ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-MOPTS016

COMPRESSION AND NOISE REDUCTION OF FIELD MAPS

X. Du and L. Groening,
GSI Helmholtzzentrum für Schwerionenforschung GmbH, Darmstadt, Germany

Abstract

Errors from discretization and large data volume of field maps are a concern for beam dynamics simulations w.r.t. achievable accuracy and to the required amount of time. High-order Singular Value Decomposition (HOSVD) has recently emerged as simple, effective, and adaptive tool to extract the essentials from multi-dimensional data. This paper is on the feasibility of compression and noise reduction of electromagnetic field map data with HOSVD. The method has been applied to an electric field map of a DTL cavity with 11 m in length comprising 55 rf-gaps. The original field map data of 220 MB was converted into practically noise-free data of just 20 KB. Noise was reduced by 95% as demonstrated using a cubic cavity for which the analytical field map is available.

METHOD

HOSVD is a multi-linear generalization of matrix SVD to high-order tensors, and it can provide for adequate preservation of the data's essentials through representation of a tensor using appropriate bases [1,2].

The original field map stored in the N^{th} order ($I_1 \times I_2 \times \cdots I_N$) tensor A. It is decomposed with HOSVD and expressed through the new base using the matrices (a set of bases) $U^{(n)}$ as

$$A = S \times_1 U^{(1)} \times_2 U^{(2)} \cdots \times_N U^{(N)}$$

Where $\times_n$ denote the n-mode product treating n-mode vectors of left tensor as horizontal vectors and multiplying with right side matrix. The n-mode SVD found a set of orthogonal bases $U^{(n)} \in R^{I_n \times I_n}$, which is sorted according to their singular values. $S \in R^{I_1 \times I_2 \times \cdots I_N}$ is a tensor with same size of A, i.e., the projection of A onto new basis $U^{(1)}, U^{(2)} \ldots U^{(N)}$.

Data Trimming

The contents of the matrices $U^{(n)}$ are reduced by deleting from them the vectors (listed in the rows) that correspond to low singular values. Doing so the dimensions of $U^{(n)}$ are reduced. The cut pendants of $U^{(n)}$ are introduced as $u^{(n)} \in R^{(C_n < I_n) \times I_n}$ with their amount of rows being reduced to C_n, i.e., the vectors corresponding to the lowest $I_n - C_n$ singular values where removed since they are attributed to noise and/or errors. Accordingly, the corresponding projections comprising the core tensor S need to be deleted as well, hence cutting S to s. This is done by reducing the "thickness" of S along dimension n by removing those layers along n that correspond to the rows being deleted from $U^{(n)}$.

Compressed and De-Noised Data

After this noise reduction the remaining de-noised data is stored into the tensor $\hat{A}$ ($\hat{A} \approx A$ and $\hat{A} \in R^{I_1 \times I_2 \times \cdots I_N}$) expressed through

$$\hat{A} = s \times_1 u^1 \times_2 u^2 \cdots \times_N u^N$$

Where $\hat{A}$ contains the "recuperated from compressed" (RC) data. It is supposed to be the data's essence without noise, as long as the thresholds C_n is chosen properly. Data compression is through storage of the trimmed data s and u^n instead of the initial data S and U^n. For further processing of the RC data, field components values f at specific locations are needed. They are retrieved through

$$f_{j_1 j_2 \cdots j_N} = s \times_1 u_{j_1}^{(1)} \times_2 u_{j_2}^{(2)} \cdots \times_N u_{j_N}^{(N)}$$

The extraction of de-noised data is done through multi-linear mapping rather than through time consuming memory search. The efficiency of beam dynamic simulation codes can by improved accordingly.

APPLICATION

The method is applied to the electric field map of a drift tube linac (DTL) cavity which is currently under design for the upgrade of the Universal Linear Accelerator (UNILAC) at GSI [3,4]. The map is from CST simulations [5] of the field between drift tubes.

