^T.
$$

As a result, we obtain both matrices $B_k$ and $C_k$ that are involved in the low-rank approximation of $A$.

Like for the incomplete LU factorization algorithm, we have the following options for this algorithm:

I \ Total pivoting: The stopping criterion is the following:

$$
\text{stopping-criterion}(R_n) = \max_{i \leq I, j \leq J} |(R_n)_{i,j}|
$$

At each iteration, the pivot is chosen by maximizing $|(R_n)_{i,j}|$ over whole the matrix

$$
|(R_n)_{i,j}| = \max_{i \leq I, j \leq J} |(R_n)_{i,j}|
$$

We then define a panel $J_\square \subset [1, J]$ of width $2K + 1 \in [1, J]$:

$$
J_\square = \begin{cases} 
[1, 2K + 1] & \text{if } j_\square \leq K, \\
[j_\square - K, j_\square + K] & \text{if } K < j_\square \leq J - K, \\
[J - 2K, J] & \text{if } j_\square > J - K.
\end{cases}
$$

As long as a maximum of $|R_n|$ over this panel is larger than $\varepsilon$

$$
\max_{i \leq I, j \in J_\square} |(R_n)_{i,j}| > \varepsilon,
$$

the pivot $(i_*, j_*)$ will be chosen into this panel of columns

$$
|(R_n)_{i_*,j_*}| = \max_{i \leq I, j \in J_\square} |(R_n)_{i,j}|.
$$

When (14) is not fulfilled anymore, another panel is considered in the same way until

$$
\max_{i \leq I, j \leq J} |(R_n)_{i,j}| < \varepsilon.
$$

II \ Dynamic panel strategy: The algorithm is more complex to describe. This corresponds to a partial pivoting. First, as for the total pivoting we define $i_\square$ and $j_\square$ such that

$$
|(R_n)_{i_\square,j_\square}| = \max_{i \leq I, j \leq J} |(R_n)_{i,j}|
$$

We then define a panel $J_\square \subset [1, J]$ of width $2K + 1 \in [1, J]$:

$$
J_\square = \begin{cases} 
[1, 2K + 1] & \text{if } j_\square \leq K, \\
[j_\square - K, j_\square + K] & \text{if } K < j_\square \leq J - K, \\
[J - 2K, J] & \text{if } j_\square > J - K.
\end{cases}
$$

As long as a maximum of $|R_n|$ over this panel is larger than $\varepsilon$

$$
\max_{i \leq I, j \in J_\square} |(R_n)_{i,j}| > \varepsilon,
$$

the pivot $(i_*, j_*)$ will be chosen into this panel of columns

$$
|(R_n)_{i_*,j_*}| = \max_{i \leq I, j \in J_\square} |(R_n)_{i,j}|.
$$

When (14) is not fulfilled anymore, another panel is considered in the same way until

$$
\max_{i \leq I, j \leq J} |(R_n)_{i,j}| < \varepsilon.
$$

III \ Cross pivoting: The pivot is chosen in the following way: Pick by hazard a non-zero column $(R_n)_{i,j_\Delta}$ of $R_n$, with $j_\Delta \in J$. Define the integer $i_* \in I$ by looking for a maximum of $|R_n|$ in this column

$$
|(R_n)_{i_*,j_\Delta}| = \max_{i \in I} |(R_n)_{i,j_\Delta}|.$$
Define the integer \( j_\ast \in J \) by looking for a maximum of \(|R_n|\) in this row
\[
|(R_n)_{i,j}| = \max_{j \in J} |(R_n)_{i,j}|
\]

The stopping criterion is then the following:
\[
|(R_n)_{i,j}| > \varepsilon
\]

**Remark 1.** The search for a maximum and the update of the matrix \( R_n \) is the performance of the bottle-neck of the CA algorithm. The total pivoting strategy is much slower than the two other strategies because of
- the search for the maximum is made over the full matrix.
- at each iteration the matrix should be fully updated

The dynamical panel strategy is more efficient since
- the search for the maximum is made over a small subset of the full matrix.
- only the panel of the matrix needs to be updated at each iteration

Since the panel is formed of a group of columns, it is important to optimize the access to memory in storing the matrix in the RAM column-by-column. If a matrix is stored row-by-row, the panel should be constructed of a group of rows.

