

## Tensor Methods for Global Sensitivity Analysis

Presentation of work originally published in the *Journal of Reliability Engineering and System Safety*, as well as in the *SIAM/ASA Journal on Uncertainty Quantification*

Rafael Ballester-Ripoll<sup>1</sup> Enrique G. Paredes<sup>2</sup> Renato Pajarola<sup>3</sup>

**Abstract:** Sobol indices and other, more recent quantities of interest (such as the effective and mean dimensions, the dimension distribution, or the Shapley values) are of great aid in sensitivity analysis, uncertainty quantification, and model interpretation. Unfortunately, computing such indices is still challenging for high-dimensional systems. We propose the tensor train decomposition (TT) as a unified framework for surrogate modeling and sensitivity analysis of independently distributed variables. To this end, we introduce the Sobol tensor train (Sobol TT) data structure, which compactly represents variance components for all possible joint variable interactions of any order. Our formulation allows efficient aggregation and subselection operations, and we are able to obtain related Sobol indices and other related quantities at low computational cost.

**Keywords:** sensitivity analysis; sobol indices; surrogate modeling; data visualization; multidimensional data analytics; tensor approximation

### 1 Motivation

Many models in computational science admit the general form  $f(x_1, \dots, x_N) \rightarrow \mathbb{R}$ , given by an analytical formula or an approximator (a regressor) trained over some sample set. Variance-based sensitivity analysis (SA), and *Sobol's method* [Sa04] are powerful tools for model interpretation, where the *Sobol indices* measure the variance in  $f$  that is due to each individual input  $x_1, \dots, x_N$ , as well as any combinations (interactions). Such interactions are crucial to gain insights on the underlying model, as some variables may be important only for certain values of one or more of the remaining variables. Among many other applications, Sobol indices can detect irrelevant variables or variables whose effect can be separated, e.g.  $f(x_1, x_2, \dots) = g(x_1) + h(x_2, \dots)$ . Those indices, however, arise from multidimensional integrals and are thus computationally challenging.

---

<sup>1</sup> Swiss Federal Institute of Technology Zurich, Switzerland barafael@ethz.ch

<sup>2</sup> Swiss National Supercomputing Center enrique.gonzalez@cscs.ch

<sup>3</sup> Department of Informatics, University of Zurich, Switzerland pajarola@ifi.uzh.ch

## 2 Tensor-based Sensitivity Analysis

In order to conduct effective and affordable SA, we propose to use *tensor decompositions*, i.e. the *tensor train* (TT) model [Os11]. It is a numerical framework that scales well with the number of dimensions, and operates on the assumption that the model of interest  $f$  can be well-approximated by a *low-rank tensor*  $\mathcal{T}$  (a *surrogate model*):  $f(x_1, \dots, x_N) \approx \mathcal{T}[i_1, \dots, i_N]$ . The low-rank assumption generalizes the notion of low-rank matrix factorization to three and more dimensions and entails two benefits: it acts as a regularization prior during model fitting, and it drastically reduces the computational cost of model postprocessing. Fortunately, a wide class of functions are well-approximable by low-rank tensor formats.

We first obtain an approximation  $\mathcal{T}$  of a function  $f$  using the *cross-approximation* sampling method [OT10], then derive a new *Sobol tensor train*  $\mathcal{S}$  which is the first data structure that gathers *all*  $2^N$  Sobol indices of an  $N$ -dimensional model. In addition, the Sobol TT can be further manipulated in compact form to (a) extract more advanced sensitivity metrics and statistics, and (b) answer a wide range of queries including e.g. *What are the  $k$  most important variables?*, *What variable interacts the most with  $x_n$ ?*, and many others. To satisfy such queries we propose a new class of so-called *tensor automata*, which are able to select and combine SA indices in many possible ways. For example, the *Hamming mask tensor* computes the combined importance of all  $k$ -plets of variables.

This extended abstract summarizes our work from [BRPP18] and [BRPP19].

### Acknowledgments

This work was partially supported by the UZH Forschungskredit “Candoc”, grant number FK-16-012.

### Bibliography

- [BRPP18] Ballester-Ripoll, Rafael; Paredes, Enrique G.; Pajarola, Renato: Tensor Algorithms for Advanced Sensitivity Metrics. *SIAM/ASA Journal on Uncertainty Quantification*, 6(3):1172–1197, 2018.
- [BRPP19] Ballester-Ripoll, Rafael; Paredes, Enrique G.; Pajarola, Renato: Sobol Tensor Trains for Global Sensitivity Analysis. *Reliability Engineering and System Safety*, 183:311–322, March 2019.
- [Os11] Oseledets, Ivan V.: Tensor-Train Decomposition. *SIAM Journal on Scientific Computing*, 33(5):2295–2317, September 2011.
- [OT10] Oseledets, I. V.; Tyrtyshnikov, E. E.: TT-cross Approximation for Multidimensional Arrays. *Linear Algebra Applications*, 432(1):70–88, 2010.
- [Sa04] Saltelli, Andrea; Tarantola, Stefano; Campolongo, Francesca; Ratto, Marco: *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. Halsted Press, New York, NY, USA, 2004.