Progress of Tensor-Based High-Dimensional Uncertainty Quantification of Process Variations

Zichang He and Zheng Zhang
Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106
zichanghe@ucsb.edu, zhengzhang@ece.ucsb.edu

10.1109/ACES53325.2021.00020

***Abstract*—Uncertainty quantification under process variations is an important topic in computational electromagnetics (EM) and electronic design automation (EDA). The popular stochastic methods become computationally inefficient when the number of uncertain parameters grows. A promising technique to address this challenge is tensor computation. This short paper reviews some recent tensor methods for high-dimensional uncertainty quantification in the field of EDA and computational EM.**

***Keywords*—Uncertainty quantification; tensor; tensor decomposition; high dimensionality; process variation.**

I. Introduction

Process variations, such as intrinsic randomness of material properties, geometrical parameters and temperature, have been a major concern in computational electromagnetics (EM) and electronic design automation (EDA). It is essential to accurately estimate and quantify the uncertainty caused by process variations. The brute-force Monte Carlo (MC) method require lots of simulation data, and each simulations if often time-consuming. Therefore, the past decades have seen a great success of applying advanced uncertainty quantification (UQ) models (like stochastic spectral methods based on generalized polynomial chaos [1]) to the EDA and computational electromagnetic problems [2, 3]. However, the efficiency of most UQ models degrades while the number of uncertain parameters increases since the required number of simulation samples often grow very fast.

Tensor computation is a powerful tool to address the curse of dimensionality in many engineering domains. Leveraging low-rank models in a high dimension, tensor methods may reduce the problem size from exponential function to a (nearly) linear one on the number of uncertain parameter. Due to their impressive performance, tensor methods have been widely used in EM [4–6] and EDA [7–12]. This paper will briefly summarize some recent progress of this topic, emphasizing tensor methods for building high-dimensional generalized polynomial chaos (gPC) models.

II. Short Background

In most computational EM and EDA problems, random parameters $\boldsymbol{\xi} \in \mathbb{R}^d$ describe a set of process variations. The stochastic spectral method approximates the interested metric $y(\boldsymbol{\xi})$ as the summation of a series of orthonormal polynomial basis functions [1]:

$$y(\boldsymbol{\xi}) \approx \sum_{\boldsymbol{\alpha} \in \Theta} c_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \tag{1}$$

where $\boldsymbol{\alpha}$ is an index vector, Θ is the index set, and $\Psi_{\boldsymbol{\alpha}}$ is a polynomial basis function of degree $|\boldsymbol{\alpha}| = \sum_{k=1}^{d} \alpha_k$. In the commonly used total degree scheme, $|\boldsymbol{\alpha}|$ is bounded by order p, leading to a total of $\frac{(d+p)!}{d!p!}$ terms of expansion.

A d-dim tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots n_d}$ represents a d-dimensional data array, and it becomes a matrix when $d = 2$. A d-dim rank-R tensor can be written as the sum of R rank-1 tensors, known as a CP decomposition:

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \cdots \circ \mathbf{u}_r^{(d)}. \tag{2}$$

While many other tensor decomposition formats exist [13], we mainly focus on CP-format-based UQ models in this survey.

III. Survey of Recent Work

As summarized in Table I, recently three tensor methods have been developed to efficiently construct a high-dimensional gPC-like surrogate model for device and circuit simulation with uncertainties.

A. Pseudo Projection via Tensor Recovery [7, 8]

For a standard gPC expansion like Eq. (1), pseudo projection is an accurate approach to compute its coefficients. Specifically, some quadrature points and weights corresponding to $\boldsymbol{\xi}$ and their simulation values are needed to calculate a series of numerical integration.

In [7, 8], the numerical integration is formulated as the inner product of two tensors $c_{\boldsymbol{\alpha}} = \langle \mathcal{Y}, \mathcal{W}_{\boldsymbol{\alpha}} \rangle$, where $\mathcal{Y}$ is filled with $(p+1)^d$ sample simulations and $\mathcal{W}_{\boldsymbol{\alpha}}$ is a known rank-one tensor constructed by a gPC basis function and all quadrature weights. Given limited simulations to fill the entries of $\mathcal{Y}$, the key idea of tensor recovery is to predict the missing elements based on the low-rankness of $\mathcal{Y}$. Combining the low-rank constraint via specifying a predefined rank and the sparse constraint of coefficients $c_{\boldsymbol{\alpha}}$, the sample simulation tensor $\mathcal{Y}$ is approximated by a rank-R tensor $\mathcal{X}$ via solving:

$$\begin{aligned} \min_{\{\mathbf{u}_1^{(k)},\ldots,\mathbf{u}_R^{(k)}\}_{k=1}^d} & \frac{1}{2} \|\mathbb{P}_\Omega(\sum_{r=1}^{R} \mathbf{u}_r^{(1)} \cdots \circ \mathbf{u}_r^{(d)} - \mathcal{Y})\|_F^2 \\ & + \lambda \sum_{|\boldsymbol{\alpha}|=0}^{p} |\langle \sum_{r=1}^{R} \mathbf{u}_r^{(1)} \cdots \circ \mathbf{u}_r^{(d)}, \mathcal{W}_{\boldsymbol{\alpha}} \rangle|, \end{aligned} \tag{3}$$

where $\mathbb{P}$ is a projection operator and Ω is the sampling set. An alternating direction method of multipliers (ADMM) algorithm