The cross partial pivoting strategy is also very efficient since
- the search for the maximum is made over a cross which is a small subset of the full matrix.
- the matrix \( R_n \) does not need not to be updated but only needs to be evaluated for a small number of indices at each iteration.

### 3. DESCRIPTION OF THE ALGORITHM

The algorithm can be decomposed into four steps.

The first step consists in decomposing vertically the matrix \( A \) into blocks \( A_i \),

\[
A = \begin{bmatrix}
A_1 \\
\vdots \\
A_6
\end{bmatrix}
\]

**Fig. 1.** Decomposition of the matrix \( A \) by blocks
(Figure 1), and in performing a low-rank approximation of each block $A_i \in C^{m_i \times J}$ (Figure 2),

$$A_i \simeq B_i C_i^T$$ with $B_i \in C^{m_i \times k_i}$ and $C_i \in C^{J \times k_i}$.

To compute a low-rank approximation of blocks we have the three options:

i) T-SVD in the full arithmetic;

ii) RRQR-piv algorithm;

iii) Cross Approximation (CA) technique.

**Remark 2.** The algorithm will only be efficient if the integer $k_i$ is less than $J$. In practice, this number is small.

**Remark 3.** The accuracy of the low-rank approximation of the matrix $A$ is characterized by a small parameter $\varepsilon$ and by the stopping criterion. For SVD-compression, QR-piv and CA algorithm, it takes the form

$$\begin{align*}
\|A_i - B_i C_i^T\|_2 & \leq \varepsilon \text{ for T-SVD and RRQR algorithm,}
\|A_i - B_i C_i^T\|_\infty & \leq \varepsilon \text{ for CA algorithm},
\end{align*}$$

with $\| \cdot \|_2$ the euclidean matrix norm and $\|A\|_\infty = \max_{i \leq I, j \leq J} |A_{i,j}|$.

**Fig. 2.** Low-rank approximation of $A$

The result of the first step is depicted in Figure 3. In this picture and in the next ones, the "plotted" parts of matrices mean dense non-zero blocks. The "white" blocks mean zero fill-in.

At the second step, we orthogonalize the matrices $B$ and $C$. More precisely, we perform a QR decomposition of the matrices $B_i$ and $C$

$$B_i = \tilde{B}_i R_i \text{ and } C^T = L \tilde{C}^T,$$

with $\tilde{B}_i \in R^{m_i \times k_i}$, $\tilde{C} \in R^{J \times k}$ being orthogonal, $R_i \in R^{k_i \times k_i}$ — the upper triangular and $L \in R^{k \times k}$ — the lower triangular, where $k = \sum k_i$. The matrices $\tilde{B}_i$ and $R_i$ are collected into the orthogonal matrix $\tilde{B}$ and into the upper triangular matrix $R$, (see Figure 4). This result in that the Low-rank approximation of $A$ should be

$$A \simeq \tilde{B} (R L) \tilde{C}^T,$$

with $\tilde{B} \in R^{I \times k}$, $\tilde{C} \in R^{J \times k}$ and $RL \in R^{k \times k}$ being full.

At the third step, a robust T-SVD with the accuracy $\delta$ of the product $RL$ is performed

$$RL = U_{RL} D_{RL} V_{RL}^*$$ with $\|U_{RL} D_{RL} V_{RL}^* - RL\| < \delta$.

The result of the third step is presented in Figure 5.

**Remark 4.** When the matrix $RL$ is much smaller than the initial matrix $A$ (this is a wide-spread in the practice case), the computation of the T-SVD in the full arithmetic of the product $RL$ is less expensive than the computation of the T-SVD of the matrix $A$. 
At the **fourth step**, we construct the final matrices by computing the products 
\( \bar{U} = B U_{RL} \), \( \bar{V} = V_{RL} \tilde{C} \) and \( \bar{D} = D_{RL} \). As a result, the matrices \( \bar{U} \) and \( \bar{V} \) have orthogonal columns.

