This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

SAND2019-3656C

Sparse PCE surrogate assisted inversion algorithm for ultra-deep electromagnetic resistivity logging-while-drilling data

Han Lu , Jiefu Chen, Xuqing Wu, Xin Fu, University of Houston; Mohammad Khalil, Cosmin Safta, Sandia National Laboratories; Yueqin Huang, Cyentech Consulting LLC

Summary

The recent development of azimuthal electromagnetic (EM) resistivity logging-while-drilling (LWD) tool has greatly extended the depth of investigation and brought new challenges to inversion and interpretation of logging measurements. The results by deterministic methods like Levenberg-Marquadt algorithm (LMA) are oftentimes strongly affected by chosen initial models. The development of surrogates/meta-models in optimization problems, combined with deterministic methods, offers the opportunity to enhance the performance of the earth model interpretation. In this paper, we construct a surrogate based on polynomial chaos expansion (PCE) for the ultra-deep azimuthal electromagnetic resistivity LWD measurements and use it to perform a two-stage pixel-based inversion. A Bayesian compressive sensing (BCS) technique is employed to efficiently build a sparse PCE surrogate for high-dimensional systems. The results show that the sparse PCE surrogates are helpful in avoiding local minima during the inversion.

Introduction

Geosteering is the proactive control of a wellbore placement based on the downhole LWD measurements, aiming at maximizing the economic production from the well. The earth model parameters can be inferred by the inversion of LWD measurements such as the azimuthal electromagnetic resistivity logging data. The geosteering inversion problem is always challenging because (a) the inverse problems in geosteering are inherently nonlinear and ill-posed; (b) measurements are usually sparse; (c) LWD measurements are noisy. In 1D pixel-based inversion, the earth model is divided into dozens or even hundreds of thin layers, resulting in large number of model parameters to be inferred. The traditional gradient-based (local) optimization methods such as LMA takes the advantage of differentiable objective function, converges rapidly but only finds the local minimum, so usually the results are strongly affected by the initial guess.

Surrogates or meta-models, as the approximation of the original experiments or computer simulations, have been extensively used in optimization problems involving computationally-intensive models (Goel et al., 2007). There are a number of studies comparing the performance of different surrogate models, including radial basis functions, kriging, polynomial chaos expansions, and artificial neural networks, as well as their recommendations for specific applications (see Simpson et al., 2001, for example). In this paper, the original forward model is replaced by a PCE surrogate to expedite the inversion process. The PCE is used to represent an arbitrary random variable, under certain conditions, as a spectral expansion in terms of known random variables. The high-dimensionality of the set of input (design) variables leads to a prohibitively large number of basis terms in the PCE representation, and thus the number of required sample points increases exponentially. This is referred as the "curse of dimensionality". To combat this challenge, Sargsyan et al. developed an algorithm to construct a sparse PCE surrogate (Sargsyan et al., 2014) by "learning" and retaining the most relevant basis terms with the aid of Bayesian compressive sensing (Babacan et al., 2010).

In general, the accuracy, robustness, efficiency, and computational simplicity are important criteria for evaluating the usefulness of surrogates. However, empirical observations show that large approximation errors do not always mislead the optimization (Jin, 2011). Furthermore, one can complement surrogates with high-fidelity forward model simulations. This is often known as surrogate model management. Based on these observations, we propose a two-stage optimization method for the geosteering inversion assisted by sparse PCE surrogates. We will show that the PCE surrogates aid in finding a better initial set of parameters and consequently the inversion accuracy is greatly improved.