TABLE I
SUMMARY OF THREE TENSOR METHODS FOR UNCERTAINTY QUANTIFICATION

Method	Tensorized objective	Rank determination	Adaptive sampling	Application
Tensor recovery [7, 8]	Sample simulations	Cross validation	×	MEMS, Electronics & Photonics IC
Functional approximation [9]	Interested function	Greedy	✓ (application-specified)	SRAM failure rate estimation
Tensor regression [10, 11]	Coefficients	Automatic determined	✓	Electronics & Photonics IC

is proposed to solve (3). This method has been applied to solve high-dimensional uncertainty quantification problems in analog IC, photonic IC and MEMS design.

However, there are two unsolved fundamental questions [7]. Firstly, the tensor rank R needs to be tuned, *e.g.* via cross validation. Secondly, how to select the sample set Ω is unknown.

B. Tensorized Functional Approximation for Analog IC [9]

Assume the interested metric $y(\boldsymbol{\xi})$ is in a tensor space $\mathcal{S} = \mathcal{S}^1 \otimes \mathcal{S}^2 \otimes \mathcal{S}^d$, with $\otimes$ denotes tensor product. Let $\mathcal{T}_1 \in \mathcal{S}$ be a rank-one tensor subset. The tensor space has the property that $\mathcal{S} = \text{span}(\mathcal{T}_1)$, such that each element $y(\boldsymbol{\xi})$ in $\mathcal{S}$ can be expressed as a linear combination of rank-one tensors. In [9], the interested metric is approximated by R rank-one tensors:

$$y(\boldsymbol{\xi}) \approx \sum_{r=1}^{R} b_r \eta_r(\boldsymbol{\xi}),\ \eta_r \in \mathcal{T}_1, b_r \in \mathbb{R}, \quad (4)$$

where b_r denotes the corresponding r-th normalization constant, and η_r is a rank-one function represented as the expansion of polynomial basis functions. Since the optimal rank $R^\star$ is not known as *a-priori*, a greedy scheme is used to search for the optimal rank via increasing rank iteratively until convergence. An adaptive algorithm with sparsity constraint is used to solve each rank-one update. The low-rank functional approximation model is validated in estimating high-dimensional SRAM circuit failure rate, showing significantly faster than MC method. A domain-specified adaptive sampling method is also proposed to accelerate the failure rate estimation.

C. Tensor-Regression-based Coefficient Estimation [10, 11]

The two unsolved challenges of [7, 8] are investigated in [10, 11]. In [10, 11] the index set of Eq. (1) is chosen as the full tensor product, leading to $(p+1)^d$ basis functions. The interested metric is approximated as the inner product of two tensors $y(\boldsymbol{\xi}) \approx \langle \mathcal{X}, \mathcal{B}(\boldsymbol{\xi})\rangle$, where $\mathcal{X}$ is the coefficient tensor, and $\mathcal{B}(\boldsymbol{\xi})$ is a rank-one tensor constructed by the tensor product of univariate basis functions. To reduce the number of unknown variables, the coefficient tensor $\mathcal{X}$ is approximated by a rank-R tensor. To estimate the optimal tensor rank, a group ℓ_q/ℓ_2-norm regularization function $g(\mathcal{X})$ is used to shrink the rank. Given N sample pairs $\{\boldsymbol{\xi}_n, y(\boldsymbol{\xi}_n)\}_{n=1}^N$, the coefficients $\mathcal{X}$ can be estimated via solving a regression problem:

$$\begin{aligned} &\min_{\{\mathbf{u}_1^{(k)},\ldots,\mathbf{u}_R^{(k)}\}_{k=1}^d} h(\mathcal{X}) + \lambda g(\mathcal{X}) \text{ with} \\ &h(\mathcal{X}) = \frac{1}{2}\sum_{n=1}^{N}\left(y(\boldsymbol{\xi}_n) - \langle \textstyle\sum_{r=1}^R \mathbf{u}_r^{(1)} \cdots \circ \mathbf{u}_r^{(d)}, \mathcal{B}(\boldsymbol{\xi}_n)\rangle\right)^2, \\ &g(\mathcal{X}) = \|\mathbf{v}\|_q, v_r = \left(\sum_{k=1}^{d} \|\mathbf{u}_r^{(k)}\|_2^2\right)^{\frac{1}{2}}\ \forall r = [1, R],\ q \in (0, 1]\,. \end{aligned} \quad (5)$$

To make (5) tractable, a variational equality is employed to reformulate the problem. An efficient alternating minimization solver with analytical solutions in each sub-step is proposed to solve the optimization problem. The initial rank is set as its upper bound, and the algorithm can reduce it automatically. To further reduce the simulation cost, an adaptive sampling method combining a Voronoi-cell-based exploration and nonlinearity-based exploitation is adopted. The tensor regression method has been verified by high dimensional synthetic and realistic circuit benchmarks for both surrogate building and sensitivity analysis.