Our statement is that the decomposition \( \bar{U} \bar{D} \bar{V} \) approximates the exact T-SVD of \( A \) in the sense of (1) and (2).
4. NUMERICAL EXPERIMENTS

The numerical experiments are aimed at demonstrating the performance of our algorithm in terms of computation time and precision.

We have not developed our own linear algebra library but have used the LAPACK and BLAS functions of Intel MKL. The performance was measured on Intel Core i7-3770K CPU 3.5 GHz, (Ivy Bridge). We have avoided the impact of OMP parallelization of all MKL functions by switching off threading by setting OMP_NUM_THREADS=1.

The Born matrix $A$ associated with a 2D elasticity vertically inhomogeneous (layered) isotropic medium, one source, 1450 receivers and 10 different frequencies, has been considered. The target domain, containing 120×20 points, is a part of a huge real model which we would like to image. Details of this model are described in [13]. The full matrix $A$ has 29,000 lines and 7,200 columns. At the preliminary step, the matrix $A$ is separated into $p = 10$ blocks.

4.1. Computational time. In the first test, the accuracy $\varepsilon$ of the low-rank approximation is $10^{-6}$ and the threshold $\delta$ of the cropped exact SVD $U_k D_k V_k^T$ and of the cropped low-rank SVD $\tilde{U} \tilde{D} \tilde{V}$ is $10^{-6}$. Different options for the first step are tested: the computational time for the SVD, the QR and the CA compression methods are computed. These computational times should be compared to 970 s, i.e. the time of a robust SVD by the gesvd computational optimized routine of Intel MKL.

The performance results show the advantage of using partial pivoting in the CA over the cross pivoting and classical total pivoting. Additionally, the performance results confirm the high acceleration of the CA approach in comparison with the SVD and the QR-piv techniques.
Finally, we point out that the first step of the method proposed is easily parallelizable: since the compression of all blocks can be done by different processors as opposed to the last three steps.

4.2. Precision. In the second test our attention will be focused on two different error indicators. Only the results for the most efficient case (Cross Approximation with dynamic panel partial pivoting) will be presented.

First, we are interested in the error resulting from approximate singular values. Denoting by \( I_k \) the \( \varepsilon \)-rank of the matrix \( A \), the following quantities

\[
\begin{align*}
\max_{i \leq I_k} \frac{|d_i - \tilde{d}_i|}{|d_i|} & \quad \text{(absolute error)} \\
\max_{i \leq I_k} \frac{|d_i - \tilde{d}_i|}{|d_i|} & \quad \text{(relative error)}
\end{align*}
\]

(23)

have been plotted with respect to \( \varepsilon \), the parameter relating to the accuracy of the Cross Approximation parameter and with respect to \( \delta \), i.e. the final threshold parameter. The results are reported in Figure 6 for the absolute error and in Figure 7 — for the relative error. It seems that the absolute error does mostly depend on \( \varepsilon \). On contrary, the relative error is sufficiently related to the ratio \( \varepsilon/\delta \). Second, we quantify the quality of the approximation of the left singular space. The results for the right singular values are quite similar and are not presented. We compute the angle (in degrees) between the subspace generated by the columns of \( U \) and the subspace generated by the columns of \( U_k \).

\[
\angle(U, U_k) = \arccos(\sigma) \frac{180}{\pi}
\]

(24)

with the smallest singular value \( \sigma \) of the matrix \( U^T U_k \).

Numerical measurements (Figure 8) show that the angle \( \alpha = \angle(U, U_k) \), for any threshold \( \delta \), can be decreased via improving the accuracy \( \varepsilon \) of the low-rank approximation. The numerical results reveal that the error depends mostly on \( \varepsilon/\delta \) like the relative error for the singular values (Figure 7).

