



**HAL**  
open science

# Orthogonalization schemes in tensor train format

Martina Iannacito, Olivier Coulaud, Luc Giraud

► **To cite this version:**

Martina Iannacito, Olivier Coulaud, Luc Giraud. Orthogonalization schemes in tensor train format. SIAM Conference on Applied Linear Algebra (LA24), May 2024, Paris, France. hal-04587668

**HAL Id: hal-04587668**

**<https://inria.hal.science/hal-04587668>**

Submitted on 24 May 2024

**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.



Distributed under a Creative Commons Attribution 4.0 International License

*Inria*

# Orthogonalization schemes in tensor train format

SIAM Conference on Applied Linear Algebra,  
May 14, 2024, Paris

Martina Iannacito\*  
joint work with Olivier Coulaud<sup>†</sup> and Luc Giraud<sup>†</sup>

\* Università di Bologna

<sup>†</sup> Centre Inria at the University of Bordeaux



Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



Italiadomani  
PIANO NAZIONALE  
DI RIPRESA E RESILIENZA



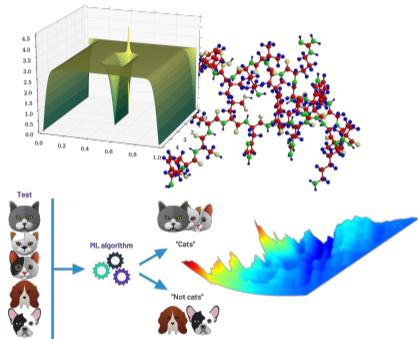
ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

Computing an orthogonal basis is a fundamental sub-problem for numerical simulations

Computing an orthogonal basis is a fundamental sub-problem for numerical simulations

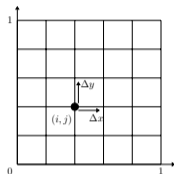
**High-dimensional** simulations arise in

- Stochastic equations
- Uncertainty quantification problems
- Quantum and vibration chemistry
- Optimization
- Machine learning



## Example

$$\begin{cases} \mathcal{L}(u) = f & \text{in } \Omega \\ u = f_0 & \text{in } \partial\Omega \end{cases} \quad \text{for } \Omega \subseteq \mathbb{R}^{n_1 \times \dots \times n_d}$$



$$\mathcal{A}\mathcal{X} = \mathcal{B}$$

where  $\mathcal{A}$  is a multilinear operator and  $\mathcal{B}$  is a tensor.

$$\begin{bmatrix} 1 & 8 & 4 \\ 9 & 2 & 3 \\ 7 & 6 & 7 \end{bmatrix}$$

# The curse of dimensionality

The storage cost increases **exponentially** with respect to the number of variables.



Several tensor decomposition techniques exist

- Canonical Polyadic
- Tucker
- Hierarchical Tucker
- Tensor-Train

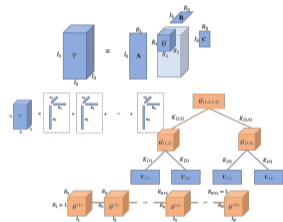


Figure: from [Bi et al. 2022]

# The curse of dimensionality

The storage cost increases **exponentially** with respect to the number of variables.



Several tensor decomposition techniques exist

- Canonical Polyadic
- Tucker
- Hierarchical Tucker
- **Tensor-Train**

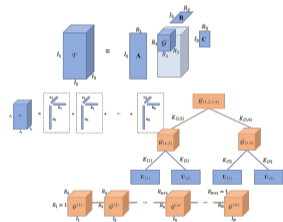




Figure: from [Bi et al. 2022]



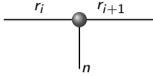
scalar



vector



matrix




order-3 tensor






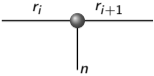
scalar



vector

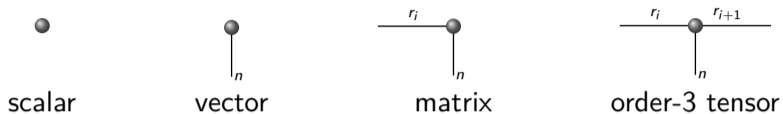


matrix

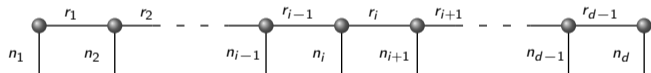


order-3 tensor

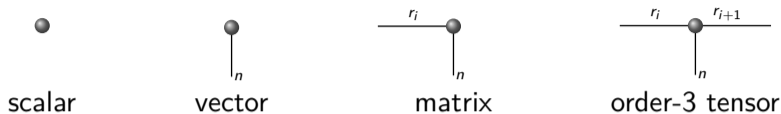
The TT-representation of  $\mathcal{X}$  an  $(n_1 \times \dots \times n_d)$  tensor is