IV. REMARKS

The paper has briefly surveyed some recent tensor methods for high-dimensional uncertainty quantification of process variations in EDA and computaitonal EM. There exist some open questions that are worth further investigations. Firstly, it is desired to discover application-specific optimal tensor network structure. Secondly, although some heuristic sampling methods for tensor-based models have been proposed in some literature [10, 11], the statistical and theoretical guarantees still need further investigation. Finally, it is worth extending the tensor-based uncertainty quantification models to fit multiple interested metrics $\mathbf{y}(\boldsymbol{\xi})$ simultaneously.

We would also like to remark that tensor methods have attracted growing interest in the broader community of computational EM [4–6]. In the community of scientific computing, there have also been increasing interesting works of applying tensor methods to solve function approximation [14, 15], stochastic PDE [16], stochastic optimal control [17] problems and so forth.

ACKNOWLEDGMENT

Authors acknowledge support from NSF grants #1763699 and #1846476.

REFERENCES

[1] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*. Princeton University Press, 2010.

[2] A. Kaintura, T. Dhaene, and D. Spina, "Review of polynomial chaos-based methods for uncertainty quantification in modern integrated circuits," *Electronics*, vol. 7, no. 3, p. 30, 2018.

[3] J. Zhang, C. Zhang, F. Feng, W. Zhang, J. Ma, and Q.-J. Zhang, "Polynomial chaos-based approach to yield-driven EM optimization," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 7, pp. 3186–3199, 2018.

[4] A. C. Yücel, L. J. Gomez, and E. Michielssen, "Compression of translation operator tensors in FMM-FFT-accelerated SIE solvers via Tucker decomposition," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 2667–2670, 2017.

[5] Z. Chen, L. J. Gomez, S. Zheng, A. C. Yucel, Z. Zhang, and V. I. Okhmatovski, "Sparsity-aware precorrected tensor train algorithm for fast solution of 2-D scattering problems and current flow modeling on unstructured meshes," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 12, pp. 4833–4847, 2019.

[6] C. Qian and A. C. Yucel, "On the compression of translation operator tensors in FMM-FFT-accelerated SIE simulators via tensor decompositions," *IEEE Trans. Antennas Propag.*, pp. 1–1, 2020.

[7] Z. Zhang, T.-W. Weng, and L. Daniel, "Big-data tensor recovery for high-dimensional uncertainty quantification of process variations," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 7, no. 5, pp. 687–697, 2016.

[8] Z. Zhang, K. Batselier, H. Liu, L. Daniel, and N. Wong, "Tensor computation: A new framework for high-dimensional problems in EDA," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 36, no. 4, pp. 521–536, 2016.

[9] X. Shi, H. Yan, Q. Huang, J. Zhang, L. Shi, and L. He, "Meta-model based high-dimensional yield analysis using low-rank tensor approximation," in *Proc. Design Autom. Conf*, 2019, pp. 1–6.

[10] Z. He and Z. Zhang, "High-dimensional uncertainty quantification via active and rank-adaptive tensor regression," in *Proc. Electr. Perform. Electron. Packag. Syst.*, 2020, pp. 1–3.

[11] ——, "High-dimensional uncertainty quantification via tensor regression with rank determination and adaptive sampling," *accepted by IEEE Trans. Compon. Packag. Manuf. Technol., doi: 10.1109/TCPMT.2021.3093432, arXiv:2103.17236*.

[12] Z. Zhang, X. Yang, I. V. Oseledets, G. E. Karniadakis, and L. Daniel, "Enabling high-dimensional hierarchical uncertainty quantification by ANOVA and tensor-train decomposition," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 34, no. 1, pp. 63–76, 2014.

[13] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, 2017.

[14] A. A. Gorodetsky and J. D. Jakeman, "Gradient-based optimization for regression in the functional tensor-train format," *J. Comput. Phys.*, vol. 374, pp. 1219–1238, 2018.

[15] C. Cui and Z. Zhang, "Uncertainty quantification of electronic and photonic ICs with non-Gaussian correlated process variations," in *Proc. Intl. Conf. Computer Aided Design*, 2018, pp. 1–8.

[16] S. Dolgov, B. N. Khoromskij, A. Litvinenko, and H. G. Matthies, "Polynomial chaos expansion of random coefficients and the solution of stochastic partial differential equations in the tensor train format," *SIAM/ASA J. Uncertain. Quantif.*, vol. 3, no. 1, pp. 1109–1135, 2015.

[17] A. Gorodetsky, S. Karaman, and Y. Marzouk, "High-dimensional stochastic optimal control using continuous tensor decompositions," *Int. J. Robot. Res.*, vol. 37, no. 2-3, pp. 340–377, 2018.