The DTL cavity which serves as test bench for the method comprises 55 cells. To each cell 201*41*41 (long.*hor.*ver.) mesh nodes are assigned. On each node the electric field has three components (E_x, E_y, E_z). The entire field map is stored as a 5th-order tensor and its number of entries along each dimension is (n-cell, z-step, x-step, y-step, spacial component). This tensor is referred to as A in following sections. Accordingly, each scalar field value is uniquely labelled by five numbers, namely: cell, long./hor./ver. node, and spacial component. The field map data of the whole cavity is stored into the 5th-order tensor $A \in R^{55 \times 201 \times 41 \times 41 \times 3}$.

The tensor A is decomposed by HOSVD as

$$A = S \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \times_4 U^{(4)} \times_5 U^{(5)},$$

where $S \in R^{55 \times 201 \times 41 \times 41 \times 3}$, $U^{(1)} \in R^{55 \times 55}$, $U^{(2)} \in R^{201 \times 201}$, $U^{(3)} \in R^{41 \times 41}$, $U^{(4)} \in R^{41 \times 41}$, $U^{(5)} \in R^{3 \times 3}$.

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2019). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI

ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-MOPTS016

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2019). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI

Thresholding

The data trimming is according to the magnitude of the single values corresponding to the vectors forming $U^{(n)}$. For n=1,2,3 these singular values have been sorted according to their magnitude and they are partially shown in Fig. 1. The plot for n=4 has been omitted since this case is equivalent to n=3 and for n=5 the plot is obsolete since $U^{(5)}$ has been replaced by the Identity and no reduction is performed. Figure 1 shows that just few singular values stick out w.r.t. their magnitude from the others.

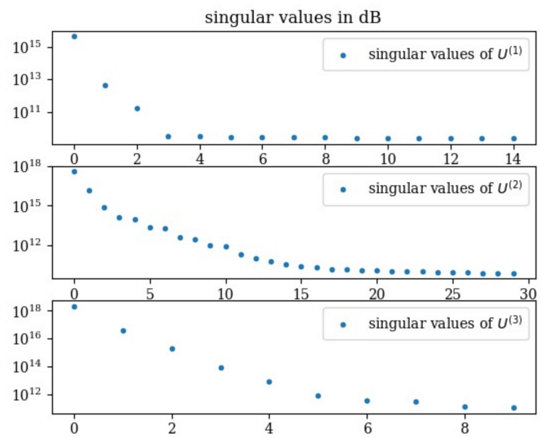

Figure 1: Highest singular values for $U^{(1)}$, $U^{(2)}$ and $U^{(3)}$ for the field map of the DTL cavity. They are sorted according to their magnitude being extracted through HOSVD.

The two largest singular values already differ to each other by about two orders of magnitude. In view of the distribution of singular values the threshold is set to 1/10000 of the largest singular value for each set of bases and the size of the preserved (new) core tensor s is accordingly (3,8,4,4,3) as anticipated in the previous subsection. Accordingly, the de-noised data is expressed as

$$\hat{A} = s \times_1 u^1 \times_2 u^2 \times_3 u^3 \times_4 u^4 \times_5 u^5$$

where $\hat{A} \in R^{55\times201\times41\times41\times3}$, $s \in R^{3\times8\times4\times4\times3}$, $u^{(1)} \in R^{3\times55}$, $u^{(2)} \in R^{8\times201}$, $u^{(3)} \in R^{4\times41}$, $u^{(4)} \in R^{4\times41}$, $u^{(5)} \in R^{3\times3}$. Binary storage of the original data A takes 220 MB of storage volume, while storage of the compressed and de-noised data (s and $u^{(n)}$) requires just 20 kB. Accordingly, reduction of storage place by more than four orders of magnitude has been achieved.