Polynomial chaos surrogate

Consider a forward computational model $z = f(\lambda)$, where $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_{\tilde{d}})$ is a d-dimensional input vector and z is a scalar output. In PCE theory (Wiener, 1938; Ghanem and Spanos, 2003), both the input parameters and the output of interest are represented as a series involving orthogonal polynomials $\Psi_k(\xi)$ in terms of independent and identically distributed random variables $\xi = (\xi_1, \xi_2, \dots, \xi_d)$. As such, the inputs can be written as

$$\lambda_i \cong \sum_{k=0}^{K_{in}-1} \lambda_{i,k}\, \Psi_k(\xi) \tag{1}$$

and the output is written as

$$z(\lambda(\xi)) \cong \sum_{l=0}^{K-1} c_l\, \Psi_l(\xi) \tag{2}$$

where $\lambda_{i,k}$, $i = 1, 2, \dots, d$, $k = 0, 1, \dots, K_{in} - 1$, and c_l, $l = 0, 1, \dots, K - 1$, are the coefficients that completely characterize the random variables. K and K_{in} are the number of basis terms in the input and output PC expansions, respectively. For surrogate construction, K_{in} is usually fixed, while the value of K is decided by the output accuracy requirements.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

It is noted that the dimension of the input vector $\tilde{\boldsymbol{d}}$ does not need to be equal to $\boldsymbol{d}$. The finite PCEs tend to lead $\boldsymbol{d}$ to be at least as large as, and often larger than $\tilde{\boldsymbol{d}}$ (O'Hagan, 2013). The type of polynomials $\boldsymbol{\Psi}_k(\boldsymbol{\xi})$ used is chosen based on the distribution of $\boldsymbol{\xi}$ to satisfy orthogonality. For example, Hermite polynomials are used for normally distributed random variables and Legendre polynomials are used for uniformly distributed random variables. For surrogate construction purposes, one can assume that $\boldsymbol{\lambda}_i$ is a linear PCE of the individual $\boldsymbol{\xi}_i$ to sufficiently cover the range of interest for each parameter. Subsequently, an available dataset of input-output pairs $\{\boldsymbol{\lambda}_i, \boldsymbol{z}(\boldsymbol{\lambda}_i)\}_{i=1}^{N}$ can be easily transferred to $\boldsymbol{\mathcal{D}} = \{\boldsymbol{\xi}, \boldsymbol{z}_i\}_{i=1}^{N}$ and what remains is determining $\boldsymbol{c}_k$ with the given data.

There exist intrusive and non-intrusive methods for the calculation of the polynomial coefficients $\boldsymbol{c}_k$. Intrusive methods are only applicable to those forward models that have explicit equations, and thus not suitable for geosteering logging problems.

Alternatively, non-intrusive methods do not require an explicit representation of the forward model but treat it as a black box. Projection-based non-intrusive methods provide the coefficients as

$$c_k = \langle z, \psi_k \rangle / \langle \psi_k, \psi_k \rangle \tag{3}$$

where $\langle \psi_k, \psi_k \rangle$ is in practice known exactly, $\langle z, \psi_k \rangle$ can be calculated by numerical integration

$$\langle z, \psi_k \rangle = \int z\, \psi_k(\xi) p(\xi) d\xi \tag{4}$$

One can use Monte Carlo sampling or quadrature rules to sample from the distribution $p(\xi)$. However, in high-dimensional problems, both methods require a large number of simulation runs even with sparse sampling techniques, rendering those approaches impractical for the inversion problems under consideration.

Alternatively, eq. (2) can be treated as a regression model. Since the PCE is only an approximation of an error term is added to the equation. Therefore, we have

$$\boldsymbol{z} = \sum_{k=0}^{K-1} \boldsymbol{c}_k \boldsymbol{\Psi}_k(\boldsymbol{\xi}) + \boldsymbol{\varepsilon} \tag{5}$$

The regression coefficients $\boldsymbol{c}_k$ can then be computed using linear regression.