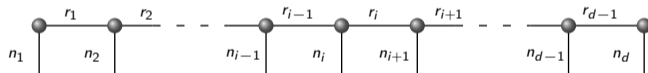
The TT-representation of  $\mathcal{X}$  an  $(n_1 \times \dots \times n_d)$  tensor is



+ The storage cost is  $\mathcal{O}(dnr^2)$  with  $r = \max\{r_i\}$ , said **TT-ranks**



The TT-representation of  $\mathcal{X}$  an  $(n_1 \times \dots \times n_d)$  tensor is



- + The storage cost is  $\mathcal{O}(dnr^2)$  with  $r = \max\{r_i\}$ , said **TT-ranks**
- Linear combinations increase the TT-ranks

## TT-rounding [Oseledets 2011]

If  $\mathcal{X}$  is a tensor in TT-format and  $\delta \in (0, 1)$ , then  $\bar{\mathcal{X}} = \text{TT-rounding}(\mathcal{X}, \delta)$  such that

$$\|\mathcal{X} - \bar{\mathcal{X}}\| \leq \delta \|\mathcal{X}\|$$

where  $\max(\text{TT-ranks}(\bar{\mathcal{X}})) < (\max \text{TT-ranks}(\mathcal{X}))$

TT-rounding [Oseledets 2011]

If  $\mathcal{X}$  is a tensor in TT-format and  $\delta \in (0, 1)$ , then  $\bar{\mathcal{X}} = \text{TT-rounding}(\mathcal{X}, \delta)$  such that

$$\|\mathcal{X} - \bar{\mathcal{X}}\| \leq \delta \|\mathcal{X}\|$$

where  $\max(\text{TT-ranks}(\bar{\mathcal{X}})) < (\max \text{TT-ranks}(\mathcal{X}))$



## Question

How the introduction of compression steps affects the **loss of orthogonality**

$$\|\mathbb{I}_k - \mathcal{Q}_k^\top \mathcal{Q}_k\|$$

of computed basis  $\mathcal{Q}_k$ ?

# Classical matrix results

The **loss of orthogonality** of the computed orthogonal basis  $\mathbf{Q}_k$

- depends on the computational precision  $u$
- may be affected by the collinearity of the input vectors  $\mathbf{A}_k = [\mathbf{a}_1, \dots, \mathbf{a}_k]$ , expressed by  $\kappa(A_k)$

# Classical matrix results

The **loss of orthogonality** of the computed orthogonal basis  $\mathbf{Q}_k$

- depends on the computational precision  $u$
- may be affected by the collinearity of the input vectors  $\mathbf{A}_k = [\mathbf{a}_1, \dots, \mathbf{a}_k]$ , expressed by  $\kappa(\mathbf{A}_k)$

<i>Source</i>	<i>Algorithm</i>	$\ \mathbb{I}_k - \mathbf{Q}_k^T \mathbf{Q}_k\ $
[Stathopoulos and Wu 2002]	<b>Gram</b>	$\mathcal{O}(u\kappa^2(\mathbf{A}_k))$
[Giraud, Langou, and Rozložník 2005]	<b>CGS</b>	$\mathcal{O}(u\kappa^2(\mathbf{A}_k))$
[Björck 1967]	<b>MGS</b>	$\mathcal{O}(u\kappa(\mathbf{A}_k))$
[Giraud, Langou, and Rozložník 2005]	<b>CGS2</b>	$\mathcal{O}(u)$
[Giraud, Langou, and Rozložník 2005]	<b>MGS2</b>	$\mathcal{O}(u)$
[Wilkinson 1965]	<b>Householder</b>	$\mathcal{O}(u)$