To verify the quality of de-noised field data, the smoothness of field derivatives from simulated data and from HOSVD-processed data has been evaluated and compared. Figure 2 shows the derivatives of E_z and E_y along a DTL cell at an off-axis distance of 5 mm. While the originally simulated field shows non-physical fluctuations violating the Laplace equation, the field from HOSVD-trimming is perfectly smooth and reflects the scenarios' symmetries. Accordingly, HOSVD-trimming can provide for field maps that are very close to analytical solutions even for very complex geometries and large volumes.

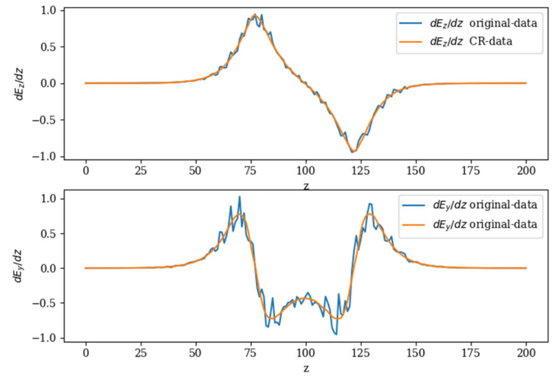

Figure 2: Comparison of the derivatives $\delta E_z/\delta z$ (upper) and $\delta E_y/\delta z$ (lower) along the z-axis along a DTL cell for simulated data and for simulated data being processed through HOSVD. The latter is much smoother and features the symmetries imposed by the geometry.

BENCHMARKING

The HOSVD-based method was benchmarked by comparing its results to an analytical solution using the well-understood cubic cavity. The analytic expressions for the electric field components inside a cubic cavity are

$$E_x = E_1 \cos\left(\frac{m\pi}{a}x\right)\sin\left(\frac{n\pi}{b}y\right)\sin\left(\frac{p\pi}{l}z\right),$$
$$E_y = E_2 \sin\left(\frac{m\pi}{a}x\right)\cos\left(\frac{n\pi}{b}y\right)\sin\left(\frac{p\pi}{l}z\right),$$
$$E_z = E_3 \sin\left(\frac{m\pi}{a}x\right)\sin\left(\frac{n\pi}{b}y\right)\cos\left(\frac{p\pi}{l}z\right),$$
$$\mathrm{E}_1 m + \mathrm{E}_2 n + \mathrm{E}_3 p = 0,$$

where a, b, l are the width, height, and length of the cavity, and m, n, p are the corresponding wave numbers. The analytical field has been scaled to match the amplitude of the simulated field map. In order to approximately match the size of a DTL cell, the dimensions of the cubic cavity were chosen as

$$a = 100\,mm, b = 130\,mm, l = 200\,mm,$$
$$m = 1,\ \ n = 1,\ \ p = 1$$

and the cavity has been simulated using CST-MWS thus providing the discrete numerical field map.

Field maps from both fine mesh and rough mesh simulations were de-noised applying the method. The original field map and the RC field data have been compared to the analytical field. Figure 3 depicts these comparisons revealing that the algorithm reduces significantly the amount of noise for both mesh types and reproduces the analytical field values very well. Using fine mesh simulations, transverse field values differ less than 1% from the analytical field values, being sufficiently accurate for most beam dynamics calculations.

ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-MOPTS016

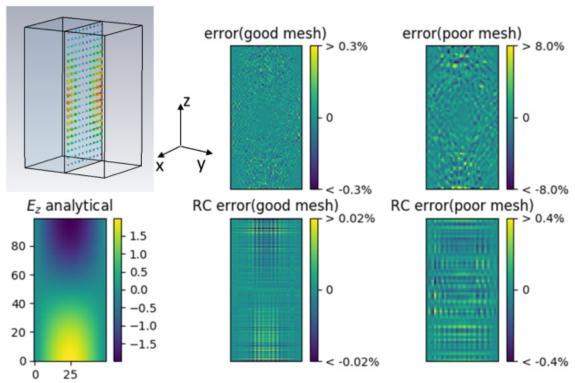

Figure 3: Upper left: definition of the cutting plane within the cubic cavity. Lower left: analytic E_z at the cutting plane. Upper centre: relative difference of E_z from analytical solution and from simulations with fine mesh. Upper right: relative difference of E_z from analytical solution and from simulations with rough mesh. Lower center: relative difference of E_z from analytical solution and from HOSVD starting from fine mesh simulation. Lower right: relative difference of E_z from analytical solution and from HOSVD starting from rough mesh simulations.