4.3. Evolution of \( \varepsilon \)-rank. In the last test we investigate the ranks of truncated matrices. They are presented in Tables 2 and 3 for different low-rank approximations. The SVD approach is the most optimal in terms of \( \varepsilon \)-rank, whereas the CA technique is the worst (the rank after the first step in the Tables). For \( \varepsilon = 10^{-6} \) and \( \delta = 10^{-6} \), the final ranks associated with different compressions are slightly different. For
\[ \varepsilon = 10^{-9} \text{ and } \delta = 10^{-6}, \text{ the final ranks are equal. This will be the case when the compression } \varepsilon \text{ is much smaller than } \delta. \text{ The final rank of a low-rank SVD should be compared to a rank obtained by truncated full arithmetic SVD, i.e. 1984. In the context of seismic imaging, a small misfit of the rank will not have, in our opinion, a large impact on the solution of the inverse problem.} \]

5. CONCLUSION

We have presented an algorithm to compute the truncated SVD of the Born matrix. This method is based on a low-rank arithmetic and the CA technique. To perform the low-rank approximation, we have proposed a dynamic panel CA algorithm. This approach is similar to the panel local pivoting $LU$ decomposition.
Fig. 8. The CA partial pivoting: Dependence of the angle between subspaces spanned on singular vectors of the cropped exact SVD and Low-rank SVD.

<table>
<thead>
<tr>
<th>Matrix rank \ Approaches</th>
<th>SVD</th>
<th>RRQR</th>
<th>Total pivot.</th>
<th>Cross pivot.</th>
<th>Partial pivot.</th>
</tr>
</thead>
<tbody>
<tr>
<td>After the 1-st step</td>
<td>2966</td>
<td>2424</td>
<td>2444</td>
<td>2454</td>
<td>2795</td>
</tr>
<tr>
<td>After the 2-nd step</td>
<td>2041</td>
<td>2041</td>
<td>2022</td>
<td>1961</td>
<td>2013</td>
</tr>
<tr>
<td>After the 3-th step</td>
<td>1963</td>
<td>1963</td>
<td>1953</td>
<td>1936</td>
<td>1950</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Matrix rank \ Approaches</th>
<th>SVD</th>
<th>RRQR</th>
<th>Total pivot.</th>
<th>Cross pivot.</th>
<th>Partial pivot.</th>
</tr>
</thead>
<tbody>
<tr>
<td>After the 1-st step</td>
<td>2963</td>
<td>3030</td>
<td>3052</td>
<td>3052</td>
<td>3565</td>
</tr>
<tr>
<td>After the 2-nd step</td>
<td>2619</td>
<td>2619</td>
<td>2586</td>
<td>2541</td>
<td>2581</td>
</tr>
<tr>
<td>After the 3-th step</td>
<td>1964</td>
<td>1964</td>
<td>1964</td>
<td>1964</td>
<td>1964</td>
</tr>
</tbody>
</table>

Table 2. The rank of matrices at intermediate steps, \( \varepsilon = \delta = 10^{-6} \)

Table 3. The rank of matrices at intermediate steps, \( \varepsilon = 10^{-9} \), \( \delta = 10^{-6} \)

The algorithm proposed is an alternative to a very popular and efficient randomized SVD approach proposed by Rokhlin [6]. The main advantages are: (i) the \( \varepsilon \)-rank of a matrix has not to be known in advance, (ii) the computation of a reduced matrix is less expensive (this has been confirmed by preliminary numerical tests which are not included in this paper).

For a representative configuration we have compared the results generated by the proposed truncated SVD algorithm to the results obtained by an exact SVD. We have observed that the method is accurate and the acceleration of the computation has increased by the factor 10 on one-thread systems. The algorithm has a good
Appendix A. The Singular Value Decomposition and the Truncated Singular Value Decomposition

The SVD of a matrix $A \in \mathbb{C}^{I \times J}$ with $J \leq I$ is a factorization of the form

$$A = UDV^*,$$

where the matrices $U$ and $V$ contain the left and right singular vectors $u_i$ and $v_i$; the matrix $D$ is diagonal and contains the singular values $d_i$

$$U = \{u_1, \cdots, u_J\} \in \mathbb{C}^{I \times J} \quad \text{with} \quad U^*U = I \in \mathbb{C}^{J \times J}$$

$$V = \{v_1, \cdots, v_J\} \in \mathbb{C}^{J \times J} \quad \text{with} \quad V^*V = I \in \mathbb{C}^{J \times J}$$

$$D = \text{diag}(d_i)_{i=1}^J \in \mathbb{R}^{J \times J}.$$

The singular values $d_i$ are all positive and ordered

$$d_1 \geq d_2 \geq \cdots \geq d_J \geq 0.$$  

The T-SVD of the matrix $A$ is obtained from the SVD by removing the singular values $d_{k+1}, d_{k+2}, \ldots$ which are lower than a small parameter $\delta$

$$A_k = U_k D_k V_k^*,$$

$$U_k = \{u_1, \cdots, u_k\} \in \mathbb{C}^{I \times k} \quad \text{with} \quad U_k^*U_k = I \in \mathbb{C}^{k \times k}$$

$$V_k = \{v_1, \cdots, v_k\} \in \mathbb{C}^{J \times k} \quad \text{with} \quad V_k^*V_k = I \in \mathbb{C}^{k \times k}$$

$$D_k = \text{diag}(d_i)_{i=1}^k \in \mathbb{C}^{k \times k}.$$

The matrix $A_k$ is an approximation of the matrix $A$ in the following sense

$$\|A - A_k\|_2 \leq \delta$$

with the Euclidean matrix norm $\| \cdot \|_2$:

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$$

Appendix B. Some theoretical results involving the acoustic Born matrix

Let us show on a simple example that the Born matrix of an acoustic problem can be approximated by a low-rank approximation derived thanks to a kernel independent multipole expansion.

The model parameters. The considered propagation domain consists of unbounded three-dimensional (3D) acoustic media governed by the Helmholtz equation with varying physical characteristics $\mu(y)$ (the square of the wave-number).

The model is parameterized on a regular grid with the spatial step $\delta_y \in \mathbb{R}_+$ composed of $J$ cells $K_j \subset \mathbb{R}^3$ (Figure 9),

$$K_j = [j_1 \delta_y, (j_1 + 1)\delta_y] \times [j_2 \delta_y, (j_2 + 1)\delta_y] \times [j_3 \delta_y, (j_3 + 1)\delta_y]$$

with the integer $j \in [1, J]$ related to the integers $j_1 \in [0, J_1 - 1]$, $j_2 \in [0, J_2 - 1]$ and $j_3 \in [0, J_3 - 1]$ by the relation

$$j = j_3 J_2 J_1 + j_2 J_1 + j_1 + 1 \quad \text{and} \quad J = J_1 J_2 J_3.$$
The function $\mu$ is chosen to be constant outside the regular grid and piecewise constant on the regular grid with the value $\mu_j \in \mathbb{R}_+$ on $K_j$

$$
\mu(y) = \mu_j \text{ if } y \in K_j \quad \text{and} \quad \mu(y) = \mu_0 \text{ else.}
$$

The model parameters $\mu_j$, $1 \leq j \leq J$, are collected into a vector $\mu \in \mathbb{R}^J$.

The data are obtained thanks to $I_1$ experiments, each corresponding to a source located at a point $x_{1i}^1 \in \mathbb{R}^3$, with $0 \leq i_1 \leq I_1 - 1$. For every experiment, $I_2$ measurements are realized by receivers located at a point $x_{2j}^2 \in \mathbb{R}^3$, with $0 \leq i_2 \leq I_2 - 1$. This gives rise to a set of data composed of $I = I_1 I_2$ measurements. To organize these data, every receiver–source couple is indexed by an integer $i \in [1, I]$ given by

$$
i = i_1 I_1 + i_2 + 1.
$$

The full wave inverse problem takes the form:

Given $f \in \mathcal{C}^I$, find $\mu \in \mathbb{R}^J_+$ such that

$$
F_i(\mu) = f_i \quad \text{for } 1 \leq i \leq I
$$

with $F_i(\mu) = u_{i_1}(\mu; x_{i_2})$ where $x \mapsto u_{i_1}(\mu; x)$ is defined over all $\mathbb{R}^3$ as the outgoing solution of the direct acoustic problem for the $i_{th}$ source:

$$
\Delta u_{i_1}(\mu; x) + \mu(\mathbf{x}) \ u_{i_1}(\mu; x) = -\delta_{x_{i_1}}
$$

with $\delta_{x_{i_1}}$ the Dirac delta function located at the point $x_{i_1}^1$.