# 1

## TT-orthogonalization kernels



# Classical and Modified Gram-Schmidt

---

$$\mathcal{Q}, \mathbf{R} = \text{CGS}(\mathcal{A}, \delta)$$

---

**Input:**  $\mathcal{A} = \{\mathcal{A}_1, \dots, \mathcal{A}_m\}$ ,  $\delta \in \mathbb{R}_+$

```
1 for  $i = 1, \dots, m$  do
2    $\mathcal{P} = \mathcal{A}_i$ 
3   for  $j = 1, \dots, i - 1$  do
4      $\mathbf{R}(i, j) = \langle \mathcal{A}_i, \mathcal{Q}_j \rangle$ 
5      $\mathcal{P} = \mathcal{P} - \mathbf{R}(i, j)\mathcal{Q}_j$ 
6   end
7    $\mathcal{P} = \text{TT-rounding}(\mathcal{P}, \delta)$ 
8    $\mathbf{R}(i, i) = \|\mathcal{P}\|$ 
9    $\mathcal{Q}_i = 1/\mathbf{R}(i, i)\mathcal{P}$ 
10 end
```

**Output:**  $\mathcal{Q} = \{\mathcal{Q}_1, \dots, \mathcal{Q}_m\}$ ,  $\mathbf{R}$

---

---

$$\mathcal{Q}, \mathbf{R} = \text{MGS}(\mathcal{A}, \delta)$$

---

**Input:**  $\mathcal{A} = \{\mathcal{A}_1, \dots, \mathcal{A}_m\}$ ,  $\delta \in \mathbb{R}_+$

```
1 for  $i = 1, \dots, m$  do
2    $\mathcal{P} = \mathcal{A}_i$ 
3   for  $j = 1, \dots, i - 1$  do
4      $\mathbf{R}(i, j) = \langle \mathcal{P}, \mathcal{Q}_j \rangle$ 
5      $\mathcal{P} = \mathcal{P} - \mathbf{R}(i, j)\mathcal{Q}_j$ 
6   end
7    $\mathcal{P} = \text{TT-rounding}(\mathcal{P}, \delta)$ 
8    $\mathbf{R}(i, i) = \|\mathcal{P}\|$ 
9    $\mathcal{Q}_i = 1/\mathbf{R}(i, i)\mathcal{P}$ 
10 end
```

**Output:**  $\mathcal{Q} = \{\mathcal{Q}_1, \dots, \mathcal{Q}_m\}$ ,  $\mathbf{R}$

---

They readily write in TT-format.

# CGS and MGS with reorthogonalization

---

 $\mathcal{Q}, \mathbf{R} = \text{CGS2}(\mathcal{A}, \delta)$ 

---

**Input:**  $\mathcal{A} = \{\mathcal{A}_1, \dots, \mathcal{A}_m\}$ ,  $\delta \in \mathbb{R}_+$

```
1 for  $i = 1, \dots, m$  do
2    $\mathcal{P}_k = \mathcal{A}_i$ 
3   for  $k = 1, 2$  do
4      $\mathcal{P}_k = \mathcal{P}_{k-1}$ 
5     for  $j = 1, \dots, i - 1$  do
6        $\mathbf{R}(i, j) = \langle \mathcal{P}_{k-1}, \mathcal{Q}_j \rangle$ 
7        $\mathcal{P}_k = \mathcal{P}_k - \mathbf{R}(i, j)\mathcal{Q}_j$ 
8     end
9      $\mathcal{P}_k = \text{TT-rounding}(\mathcal{P}_k, \delta)$ 
10  end
11   $\mathbf{R}(i, i) = \|\mathcal{P}_2\|$ 
12   $\mathcal{Q}_i = 1/\mathbf{R}(i, i)\mathcal{P}_2$ 
13 end
```

**Output:**  $\mathcal{Q} = \{\mathcal{Q}_1, \dots, \mathcal{Q}_m\}$ ,  $\mathbf{R}$

---

---

 $\mathcal{Q}, \mathbf{R} = \text{MGS2}(\mathcal{A}, \delta)$ 

---

**Input:**  $\mathcal{A} = \{\mathcal{A}_1, \dots, \mathcal{A}_m\}$ ,  $\delta \in \mathbb{R}_+$

```
1 for  $i = 1, \dots, m$  do
2    $\mathcal{P}_k = \mathcal{A}_i$ 
3   for  $k = 1, 2$  do
4      $\mathcal{P}_k = \mathcal{P}_{k-1}$ 
5     for  $j = 1, \dots, i - 1$  do
6        $\mathbf{R}(i, j) = \langle \mathcal{P}_k, \mathcal{Q}_j \rangle$ 
7        $\mathcal{P}_k = \mathcal{P}_k - \mathbf{R}(i, j)\mathcal{Q}_j$ 
8     end
9      $\mathcal{P}_k = \text{TT-rounding}(\mathcal{P}_k, \delta)$ 
10  end
11   $\mathbf{R}(i, i) = \|\mathcal{P}_2\|$ 
12   $\mathcal{Q}_i = 1/\mathbf{R}(i, i)\mathcal{P}_2$ 
13 end
```

**Output:**  $\mathcal{Q} = \{\mathcal{Q}_1, \dots, \mathcal{Q}_m\}$ ,  $\mathbf{R}$

---

## Gram approach - matrix format

Let  $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_m]$ , then we look for  $\mathbf{A} = \mathbf{QR}$  with  $\mathbf{Q}^\top \mathbf{Q} = \mathbb{I}_m$   
compute the Gram matrix

$$\mathbf{A}^\top \mathbf{A} = (\mathbf{R}^\top \mathbf{Q}^\top) \mathbf{Q} \mathbf{R} = \mathbf{R}^\top \mathbf{R}$$

this is (almost) the **Cholesky** factorization of  $\mathbf{A}^\top \mathbf{A}$  that can be written as

$$\mathbf{A}^\top \mathbf{A} = \mathbf{R}^\top \mathbf{R} = \mathbf{L} \mathbf{L}^\top$$

with the Cholesky factor  $\mathbf{L} = \mathbf{R}^\top$ . Thus, it follows

$$\mathbf{A} = \mathbf{QR} = \mathbf{QL}^\top$$

from which we obtain  $\mathbf{Q}$  by solving a linear system.

---

$$\mathcal{Q}, \mathbf{R} = \text{Gram}(\mathcal{A}, \delta)$$

---

**Input:**  $\mathcal{A} = \{\mathcal{A}_1, \dots, \mathcal{A}_m\}$ ,  $\delta \in \mathbb{R}_+$

1  $\mathbf{G}$  is  $(m \times m)$  Gram matrix from  $\mathcal{A}$

2  $\mathbf{L} = \text{cholesky}(\mathbf{G})$

3  $\{\mathcal{P}_1, \dots, \mathcal{P}_m\}$  from  $\mathcal{A}$  and  $(\mathbf{L}^\top)^{-1}$

4 **for**  $i = 1, \dots, m$  **do**

5 |  $\mathcal{Q}_i = \text{TT-rounding}(\mathcal{P}_i, \delta)$

6 **end**

**Output:**  $\mathcal{Q} = \{\mathcal{Q}_1, \dots, \mathcal{Q}_m\}$ ,  $\mathbf{R}$

---

In TT-format the following modifications occur

- $\mathbf{G}(i, j) = \langle \mathcal{A}_i, \mathcal{A}_j \rangle$
- $(\mathbf{L}^\top)^{-1}$  is explicitly computed
- $\mathcal{P}_i$  is constructed as a linear combination of  $\mathcal{A}$  elements

# Householder transformation

Let  $\mathbf{x}$  and  $\mathbf{y} \in \mathbb{R}^n$  be vectors with  $\|\mathbf{y}\| = 1$ , the Householder reflector  $\mathbf{H}$  reflects  $\mathbf{x}$  along  $\mathbf{y}$ , i.e.,  $\mathbf{H}\mathbf{x} = \|\mathbf{x}\|\mathbf{y}$ . The Householder reflector  $\mathbf{H}$  is

$$\mathbf{H} = \mathbb{I}_n - \frac{2}{\|\mathbf{z}\|^2} \mathbf{z}\mathbf{z}^\top \quad \text{with} \quad \mathbf{z} = (\mathbf{x} - \|\mathbf{x}\|\mathbf{y}).$$

## Remark

- The Householder algorithm uses some of the components of the input vectors
- The Householder algorithm does **not** readily apply to tensor in TT-formats
- The approximation is applied to
  - \* the tensor defining the Householder reflector
  - \* the reflected tensors
  - \* the tensor of the orthogonal basis

