Low-Rank Tensor Decompositions for High Dimensional Uncertainty Quantification in Electromagnetic Field Problems

Dimitrios Loukrezis[1,2], Ulrich Römer[1,2] and Herbert De Gersem[1,2]
[1]Institut für Theorie Elektromagnetischer Felder, TU Darmstadt, Schlossgartenstraße 8, 64289 Darmstadt, Germany
[2]Graduate School Computational Engineering, TU Darmstadt, Dolivostraße 15, 64293 Darmstadt, Germany

Recently developed low-rank tensor decompositions offer new tools for high dimensional uncertainty quantification. Such approaches do not inherently suffer from the so-called "curse of dimensionality", because their convergence rates depend only linearly on the number of parameters. Therefore they have a theoretical advantage over current state-of-the-art methods, e.g. stochastic collocation on sparse grids. In this work, we offer an overview on low-rank tensor decompositions, justify their use for uncertainty quantification purposes and present a relevant numerical example from electromagnetics.

***Index Terms*—Low-rank tensor decompositions, stochastic collocation, tensor train, uncertainty quantification**

I. Introduction

Uncertainty quantification (UQ) for electromagnetic field applications is primarily based either on sampling or spectral methods. Sampling methods converge slowly but their convergence rates remain unaffected from the number of uncertain parameters, resp. random variables (RVs). On the other hand, spectral methods converge rapidly, but become intractable for a large number of RVs. Combining spectral methods with *low-rank tensor decompositions* [2] can result in linear dependencies w.r.t. the number of RVs, while preserving the desired convergence properties [8]. Therefore, tensor decompositions seem attractive for high dimensional UQ problems.

In [3], the *tensor train* (TT) decomposition was applied to an electrothermal field problem with 12 input RVs. It was shown that the approximation error decreases with increasing TT ranks, yielding accurate moments at a reasonable computational cost. In this work we investigate the efficiency of this approach for a quantity of interest (QoI) of different regularity. Moreover, in the full paper, results for a larger number of input RVs will be given, in order to determine the break-even between tensor decompositions and sparse grids, for specific applications in electromagnetics.

II. Tensors and Tensor Decompositions

We call an N-dimensional array of size I_n in the n-th dimension a *tensor* $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{I_1\times\cdots\times I_N}$. The entries of $\boldsymbol{\mathcal{A}}$ are given by $\boldsymbol{\mathcal{A}}(i_1,...,i_N)$, $i_n = 1,...,I_n$. A tensor's storage and computational complexity is $\mathcal{O}\left(I^N\right)$, $I=\max_n I_n$. This *curse of dimensionality* can be circumvented with the use of low-rank tensor decompositions. The *canonical* tensor rank [2] is defined as the smallest integer R, such that

$$\boldsymbol{\mathcal{A}} = \sum_{r=1}^{R} \mathbf{u}_r^1 \otimes \mathbf{u}_r^2 \otimes \cdots \otimes \mathbf{u}_r^N, \tag{1}$$

where $\otimes$ denotes the Kronecker product and $\mathbf{u}_r^n \in \mathbb{R}^{I_n}$. Determining the canonical rank of a tensor is an NP-hard problem and while algorithms for its numerical computation exist, they are not always robust. On the other hand, using tensor *matricizations* $\mathbf{A}_n \in \mathbb{R}^{I_n\times I_1\cdots I_{n-1}I_{n+1}\cdots I_N}$, the multi-linear or *Tucker* rank $(R_1,...,R_N)$, $R_n = \operatorname{rank}(\mathbf{A}_n)$, can be robustly computed with N *singular value decompositions*. Both rank definitions yield specific tensor decomposition formats.

The *canonical polyadic decomposition* (CPD) yields a fixed-R tensor approximation in the form of (1). Its complexity is $\mathcal{O}(NIR)$, i.e. linear w.r.t. the number of dimensions. Greedy algorithms for the construction of CPDs, e.g. the proper generalized decomposition, have been found to work well for some UQ problems [4]. However, this format suffers from the same lack of robustness as the canonical-rank problem.

The Tucker decomposition, in element-wise notation, reads

$$\boldsymbol{\mathcal{A}}(i_1,...,i_N) \approx \sum_{r_1,...,r_N} \boldsymbol{\mathcal{C}}(r_1,...,r_N) \prod_{n=1}^{N} \mathbf{U}_n(i_n, r_n), \tag{2}$$

with indices $r_n = 1,...,R_n$, core tensor $\boldsymbol{\mathcal{C}} \in \mathbb{R}^{R_1\times\cdots\times R_N}$ and factor matrices $\mathbf{U}_n \in \mathbb{R}^{I_n\times R_n}$. The Tucker decomposition is robust and offers major compressions for $R_n \ll I_n$, but suffers from the curse of dimensionality, as its complexity is $\mathcal{O}\left(R^N + NIR\right)$, $R = \max_n R_n$.