Figure 3 demonstrates that HOSVD is a method that can extract the essence of field map data, i.e., field values deviating from the analytical solution by less than 1% even from poor-meshed simulations inhabiting large errors. Additionally, the data volume to be stored is significantly reduced during this extraction.

Integration of DT Field Map

The integral of dipole components is another criterion for quantifying the quality of field maps. The field map of the 55th gap is integrated along the z-axis.

$$G_x(i,j) = \sum_{1}^{201} f(55,1{\sim}201,i,j,1) \approx 0$$

$$G_y(i,j) = \sum_{1}^{201} f(55,1{\sim}201,i,j,2) \approx 0$$

$$G_z(i,j) = \sum_{1}^{201} f(55,1{\sim}201,i,j,3) \approx constant$$

G_x and G_y label the integrated dipole field components inside the cell and should be zero since transverse fields symmetrically cancel each half of the cell. G_z should be constant as the integral always start at one drift tube surface (equipotential area) and end at another one. Figure 4 shows that G_x, G_y and G_z (2D maps of 41*41 pixels) barely vary between de-noised data and original data, giving a conclusion that the algorithm does not change these integrals being global statistical properties of the field map. Integral of noises is supposed to be zero according to its randomness.

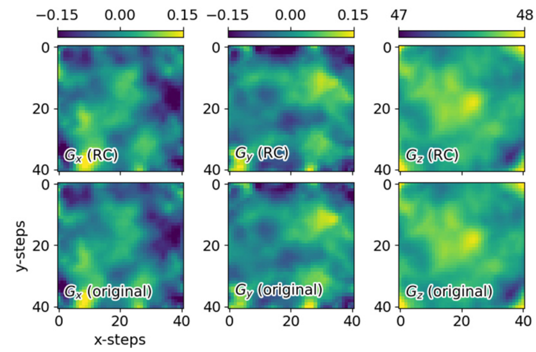

Figure 4: The plot of G_x, G_y and G_z from de-noised (RC) data and from original data. The magnitudes are arbitrary units.

CONCLUSION

Efficient de-noising and hence compression of high dimensional electric field map data by HOSVD has been proposed and successfully applied. The electric field map of an DTL cavity comprising 55 rf-gaps obtained from CST simulations was compressed from an initial data volume of 220 MB by four orders of magnitude to just 20 kB. Noise has been practically removed from the data and the data quality increased significantly w.r.t. redundancy, smoothness, and symmetry. The extraction of de-noised data is done through multi-linear mapping rather than through time consuming memory search. This method is general and applicable to all kind of maps. It will significantly increase the accuracy and hence reliability of codes that apply such maps.

REFERENCES

[1] X. Du and L. Groening, Compression and noise reduction of field maps, *Phys. Rev. Accel. Beams* **21** 064801 (2018). /doi.org/10.1103/PhysRevAccelBeams.21.084601

[2] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A Multilinear Singular Value Decomposition". *SIAM J. Matrix Anal. Appl.* 21-4, pp. 1253-1278, (2000).

[3] L. Groening and S. Mickat, "Development of new Heavy Ion Linacs at GSI", *Proceedings of the XXVIII Linear Accelerator Conference*, East Lansing, MI, USA, edited by Y. Yamazaki (MSU, East Lansing, 2016).

[4] M. Heilmann *et al.*, "FOS Vavity of The Alvarez 2.0 DTL as FAIR Injector", in *Proc. IPAC'19*, Melbourne, Australian.

[5] https:// www.cst.com/Applications/MWandRF

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2019). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI