ENHANCED RECOVERY OF SUBSURFACE GEOLOGICAL STRUCTURES USING COMPRESSED SENSING AND THE ENSEMBLE KALMAN FILTER

Furrukh Sana[a], K. Katterbauer[b], Tariq Al-Naffouri[a], I. Hoteit[a,b]

[a]Dept. of Computer, Electrical and Mathematical Sciences and Engineering,
[b]Dept. of Physical Sciences and Engineering,
King Abdullah University of Science & Technology, 23955-9600 Thuwal, Saudi Arabia

ABSTRACT

Recovering information on subsurface geological features, such as flow channels, holds significant importance for optimizing the productivity of oil reservoirs. The flow channels exhibit high permeability in contrast to low permeability rock formations in their surroundings, enabling formulation of a sparse field recovery problem. The Ensemble Kalman filter (EnKF) is a widely used technique for the estimation of subsurface parameters, such as permeability. However, the EnKF often fails to recover and preserve the channel structures during the estimation process. Compressed Sensing (CS) has shown to significantly improve the reconstruction quality when dealing with such problems. We propose a new scheme based on CS principles to enhance the reconstruction of subsurface geological features by transforming the EnKF estimation process to a sparse domain representing diverse geological structures. Numerical experiments suggest that the proposed scheme provides an efficient mechanism to incorporate and preserve structural information in the estimation process and results in significant enhancement in the recovery of flow channel structures.

Index Terms— Subsurface Characterization, Compressed Sensing, Ensemble Kalman Filter, Orthogonal Matching Pursuit, K-SVD.

1. INTRODUCTION

Geological structures are formed by complex processes within the earth's crust. This leads to the formation of flow channels that exhibit high permeability, as compared to the low permeability surrounding rock formations [1]. These channels serve as preferred paths for fluid propagation and estimating their locations and the structures holds significant importance for optimizing well placement and the productivity of reservoirs. However, reconstruction of subsurface geological features from low-resolution data has been a great challenge due to the limited number of observations available to identify their spatial distribution [2].

The Ensemble Kalman filter (EnKF) is a popular Monte Carlo variant of the standard Kalman filter (KF) for the estimation of subsurface reservoir model parameters, such as permeability [3]. However, due to its least-squares formulation, the conventional EnKF framework often fails to recover and preserve flow channel structures [4].

Compressive sensing (CS) has revolutionized the field of signal and image processing, providing efficient tools to improve the reconstruction quality of signals and fields exhibiting sparsity in some domain [5]. CS principals allow for enhanced recovery of sparse domain signals from a reduced number of observations by exploiting the prior information on signal sparsity.

For the general reservoirs under consideration, the fraction of the reservoir occupied by high permeability channels is rather limited and enables formulation of the reconstruction of subsurface geological features as a sparse field recovery problem. Jafarpour et. al. investigated the use of the well known Wavelet transforms in [2] and the Discrete Cosine Transform (DCT) in [4] for sparse representation under an EnKF framework. However, the gains from representation in these sparse domains are largely limited to the reduction in computational complexity. In [6], a comprehensive study was presented on the influence of the sparse geologic dictionaries used as prior information in an inversion process, concluding that diverse geologic dictionaries lead to robust estimation attributed to the connectivity exhibited by the dictionary elements.

We propose a new scheme based on compressed sensing framework to enhance the estimation and recovery of the subsurface geological features using the EnKF. The proposed scheme transforms the EnKF estimation to a specifically constructed sparse domain based on learned geological dictionaries. Numerical experiments suggest that the proposed method allows to efficiently incorporate and preserve the structural information in the EnKF estimation process and results in significant improvement in recovering flow channel structures.

The paper is organized as follows. Section 2 present the problem formulation with a brief overview of the EnKF and the components of sparse parameterization framework. The proposed scheme is introduced in Section 3. Section 4 describes the experimental setup and presents the simulation results from numerical experiments. The paper concludes in

Section 5 with a brief summary of the main results.

2. SPARSE PARAMETERIZATION FRAMEWORK

In order to achieve enhanced reconstruction of geological features from low resolution data, we formulate our estimation problem as a sparse field recovery problem. In this section, we first present a brief overview of the EnKF and then the components of the sparse parameterization framework, namely the Orthogonal Matching Pursuit (OMP) and the K-SVD algorithms.

2.1. The Ensemble Kalman Filter

The Ensemble Kalman Filter based reservoir characterization problem can be formulated using the state-space model

$$x_{k+1} = M_k(x_k) + \eta_k, \tag{1}$$

and the observation equation

$$y_k = h_k(x_k) + \epsilon_k, \tag{2}$$

where M_k represents the reservoir forward propagation model given by the MRST reservoir simulator [7]. The reservoir simulator couples a well model to the two-phase flow problem for the oil and water phases. The state vector x_k consists of the static parameters, permeability and porosity, and dynamic variables, pressure and saturation, and η_k is a stochastic term representing the model error. y_k is the observation that is related to the state via a nonlinear observation operator h_k, perturbed by a random noise ϵ_k. The observation is composed of the production data such as the flux, the bore hole pressure (bhp), and the water cut ratio (wct). Both η_k and ϵ_k are assumed independent and Gaussian of zero mean and with covariance matrices Q_k and R_k respectively.