Most of the algorithms that have been proposed in the literature are of the Newton type [9, 1]. They require the computation of the Born Matrix $A \in \mathbb{C}^{I \times J}$ which contains the partial derivative with respect to $\mu_j$ of the nonlinear form $F_i$.

$$
A_{i,j} = \frac{\partial F_i}{\partial \mu_j}(\mu) \quad \text{for } 1 \leq i \leq I \text{ and } 1 \leq j \leq J.
$$

The matrix $A$ can be expressed as

$$
A_{i,j} = u_{i_1}(x_{i_2}^2)
$$

Fig. 9. Discretization of the model parameter via a 3D grid.
with respect to the function $u_i^j$, which is the partial derivative of $u_i$ with respect to $\mu_j$

$$u_i^j(\mu; x) = \frac{\partial u_i}{\partial \mu_j}(\mu; x) = \lim_{h \to 0} \frac{u_i(\mu + he_j; x) - u_i(\mu; x)}{h}.$$  

Deriving (36) with respect to $\mu_j$, we obtain that the function $u_i^j$ is the unique outgoing solution of

$$\Delta u_i^j(\mu; x) + \mu(x) u_i^j(\mu; x) = -1_{K_j}(x) u_i(\mu; x)$$

with the characteristic function $1_{K_j}$ associated to $K_j$.

This Born matrix can then be related to the Green function associated with the acoustic media. This function, which depends on $x \in \mathbb{R}^3$ and $y \in \mathbb{R}^3$, is symmetric $G(\mu; x, y)$ and is defined for every $y \in \mathbb{R}^3$ as $G(\mu; x, y) = G_{\mu,y}(x)$ with $G_{\mu,y}$ the outgoing solution of

$$\Delta G_{\mu,y}(x) + \mu(x) G_{\mu,y}(x) = -\delta_y(x) \text{ on } \mathbb{R}^3$$

with the Dirac generalized function $\delta_y$ at $x = y$. It follows that the function $u_i$, which solves (36), is explicitly given by

$$u_i(\mu; x) = G(\mu, x, x_i^1).$$

Its partial derivative $u_i^j$, which solves (40) with respect to $\mu$ is given by the representation formula

$$u_i^j(x) = \int_{\mathbb{R}^3} G(\mu; x, y) 1_{K_j}(x) u_i(\mu; y) dy.$$  

Taking into account (43), we obtain

$$u_i^j(x) = \int_{K_j} G(\mu; x, y) G(\mu; y, x_i^1) dy.$$  

This leads to the following simple formula for the Born matrix

$$A_{i,j} = u_i^j(x_i^2) = \int_{K_j} G(\mu; x_i^2, y) G(\mu; y, x_i^1) dy.$$  

For large problems (a large number of sources, receivers and model parameters), the computation of this matrix can be very expensive and can be achieved only thanks to high–performance computing. However, it can be easily evaluated in the case of homogeneous media (i.e. $\mu(y) = \mu_0$)

$$G(\mu; x, y) = \frac{i kr}{4\pi r} \text{ with } k = \sqrt{\mu_0} \text{ and } r = |x - y|.$$  

Practically, this particular choice, can be seen as the initial guess for the acoustic media under study.

Remark: In practice, the number of rows of the Born matrix is much larger than the number of columns, i.e. $I \gg J$.

Let us prove that the Born Matrix defined in (46) can be approximated by a low-rank matrix under suitable assumptions on the location of receivers and sources. Most of the arguments that are present in this Section are rather similar to those developed in the multipole theory to solve the direct problem [5]. The
main ingredient to obtain a low-rank approximation of the Born matrix $A$ is the so-called kernel-independent fast multipole method [2]. This method furnishes a tensorial approximation of the Green matrix under the following assumptions, see Figure 10:

\[ \text{i) The sources } x_i^1 \text{ are included in a 3D-dimensional box } B_1 \]

\[ B_1 = x_i^1 + [-d_1, d_1]^3 \]

with $x_i^1 \in \mathbb{R}^3$, the center of the box and $d_1 > 0$ the size of the box.