The TT decomposition [5], also known as the *matrix product states* (MPS) format, combines both robustness and low complexity. In element-wise notation, it reads

$$\boldsymbol{\mathcal{A}}(i_1,...,i_N) \approx \sum_{r_0,...,r_N} \prod_{n=1}^{N} \boldsymbol{\mathcal{G}}_n(r_{n-1}, i_n, r_n) \tag{3}$$

$$= \mathbf{G}_1^{i_1}\mathbf{G}_2^{i_2}\cdots\mathbf{G}_N^{i_N}, \tag{4}$$

with indices $r_n = 1,...,R_n$, $R_0 = R_N = 1$, and TT-cores $\boldsymbol{\mathcal{G}}_n \in \mathbb{R}^{R_{n-1}\times I_n\times R_n}$. In the equivalent MPS format of (4), matrices $\mathbf{G}_n^{i_n} \in \mathbb{R}^{R_{n-1}\times R_n}$ are slices of the TT-cores $\boldsymbol{\mathcal{G}}_n$ for fixed i_n, such that $\mathbf{G}_n^{i_n} = \boldsymbol{\mathcal{G}}_n(:, i_n, :)$. The TT format's complexity is $\mathcal{O}\left(NIR^2\right)$, i.e. also linear w.r.t. the number of dimensions. Moreover, there exist TT-based *cross approximation* algorithms [6], [7] for function-related tensors, i.e. $\boldsymbol{\mathcal{A}}(i_1,...,i_N) = f\left(x_1^{(i_1)},...,x_N^{(i_N)}\right)$ with $x_n^{(i_n)}$ being discrete values of variable x_n. These algorithms yield the TT decomposition without ever computing the full tensor, resulting in computational complexities of $\mathcal{O}\left(NIR^3\right)$.

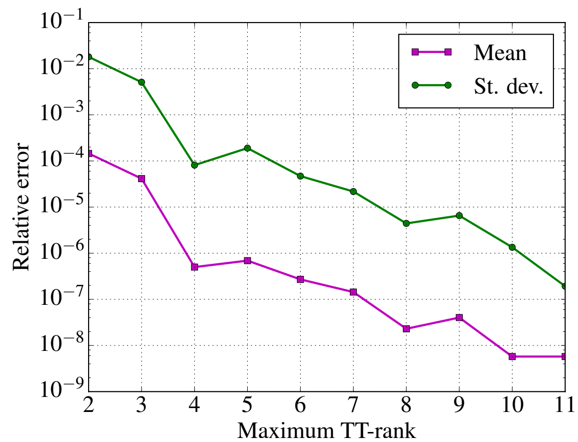

Fig. 1. Relative errors vs maximum TT-rank of $\mathcal{Q}_{\mathrm{TT}}$.

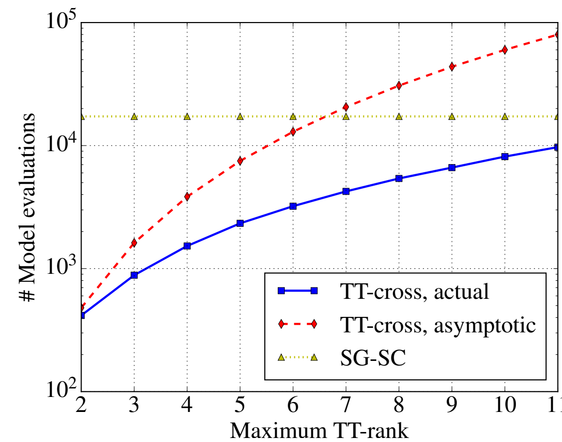

Fig. 2. Computational costs vs maximum TT-rank of $\mathcal{Q}_{\mathrm{TT}}$.

III. Tensor Decompositions for UQ

We want to estimate moments of a stochastic model's QoI $\mathcal{Q}$, dependent on a set of independent RVs $\mathbf{y} = (y_1, \dots, y_N)$, such that $\mathcal{Q} = \mathcal{Q}(\mathbf{y})$. Using the stochastic collocation (SC) method [1], the mean value is computed with a numerical quadrature scheme, such that

$$\mathbb{E}[\mathcal{Q}] \approx \sum_{i_1,\dots,i_N} w_1^{(i_1)} \cdots w_N^{(i_N)} \mathcal{Q}\left(y_1^{(i_1)}, \dots, y_N^{(i_N)}\right), \quad (5)$$

where $y_n^{(i_n)}$ are univariate collocation points, resp. quadrature abscissas, $w_n^{(i_n)}$ the corresponding weights and $i_n = 1, \dots, I_n$. The QoI is evaluated for each multivariate collocation point $\left(y_1^{(i_1)}, \dots, y_N^{(i_N)}\right)$. The *tensor product* (TP) SC results in a complexity of $\mathcal{O}\left(I^N\right)$. This complexity can be mitigated to $\mathcal{O}\left(I\left(\log I\right)^{N-1}\right)$ with the use of *sparse grids* (SG), however the curse of dimensionality remains.

