High-Dimensional Sparse Surrogate Construction via Bayesian Compressive Sensing

Khachik Sargsyan *, Cosmin Safta †, Habib N. Najm †, Bert Debusschere †, Daniel Ricciuto ‡, Peter Thornton ¶

* ‡ Sandia National Laboratories
Livermore, CA
ksargsy@sandia.gov, csafta@sandia.gov, hnnajm@sandia.gov, bjdebus@sandia.gov

‡ Oak Ridge National Laboratory
Oak Ridge, TN
ricciuto@ornl.gov, thorntonpe@ornl.gov

ABSTRACT

Due to slow convergence, classical Monte-Carlo approaches are ineffective for computationally intensive studies of complex models as they require prohibitively many sampled simulations for reasonable accuracy. Targeting high-dimensional systems, we build computationally inexpensive surrogate models in order to accelerate both forward (e.g., uncertainty propagation and sensitivity analysis) and inverse (e.g., calibration) uncertainty quantification studies. We apply Polynomial Chaos (PC) spectral expansions to build surrogate relationships between output quantities and model parameters using as few forward model simulations as possible. For a complex model with a large number of input parameters, building a PC surrogate model is challenged by high dimensionality: there is typically insufficient model simulation data as well as a prohibitively large number of spectral basis terms. Bayesian compressive sensing (BCS) approach [1] is employed in order to detect a sparse polynomial basis set that best captures the model outputs. We enhance the BCS algorithm with iterative basis growth and reweighing that effectively searches polynomial space for an optimal, sparse basis set. Besides proof-of-concept studies for synthetic models, the technique is demonstrated on the Community Land Model with more than 50 input parameters. The outcome of the algorithm is then employed for forward uncertainty propagation and variance-based sensitivity analysis, leading to dimensionality reduction [2]. Furthermore, the computationally inexpensive surrogates greatly accelerate statistical methods for parameter estimation, where one relies on observational data to estimate input parameters with quantified uncertainty, using Markov Chain Monte Carlo sampling.

REFERENCES