\[ \text{ii) The receivers } x_i^2 \text{ are included in a 3D box } B_2 \]

\[ B_2 = x_i^2 + [-d_2, d_2]^3 \]

with $x_i^2 \in \mathbb{R}^3$ the center of the box and $d_2$ the size of the box.

\[ \text{iii) The reflectors } K_j \text{ are included in a 3D box } B_3 \]

\[ B_3 = y_0 + [-d_3, d_3]^3 \]

with $y_0 \in \mathbb{R}^3$ being its center and $d_3$ the size of the box.

\[ \text{iv) The diameters of these three boxes are smaller or equal to a wavelength} \]

\[ d_1 < \lambda, \quad d_2 < \lambda \quad \text{and} \quad d_3 < \lambda. \]

\[ \text{v) The distance } D_1 \text{ between the boxes } B_1 \text{ and } B_3 \text{ and the distance } D_2 \text{ between the boxes } B_2 \text{ and } B_3 \text{ are larger than some wavelengths.} \]

\[ D_1 >> \lambda \quad \text{and} \quad D_2 >> \lambda. \]

In the box $B_\ell$, $1 \leq \ell \leq 3$, we can approximate the function $f_\ell : B_\ell \rightarrow \mathbb{C}$ in the following way

\[ f_\ell(x) = \sum_{m=1}^{M} f_\ell(X_m^\ell) p_m^\ell(x) + \varepsilon_\ell(x) \quad \forall f_\ell \in C^\infty(B_\ell), \]

where for $1 \leq m \leq M$ and $\ell \in [1, 3]$, $X_m^\ell \in \mathbb{R}^3$ are the interpolation points and $p_m^\ell : B_\ell \rightarrow \mathbb{R}$ are interpolating functions. The residual $\varepsilon_M$ is small when the
interpolating function is well chosen (see Appendix C for a possible choice) and the interpolated function is regular.

In the context of the Born matrix, these interpolation functions can be used to
define a tensorial approximation of the Green function

\[
\begin{align*}
G_{k}(x, y) & \simeq \sum_{m=1}^{M} \sum_{n=1}^{M} G_{k}(x_{m}, x_{n}) p_{m}^{i}(x) p_{n}^{j}(y), & x \in B_{1}, y \in B_{3}, \\
G_{k}(x, y) & \simeq \sum_{m=1}^{M} \sum_{n=1}^{M} G_{k}(x_{m}, x_{n}) p_{m}^{i}(x) p_{n}^{j}(y), & x \in B_{2}, y \in B_{3}.
\end{align*}
\]

This brings about residual in the following approximation of the Born matrix

\[
A_{i,j} \simeq \sum_{m_{1}=1}^{M} \sum_{m_{2}=1}^{M} \sum_{n_{1}=1}^{M} \sum_{n_{2}=1}^{M} G_{k}(x_{m_{1}}, y_{n_{1}}) G_{k}(x_{m_{2}}, y_{n_{2}})
\]

\[
p_{m_{1}}^{i}(x_{i_{1}}) p_{m_{2}}^{j}(x_{i_{2}}) \int_{K_{j}} p_{n_{1}}^{j}(y) p_{n_{2}}^{j}(y) dy.
\]

Rearranging the latter, we deduce that the Born matrix \( A \) takes the form

\[
A = A^{1} A^{2} A^{3},
\]

with \( A^{1} \in \mathbb{C}^{I \times M^{2}}, A^{2} \in \mathbb{C}^{M^{2} \times M^{2}} \) and \( A^{3} \in \mathbb{C}^{M^{2} \times J} \)

\[
\begin{align*}
A_{i,m}^{1} & = p_{m_{1}}^{i}(x_{i_{1}}) p_{m_{2}}^{j}(x_{i_{2}}) \\
A_{m,n}^{2} & = G_{k}(x_{m_{1}}, y_{n_{1}}) G_{k}(x_{m_{2}}, y_{n_{2}}) \\
A_{n,j}^{3} & = \int_{K_{j}} p_{n_{1}}^{j}(y) p_{n_{2}}^{j}(y) dy
\end{align*}
\]

where the integers \( m, n \in [1, M^{2}] \) and \( i \in [1, I] \) are related to the integers \( m_{1}, m_{2}, n_{1}, n_{2}, i_{1} \) and \( i_{2} \) by the relations

\[
m = (m_{1} - 1) M + m_{2}, \quad n = (n_{1} - 1) M + n_{2} \quad \text{and} \quad i = i_{1} + i_{2} + 1
\]

Equation (56) reveals that the Born matrix \( A \) can be approximated by a low-rank approximation when \( M^{2} \) is much smaller than \( I \) and \( J \). This is the case, when the number of receivers, sources and reflectors is very large.