First introduced by Evensen et. al. [3], the EnKF is fundamentally based on the Kalman Filter (KF) and differs from the KF in that the distribution of the system state is represented by a collection, or ensemble, of state vectors approximating the KF estimate and its error covariance matrix by the ensemble sample mean and covariance. This ensemble representation helps when dealing with nonlinear models and in overcoming prohibitive manipulation of the KF covariance matrices. For further details on the theoretical and implementation aspects of the EnKF, the reader may refer to [3].

2.2. OMP Algorithm

Given an estimate of sparsity k for a signal (or field) represented by a vector of unknown weights $\mathbf{x}$ in a d-dimensional space, Orthogonal Matching Pursuit (OMP) algorithm attempts to reconstruct the data signal $\mathbf{y}$ using a linear combination of at most k basis elements (atoms) from an overcomplete dictionary Ψ. The dictionary is a collection of basis elements representing a particular domain for signal representation and can be chosen to suit a particular reconstruction problem [8]. The OMP algorithm solves the optimization problem

$$\min_{\mathbf{x}} \parallel \mathbf{x} \parallel_{\mathbf{0}} \quad \textbf{subject to} \quad \parallel \mathbf{y} - \mathbf{\Psi}\mathbf{x} \parallel_2 < \eta, \tag{3}$$

by iteratively projecting the data onto a dictionary and finding the basis elements that are most correlated with the residuals. The term $\parallel \mathbf{x} \parallel_{\mathbf{0}}$ denotes the $\ell_0 - norm$ and represents the sparsity level of signal $\mathbf{x}$, i.e. the number of its non-zero elements, and η is the error tolerance. The inputs to the OMP algorithm are the N-dimensional data vector $\mathbf{y}$, an $N \times d$ observation matrix $\mathbf{\Psi}$ (dictionary) and the estimate of sparsity k for the signal $\mathbf{x}$. The complete algorithm can be found in [9]. OMP is used in the context of our proposed scheme to enforce a sparse representation for the unknown field.

2.3. Dictionary Learning

A key factor in sparse estimation techniques is the identification of a domain (basis) in which the signal, or field, of interest can be modeled as sparse. Sparse domain implies that any signal in that domain could be represented by a linear combination of only a few basis elements [5]. Discrete Cosine Transform (DCT) [4] and Wavelets [2] are two of such sparse domains. However, a sparse domain can also be constructed to suit a particular application, and algorithms used for such purpose are called dictionary learning algorithms [10]. Given a set of training signals (or fields) $\mathbf{Y}$ of size N, dictionary learning algorithms attempt to find the basis elements to construct a dictionary $\mathbf{\Psi}$ whose sparse linear combination can represent each of the signals (or fields) in the training set $\mathbf{Y}$. The weights of this sparse linear combination, for each signal in the training set, are then given by the corresponding vector of coefficients in a set $\mathbf{X}$. Mathematically, this is done by solving the optimization problem [10]

$$\{\mathbf{X}, \mathbf{\Psi}\} = \arg\min_{\mathbf{X}, \mathbf{\Psi}} \sum_{i=1}^{N} \parallel \mathbf{y}_i - \mathbf{\Psi}\mathbf{x}_i \parallel_2^2 + \gamma \parallel \mathbf{x}_i \parallel_1, \tag{4}$$

$$\textit{s.t.} \quad \|\psi_i\|_2^2 \leq 1, \quad \forall i = 1, ..., N,$$

where $\mathbf{y_i}$ and $\mathbf{x_i}$ represent the individual training signals and the coefficient vectors in $\mathbf{Y}$ and $\mathbf{X}$, respectively. $\psi_\mathbf{i}$ represents the atoms of the dictionary and γ is a parameter that penalizes the ℓ_1-norm of $\mathbf{x_i}$. The optimization problem in (4) consecutively attempts to find both the atoms for the dictionary $\mathbf{\Psi}$ and their corresponding weights $\mathbf{X}$ such that the representation error is minimized for all signals in the training set $\mathbf{Y}$. This is a non-convex problem that is difficult to solve, even approximately [10].

K-SVD is a dictionary learning algorithm that solves an ℓ_0-norm equivalent of the problem defined in equation (4) using a recursive two-step approach [8]. For more details on the K-SVD algorithm, the reader may refer to [8]. In Section 3 and 4, we will outline how the K-SVD algorithm can be used to generate a sparse basis for the subsurface parameter estimation problem.