Eq. (5) can be written in the form of a tensor inner product

$$\mathbb{E}\left[\mathcal{Q}\right] \approx \langle \boldsymbol{\mathcal{W}}, \boldsymbol{\mathcal{Q}} \rangle = \sum_{i_1,\dots,i_N} \boldsymbol{\mathcal{W}}\left(i_1, \dots, i_N\right) \boldsymbol{\mathcal{Q}}\left(i_1, \dots, i_N\right), \quad (6)$$

where $\boldsymbol{\mathcal{W}}\left(i_1, \dots, i_N\right) = w_1^{(i_1)} \cdots w_1^{(i_N)}$ and $\boldsymbol{\mathcal{Q}}\left(i_1, \dots, i_N\right) = \mathcal{Q}\left(y_1^{(i_1)}, \dots, y_N^{(i_N)}\right)$. Using univariate weight vectors $\mathbf{w}_n = \left(w_n^{(1)}, \dots, w_n^{(I_n)}\right)$, tensor $\boldsymbol{\mathcal{W}}$ can be written as an *exact* rank-1 CPD, i.e. $\boldsymbol{\mathcal{W}}_{\mathrm{CPD}} = \boldsymbol{\mathcal{W}}$. TT-cross approximation algorithms [6], [7], can be used to compute a TT decomposition of tensor $\boldsymbol{\mathcal{Q}}$, such that $\boldsymbol{\mathcal{Q}}_{\mathrm{TT}} \approx \boldsymbol{\mathcal{Q}}$. Due to their linear dependence w.r.t. the number of RVs N and assuming that an accurate low-rank TT approximation of $\boldsymbol{\mathcal{Q}}$ is feasible, low-rank decompositions shall outperform the SG-SC in high dimensions.

IV. Numerical Results and Conclusions

We consider an electrothermal field problem with 12 input RVs, as in [3]. As QoI we choose the ℓ^2-norm $\|\mathbf{T}\|_2$ of the discrete temperature vector $\mathbf{T}$. Norm $\|\mathbf{T}\|_2$ is a smooth function w.r.t. the input RVs. We pick $I_n = 5$ univariate collocation points per RV, thus rendering the TP-SC completely unaffordable, since $5^{12} \approx 2.5 \cdot 10^8$ model evaluations would be required. SG-SC yields approximately $17 \cdot 10^3$ collocation points, resp. weights, which is still a high computational cost.

We regard the mean and standard deviation obtained with the SG-SC as reference values and apply the approach discussed in Section III. Storage of tensor $\boldsymbol{\mathcal{W}}$ requires only $12 \cdot 5 = 60$ entries due to its exact rank-1 CPD. For the TT approximation of tensor $\boldsymbol{\mathcal{Q}}$, the rank-adaptive *greedy-TT-cross* algorithm [7] is employed. Fig. 1 presents the relative errors of the moments in relation to the maximum rank of $\boldsymbol{\mathcal{Q}}_{\mathrm{TT}}$. The error mostly decreases with increasing maximum TT-ranks, but there exist cases where stagnation or slight error increases are observed. Fig. 2 presents the computational costs of the TT cross approximation per maximum TT-rank, as well as the asymptotic costs. Because the maximum rank appears only in a few cores, the asymptotic complexity $\mathcal{O}\left(NIR^3\right)$ is, in this case, a very pessimistic upper limit.

In view of the numerical results, it can be deduced that accurate moments are computed with the approach based on tensor decompositions. Moreover, due to the rank-adaptive procedure, the computational cost is minimized w.r.t. the goal function under consideration and the desired accuracy.

References

[1] Babuška I., Nobile F. and Tempone R., *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM Rev. 52:317-355, 2010.

[2] Kolda T.G. and Bader B.W., *Tensor decompositions and applications*, SIAM Rev. 51:455-500, 2009.

[3] Loukrezis D., Römer U., Casper T., Schöps S. and De Gersem H., *High Dimensional Uncertainty Quantification for an Electrothermal Field Problem using Stochastic Collocation on Sparse Grids and Tensor Train Decompositions*, Int. J. Numer. Model. El., 2017.

[4] Nouy A., *Proper generalized decomposition and separated representations for the numerical solution of high dimensional stochastic problems*, Arch. Comput. Meth. 17:403-434, 2010.

[5] Oseledets I.V., *Tensor-train decomposition*, SIAM J. Sci. Comput. 33:2295-2317, 2011.

[6] Oseledets I.V. and Tyrtyrshnikov E.E., *TT-cross approximation for multidimensional arrays*, Linear Algebra Appl. 432:70-88, 2010.

[7] Savostyanov D.V., *Quasioptimality of maximum-volume cross interpolation of tensors*, Linear Algebra Appl. 458:217-244, 2014.

[8] Uschmajew A., *Zur Theorie der Niedrigrangapproximation in Tensorprodukten von Hilberträumen*, PhD Thesis, TU Berlin, 2013.