Bayesian compressive sensing for PCE

From the Bayesian point of view, the solution to the problem of determining the coefficients $\boldsymbol{c}_k$ is a posterior probability density function $\boldsymbol{q}(\boldsymbol{c})$. Based on Bayes' formula, we can write:

$$\boldsymbol{q}(\boldsymbol{c}) \propto \boldsymbol{L}_{\boldsymbol{\mathcal{D}}}(\boldsymbol{c}) \boldsymbol{p}(\boldsymbol{c}) \tag{7}$$

Where $\boldsymbol{L}_{\boldsymbol{\mathcal{D}}}(\boldsymbol{c})$ is the likelihood, a measure of a goodness-of-fit of the corresponding surrogate with respect to the given data $\boldsymbol{\mathcal{D}}$, and $\boldsymbol{p}(\boldsymbol{c})$ is the prior distribution of $\boldsymbol{c}$. Given a zero-mean normal distributed noise model with standard deviation $\boldsymbol{\sigma}$, we can write the likelihood as

$$\boldsymbol{L}_{\boldsymbol{\mathcal{D}}}(\boldsymbol{c}) = \left(2\pi\boldsymbol{\sigma}^2\right)^{-\frac{N}{2}} \exp\left(-\sum_{i=1}^{N} \frac{(z_i - z_{c(\xi_i)})^2}{2\sigma^2}\right) \tag{8}$$

This problem becomes intractable for high dimensional problems since the number of unknown coefficients will grow rapidly with increasing dimensionality. In many practical applications applications, most of the basis functions in a PCE will have near-zero coefficients, i.e. the vector $\boldsymbol{c}$ is sparse. It is efficient and reasonable to only recover the most significant terms of the PCE, both in the construction and evaluation of the PCE surrogates. To this end, (Sargsyan et al., 2014) proposed to use Bayesian compressive sensing (BCS) to compute a sparse representation of $\boldsymbol{c}$ given available data (Ji et al., 2008; Babacan et al., 2010). The key of inferring a sparse PCE is to impose a prior distribution on $\boldsymbol{c}$ that induces sparsity. A commonly used sparsity-inducing prior is the Laplace prior

$$\boldsymbol{p}(\boldsymbol{c}) = \left(\frac{\boldsymbol{\alpha}}{2}\right)^{K+1} \exp\left(-\boldsymbol{\alpha} \sum_{k=0}^{K} \|\boldsymbol{c}\|\right) \tag{9}$$

The vector $\boldsymbol{c}$ that maximizes the posterior $\boldsymbol{q}(\boldsymbol{c})$ in eq. (7) coincides with the solution of the classical compressive sensing problem

$$\arg\max_{\boldsymbol{c}} \left(\log \boldsymbol{L}_{\boldsymbol{\mathcal{D}}}(\boldsymbol{c}) - \boldsymbol{\alpha}\|\boldsymbol{c}\|_1\right) \tag{10}$$

where the regularization term $\boldsymbol{\alpha}\|\boldsymbol{c}\|_1$ corresponds to the sparsity-inducing prior distribution. The positive parameter $\boldsymbol{\alpha}$ is a user-defined value that controls the level of sparsity. However, the Laplace distribution is not conjugate to the Gaussian likelihood defined in eq. (8) and thus does not allow a tractable Bayesian analysis. This issue was addressed in sparse Bayesian learning, particularly with the Relevance vector machine (RVM) (Tipping, 2001). Instead of directly using the Laplace prior, a hierarchical prior distribution is constructed by with a Gaussian prior distribution on $\boldsymbol{c}$

$$\boldsymbol{p}(\boldsymbol{c}_k | \boldsymbol{s}_k^2) = \frac{\boldsymbol{1}}{\sqrt{2\pi \boldsymbol{s}_k^2}} \exp\left(-\frac{\boldsymbol{c}_k^2}{2\boldsymbol{s}_k^2}\right) \tag{11}$$

and a gamma prior to the hyper-parameter $\boldsymbol{s}_k^2$

$$\boldsymbol{p}(\boldsymbol{s}_k^2 | \boldsymbol{\alpha}^2) = \frac{\boldsymbol{\alpha}^2}{\boldsymbol{2}} \exp\left(-\frac{\boldsymbol{\alpha}^2 \boldsymbol{s}_k^2}{\boldsymbol{2}}\right) \tag{12}$$