# 2

## Numerical experiments

## Numerical experiments setting

We construct a set of 'Krylov TT-tensors',  $\{\mathcal{A}_k\}$ , such that

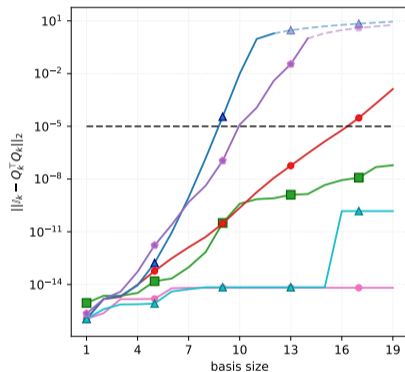
- $\mathcal{A}_1$  is randomly generated with  $\max(\text{TT-rank}) = 1$
- $\mathcal{X}_k = 1/\|\mathcal{A}_k\| \mathcal{A}_k$
- $\mathcal{A}_{k+1} = \text{TT-rounding}(\Delta_d \mathcal{X}_k, \max\_rank = 1)$  where  $\Delta_d$  is be the TT-Laplacian

The perturbations are due to

- the inexactness from the approximation at accuracy  $\delta = 10^{-5}$
- the rounding error from the floating point system at precision  $u = \mathcal{O}(10^{-16})$

# Numerical experiments results

Loss of orthogonality for  $\{\mathcal{A}_k\}$

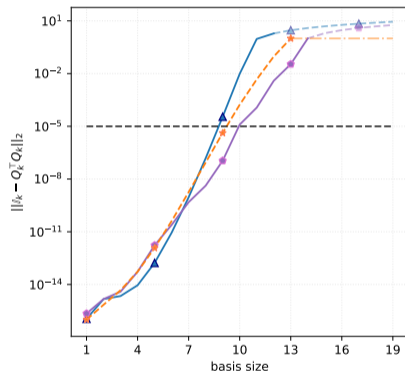


- Gram approach
- CGS
- MGS
- CGS2
- MGS2
- Householder transformation

Figure: 20 tensors of order  $d = 3$  and mode size  $n = 15$



## Loss of orthogonality for $\{\mathcal{A}_k\}$



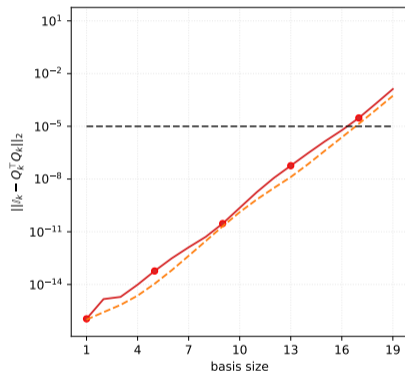
- Gram approach
- CGS
- $\delta\kappa^2(\mathbf{A}_k)$  with

$$\mathbf{A}_k = [\text{vec}(\mathcal{A}_1), \dots, \text{vec}(\mathcal{A}_k)]$$

Figure: 20 tensors of order  $d = 3$  and mode size  $n = 15$

# Numerical experiments results

Loss of orthogonality for  $\{\mathcal{A}_k\}$



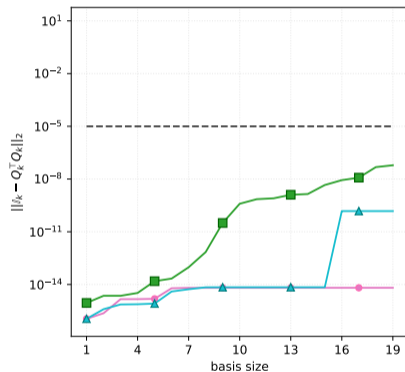
- **MGS**
- $\delta\kappa(\mathbf{A}_k)$  with