We have proved that the Born matrix associated with a homogeneous acoustic medium admits a low-rank approximation under very restrictive assumptions on the location of the receivers, sources and reflectors. These results that use similar arguments to the fast multipole method can be extended. When these assumptions are not fulfilled, a low-rank approximation can also be obtained. It relies on elaborated arguments of the Fast Multipole Method. This will not be presented here due to its complexity.

For elastic media, it is also possible to show that the Born matrix associated with a homogeneous media admits a low-rank approximation. The reader can refer to [2] for the fast multipole method for elastic media.
Appendix C. The Tensorial Chebyshev Interpolation

We would like to briefly recall classical results about the Chebyshev tensorial interpolation of a function \( f \) in the box
\[
B = x_0 + [-d, d]^3.
\]
We denote by \( C_p \) the Chebyshev polynomial of degree \( P > 0 \) which is given by the formula
\[
C_p(Z) = \cos(P \arccos(Z)), \quad Z \in [-1, 1].
\]
For \( 1 \leq p \leq P \), its zeros are denoted by \( Z_p^P = \cos \left( \left( p - \frac{1}{2} \right) \frac{\pi}{P} \right) \). To the zeros \( Z_p^P \), that all belong to \([-1, 1]\), we associate the Lagrangian interpolation polynomial
\[
I_p^P(z) = \prod_{k=1 \atop k \neq p}^P \frac{z - Z_k^P}{Z_p^P - Z_k^P} \quad \text{for } p \in [1, P] \text{ and } z \in [-1, 1].
\]
On the interval \([-1, 1]\), any function \( u \) can be approximated by the formula
\[
u(z) = \sum_{p=1}^P u(Z_p^P) I_p^P(z) + \varepsilon_P(z).
\]
These interpolation polynomials are optimal in the sense that they minimize the \( L^\infty\)-norm error
\[
\|\varepsilon_P\|_{L^\infty([-1,1])} = \left( \frac{2}{\pi} \log(P + 1) + 1 \right) \left( \frac{\pi/2}{P!} \right)^P \|u(P)\|_{L^\infty([-1,1])}.
\]
Thanks to this family of unidimensional interpolation functions, we define the tensorial interpolation functions \( p_m : B \rightarrow \mathbb{C} \) on the box \( B \)
\[
p_m(x_0 + z d) = I_{p_1}(z_1) I_{p_2}(z_2) I_{p_3}(z_3) \quad \text{with } z = (z_1, z_2, z_3) \in [-1, 1]^3.
\]
and the interpolation points \( X_m \in B \)
\[
X_m = x_0 + d (Z_{p_1}^P, Z_{p_2}^P, Z_{p_3}^P)
\]
where we have denoted by \( m \in [1, M] \), with \( M = P^3 \), the integer defined by
\[
m = p_1 P^2 + p_2 P + p_1 \quad \text{with } p_1, p_2 \text{ and } p_3 \in [1, P]\]
It follows that a function \( f : B \rightarrow \mathbb{C} \) can be approximated in the following way
\[
f(x) \approx \sum_{m=1}^M p_m(x) f(X_m)
\]

References


Sergey Alexeyevich Solovev
Institute of Petroleum Geology and Geophysics SB RAS,
pr. Koptyuga, 3,
630090, Novosibirsk, Russia
E-mail address: selevessa@ipgg.sbras.ru

Sebastien Tordeux
INRIA Bordeaux Sud-Ouest, Equipe-Projet Magique-3D IPRA-LMA
Université de Pau et des Pays de l’Adour
BP 1155, 64013 Pau Cedex, Université de Pau, France
E-mail address: sebastien.tordeux@gmail.com