with a resulting (marginalized) Laplace prior density

$$\begin{aligned} \boldsymbol{p}(\boldsymbol{c} | \boldsymbol{\alpha}^2) &= \int_0^\infty \prod_{k=0}^{K-1} \boldsymbol{p}(\boldsymbol{c}_k | \boldsymbol{s}_k^2)\, \boldsymbol{p}(\boldsymbol{s}_k^2 | \boldsymbol{\alpha}^2) d\boldsymbol{s}_k^2 \\ &= \prod_{k=0}^{K-1} \frac{\boldsymbol{\alpha}}{\boldsymbol{2}} \boldsymbol{e}^{-\boldsymbol{\alpha}|\boldsymbol{c}_k|} \end{aligned} \tag{13}$$

This procedure has been implemented in the Bayesian LASSO method (Hans, 2009). For details, see (Babacan et al., 2010; Tipping et al., 2003).

PCE-assisted pixel-based geosteering inversion

The accuracy is usually an import criterion of evaluating a surrogate. However, empirical observations show that a surrogate with large approximation errors can be helpful when used for optimization, as shown in Figure 1.

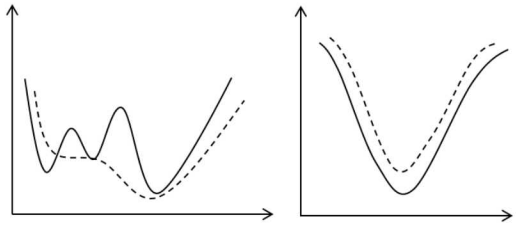

Figure 1: Examples of surrogates that have large approximation errors but are good for global minimum search.

This is particularly true a for high-dimensional non-linear optimization problem since there might be a large number of local optima in the solution space. A low-order polynomial-based surrogate (PCE), albeit possessing large errors, is smoother in nature. Therefore, such a surrogate with large approximation error should be used in combination with the original forward model in order to get an accurate solution. We propose a two-stage inversion method with the work flow shown in Figure 2.

As shown in Figure 2, a sparse PCE surrogate for the original forward model is constructed using Bayesian compressive sensing. At stage 1, the surrogate replaces the original forward model in the inversion process. The surrogate used in stage 1 does not need to be highly accurate, as the low-order sparse PCE surrogate captures the smooth trends in the response surface, and the solution at this stage is considered to be close to the global optimal solution. At stage 2, we perform the deterministic inversion process with the actual forward model using the initial values obtained from stage 1. Compared to conventional deterministic inversion algorithms, this surrogate-assisted two-stage strategy may have a better chance of escaping local minima and reaching the global minima (i.e. optimal solution).

Simulations

Our method framework utilizes the C++ library named "Uncertainty Quantification toolkit" (UQTk) (Debusschere et al., 2016), developed by Sandia National Laboratories. We construct PCE surrogates for the forward model that computes the ultra-deep azimuthal EM resistivity LWD tool responses in layered formation at 2kHz, 6kHz and 24kHz working frequencies, giving 72 measurement curves as the output. For each curve, we construct an independent sparse 2nd-order PCE using Bayesian compressive sensing. The input earth model parameters are transferred into pixel-based models, offering more flexibility compared to model-based inversion since the number of layers does not need to be prespecified (therefore avoids the risk of overestimating or underestimating the complexity of the earth model for inversion). The details of the input data are given in Table 1.

Table 1: input parameter configurations

Number of layers	52
Layer thickness	2ft
Layer resistivity	0.1 ohm·mm ~ 300 ohm·mm
Depth of investigation	-50ft ~ 50ft

The database has size $\mathbf{1.6 \cdot 10^5}$ and contains model parameters of length 52 and corresponding tool signal for each PCE. We tested the original LMA inversion and PCE-assisted LMA inversion with the tool signals generated by a 7-layer model with resistivity values of 2 ohm · m, 50 ohm·m, 200 ohm·m, 100 ohm · m, 20 ohm · m, 5 ohm·m and 1 ohm·m from the top to the bottom, as shown in the up panel of Figure 3. The thickness of the layers varies from 2ft to 12ft. The inversion is conducted at every 10ft while the total working region extends to 200ft horizontally. The drilling tool moves from the top layer to the bottom layer, as represented by the dotted lines. For deterministic methods, multi-start local searches scheme is suggested (Haftka et al., 2016), which is commonly used in industry. Therefore, we assign 8 initial models of different resistivity and number of layers to both methods and run in parallel, the final result is the one with the smallest data misfit. The results in Figure 3 indicate that the PCE-assisted inversion recovers the earth model better than the conventional LMA. Meanwhile, due to its sparsity, the PCE evaluation time of one input vector is only several milliseconds for all the 72 curves, being approximately 2% of the original forward model execution time.

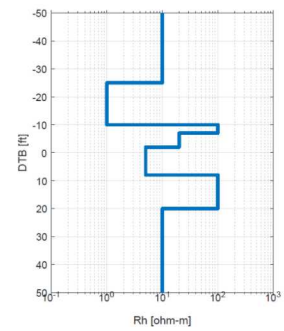

Figure 4: The vertical profile of formation resistivity.

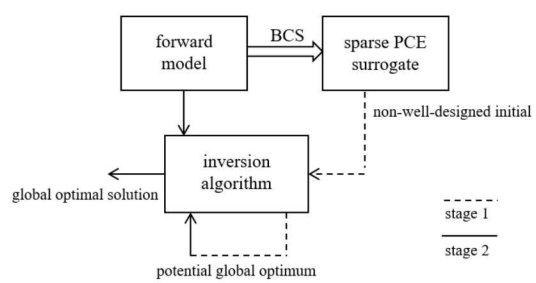

Figure 2: PCE-assisted inversion work flow.

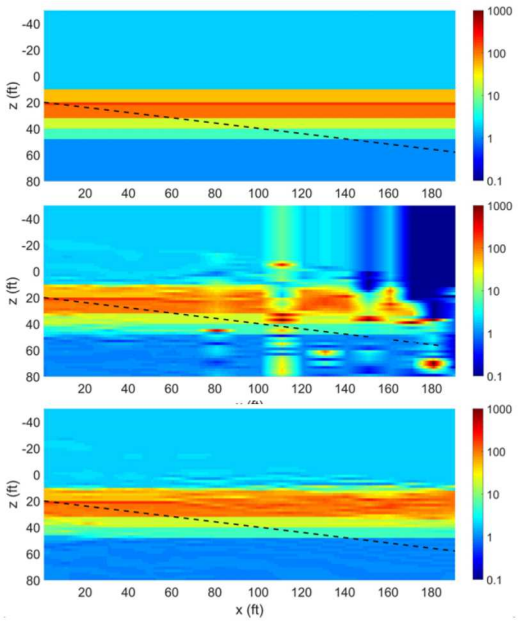

Figure 3: Top: The real earth model. Middle: Model reconstructed by only using original forward model. Bottom: Model reconstructed by PCE-assisted two-stage inversion method.

To further demonstrate how the PCE surrogates help in reaching the the optimal solution, we show the multi-start inversion results of a single logging point. The y-axis represents the distance to the boundary and y = 0 is the location of the tool. The x-axis represents the resistivity in log scale. The real model is shown in Figure 4. In this test, 15 initial models of different types are assigned to both methods. We picked the best 8 results (i.e. those with smaller data misfit) from the 15 results obtained by each method and show in Figure 5. The left panel shows that in the original inversion, different initial models lead to very different solutions. This indicates that many of these initial models lead to local optimum. On the other hand, the solutions by the proposed method shown in the right panel are highly consistent with each other and also with the ground truth.