$$\mathbf{A}_k = [\text{vec}(\mathcal{A}_1), \dots, \text{vec}(\mathcal{A}_k)]$$

Figure: 20 tensors of order  $d = 3$  and mode size  $n = 15$

# Numerical experiments results

Loss of orthogonality for  $\{\mathcal{A}_k\}$



- CGS2
- MGS2
- Householder transformation

Figure: 20 tensors of order  $d = 3$  and mode size  $n = 15$

<i>Algorithm</i>	Matrix, theoretical	TT-format, conjecture
	$\ \mathbb{I}_k - \mathbf{Q}_k^\top \mathbf{Q}_k\ $	$\ \mathbb{I}_k - \mathcal{Q}_k^\top \mathcal{Q}_k\ $
<b>Gram</b>	$\mathcal{O}(u\kappa^2(\mathbf{A}_k))$	$\mathcal{O}(\delta\kappa^2(\mathcal{A}_k))$
<b>CGS</b>	$\mathcal{O}(u\kappa^2(\mathbf{A}_k))$	$\mathcal{O}(\delta\kappa^2(\mathcal{A}_k))$
<b>MGS</b>	$\mathcal{O}(u\kappa(\mathbf{A}_k))$	$\mathcal{O}(\delta\kappa(\mathcal{A}_k))$
<b>CGS2</b>	$\mathcal{O}(u)$	$\mathcal{O}(\delta)$
<b>MGS2</b>	$\mathcal{O}(u)$	$\mathcal{O}(\delta)$
<b>Householder</b>	$\mathcal{O}(u)$	$\mathcal{O}(\delta)$

with  $u$  the computational precision,  $\delta$  the compression accuracy

## Computational costs: matrix vs tensor

	cost in fp operations	cost in TT-rounding
<b>Gram</b>	$\mathcal{O}(2nm^2)$	$m$
<b>CGS</b>	$\mathcal{O}(2nm^2)$	$m$
<b>MGS</b>	$\mathcal{O}(2nm^2)$	$m$
<b>CGS2</b>	$\mathcal{O}(4nm^2)$	$2m$
<b>MGS2</b>	$\mathcal{O}(4nm^2)$	$2m$
<b>Householder</b>	$\mathcal{O}(2nm^2 - 2m^3/3)$	$4m$

where  $m$  is the basis size and  $n$  is the length of each element

# 3

## Outlook

# Conclusions and perspectives

- All the 6 kernels can be generalized to the TT-framework
- Loss of orthogonality bounds appears to hold true with
  - \* the compression accuracy  $\delta$  replacing the computational precision  $u$
  - \* the compression accuracy  $\delta$  can be read as a computational precision with two peculiarities
    - > it is independent from the machine architecture
    - > it controls the perturbation norm-wise rather than component-wise

More detailed results can be found at [**Coulaud2022-Ortho**]

# Conclusions and perspectives

- All the 6 kernels can be generalized to the TT-framework
- Loss of orthogonality bounds appears to hold true with
  - \* the compression accuracy  $\delta$  replacing the computational precision  $u$
  - \* the compression accuracy  $\delta$  can be read as a computational precision with two peculiarities
    - > it is independent from the machine architecture
    - > it controls the perturbation norm-wise rather than component-wise

More detailed results can be found at [**Coulaud2022-Ortho**]

What is left out?

- Theoretical proof of the loss of orthogonality bounds in the inexact framework
- Investigating the quality of the tensor subspace spanned by the orthogonal basis



- [1] Y. Bi et al. "Chapter 1 - Tensor decompositions: computations, applications, and challenges". In: *Tensors for Data Processing*. Ed. by Y. Liu. Academic Press, 2022, pp. 1–30. ISBN: 978-0-12-824447-0. DOI: <https://doi.org/10.1016/B978-0-12-824447-0.00007-8>. URL: <https://www.sciencedirect.com/science/article/pii/B9780128244470000078>.
- [2] Å. Björck. "Solving linear least squares problems by Gram-Schmidt orthogonalization". In: *BIT Numerical Mathematics* 7.1 (Mar. 1967), pp. 1–21. DOI: 10.1007/BF01934122.
- [3] L. Giraud, J. Langou, and M. Rozložník. "The loss of orthogonality in the Gram-Schmidt orthogonalization process". In: *Computers & Mathematics with Applications* 50.7 (2005). Numerical Methods and Computational Mechanics, pp. 1069–1075. DOI: <https://doi.org/10.1016/j.camwa.2005.08.009>.
- [4] I. V. Oseledets. "Tensor-Train Decomposition". In: *SIAM Journal on Scientific Computing* 33.5 (2011), pp. 2295–2317. DOI: 10.1137/090752286.
- [5] A. Stathopoulos and K. Wu. "A Block Orthogonalization Procedure with Constant Synchronization Requirements". In: *SIAM Journal on Scientific Computing* 23.6 (2002), pp. 2165–2182. DOI: 10.1137/S1064827500370883.
- [6] J. H. Wilkinson. *The algebraic eigenvalue problem*. en. Numerical Mathematics and Scientific Computation. Oxford, England: Clarendon Press, 1965.