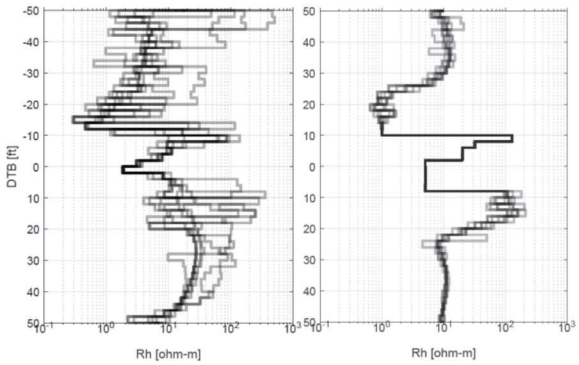

Figure 5: Left: The inversion results from the original inversion. Right: The inversion results from PCE-assisted inversion.

Conclusion

We proposed a sparse PCE surrogate assisted inversion algorithm for ultra-deep azimuthal EM resistivity logging data. Numerical results suggest that the sparse PCE surrogates aid in reaching the globally-optimal solution.

Acknowledgement

This material is based upon work supported by the U.S. Department of Energy, Office of Science, and Office go Advanced Science Computing Research, under Award Numbers DE-SC0017033. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

References

Babacan, S. Derin, Rafael Molina, and Aggelos K. Katsaggelos. "Bayesian compressive sensing using Laplace priors." *IEEE Transactions on Image Processing* 19, no. 1 (2010): 53-63.

Debusschere, Bert, Khachik Sargsyan, Cosmin Safta, and Kenny Chowdhary. "Uncertainty quantification toolkit (UQTk)." *Handbook of Uncertainty Quantification* (2016): 1-21.

Ghanem, Roger G., and Pol D. Spanos. *Stochastic finite elements: a spectral approach*. Courier Corporation, 2003.

Goel, Tushar, Raphael T. Haftka, Wei Shyy, and Nestor V. Queipo. "Ensemble of surrogates." *Structural and Multidisciplinary Optimization* 33, no. 3 (2007): 199-216.

Haftka, Raphael T., Diane Villanueva, and Anirban Chaudhuri. "Parallel surrogate-assisted global optimization with expensive functions–a survey." *Structural and Multidisciplinary Optimization* 54, no. 1 (2016): 3-13.

Hans, Chris. "Bayesian lasso regression." *Biometrika* 96, no. 4 (2009): 835-845.

Ji, Shihao, Ya Xue, and Lawrence Carin. "Bayesian compressive sensing." *IEEE Transactions on signal processing* 56, no. 6 (2008): 2346.

Jin, Ruichen, Wei Chen, and Timothy W. Simpson. "Comparative studies of metamodelling techniques under multiple modelling criteria." *Structural and multidisciplinary optimization* 23, no. 1 (2001): 1-13.

Jin, Yaochu. "Surrogate-assisted evolutionary computation: Recent advances and future challenges." *Swarm and Evolutionary Computation* 1, no. 2 (2011): 61-70.

O'Hagan, Anthony. "Polynomial chaos: A tutorial and critique from a statistician's perspective." *SIAM/ASA J. Uncertainty Quantification* 20 (2013): 1-20.

Sargsyan, Khachik, Cosmin Safta, Habib N. Najm, Bert J. Debusschere, Daniel Ricciuto, and Peter Thornton. "Dimensionality reduction for complex models via Bayesian compressive sensing." *International Journal for Uncertainty Quantification* 4, no. 1 (2014).

Simpson, Timothy W., J. D. Poplinski, Patrick N. Koch, and Janet K. Allen. "Metamodels for computer-based engineering design: survey and recommendations." *Engineering with computers* 17, no. 2 (2001): 129-150.

Tipping, Michael E. "Sparse Bayesian learning and the relevance vector machine." *Journal of machine learning research* 1, no. Jun (2001): 211-244.

Tipping, Michael E., and Anita C. Faul. "Fast marginal likelihood maximisation for sparse Bayesian models." In *AISTATS*. 2003.

Wiener, Norbert. "The homogeneous chaos." *American Journal of Mathematics* 60, no. 4 (1938): 897-936.