3. CS ENHANCED ENSEMBLE KALMAN FILTER

By spatially transforming the EnKF ensemble to a sparse domain, sparsity can be incorporated into the EnKF estimate. The K-SVD algorithm can be used for this purpose to generate a new sparse basis capable of representing any subsurface geological structure using a sparse linear combination of its basis elements. A large set of realizations, each representing a subsurface geological structure, can be used to train the K-SVD algorithm in order to generate this new sparse basis. The OMP algorithm can then be used to transform the ensemble members into the new sparse domain by representing them using a sparse linear combination of basis elements from the new dictionary. The coefficients (or weights) $\chi_{\mathbf{k}}^{\mathbf{f}}$ used for this sparse linear combination represent the ensemble members in the new sparse domain. OMP allows to enforce sparsity such that only a few coefficients in the new domain are non-zero. These coefficients can now be used in the EnKF update process. This new sparse domain of subsurface geological structures thus becomes a natural choice for the sparse representation of ensemble members and offers a more robust domain of estimation compared to the DCT and Wavelet domains due to the connectivity and structural information exhibited by its basis elements.

The proposed algorithm, referred to as *Sparse Geological Structures Domain EnKF* (SGSD-EnKF) from here onwards, is listed in Algorithm 1.

Algorithm 1 Sparse Geological Structures Domain EnKF (SGSD-EnKF) Algorithm

---

- **Input:** Forecast ensemble set $\mathbf{X}_{\mathbf{k}}^{\mathbf{f}}$, dictionary $\mathbf{\Psi}$, Observation data $\mathbf{y}$, target sparsity k,

- **Obtain Sparse Domain Coefficients:**

$$[\chi_{\mathbf{k}}^{\mathbf{f}}] = \mathbf{OMP}(\mathbf{X}_{\mathbf{k}}^{\mathbf{f}}, \mathbf{\Psi}, \mathbf{k})$$

- **Perform EnKF Analysis Step:**

$$\chi_{\mathbf{k}}^{\mathbf{a}} = \mathbf{EnKF}(\chi_{\mathbf{k}}^{\mathbf{f}}, \mathbf{y})$$

- **Transform back to Spatial Domain:**

$$\mathbf{X}_{\mathbf{k}}^{\mathbf{a}} = \mathbf{\Psi} * \chi_{\mathbf{k}}^{\mathbf{a}}$$

- **Output:** $\mathbf{X}_{\mathbf{k}}^{\mathbf{a}}$

---

4. EXPERIMENTAL RESULTS

The studied reservoir structure is resolved in dimensions of 50 spatial units in both x and y-directions. The flow channels exhibiting high permeability values are surrounded by low permeability rock formations as shown in Figure 2. The channels have permeability values of 5000 millidarcy, with the surrounding rock formation having a permeability value of 100 millidarcy.

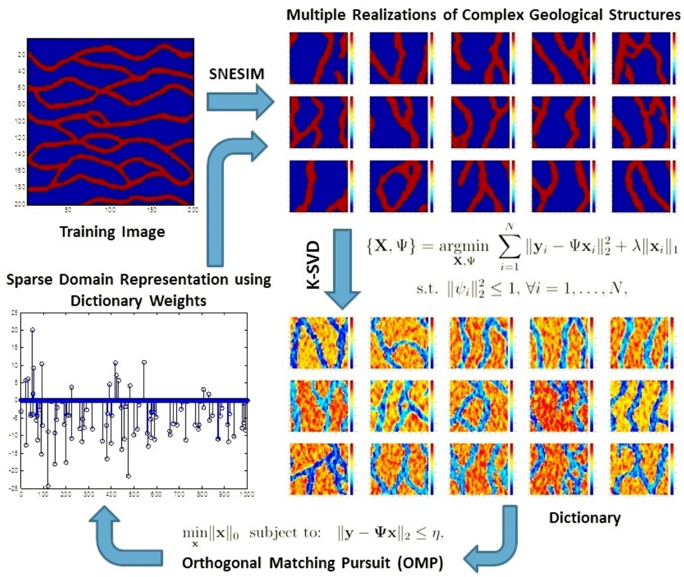

Fig. 1. Process illustrating the generation of multiple subsurface geological structures (top-right) from the training image (top-left) and the dictionary (bottom-right) for representing a field using a weighted sparse linear combination (bottom-left)

The training image in Figure 1 was used to generate a large set of 2000 images each depicting different sparse geological structures. These realizations were generated using the S-GeMS software which utilizes the SNESIM (*Single Normal Equation SIMulation*) algorithm. These realizations are used as a training set for the K-SVD algorithm to generate the dictionary of 1000 linearly independent basis elements for the SGSD scheme of Section 3. The size of the state vector in the new sparse domain is hence reduced to $N = 1000$ compared to $N = 2500$ in the spatial domain. The optimum sparsity rate for the OMP algorithm was found to be 10%. The process of generating the realizations set from the training image and the dictionary from these realizations is illustrated in Figure 1.

In order to assess the accuracy of recovering the channel structure and locations, we have utilized the SSIM criterion [11] to compare the structural information in the reference field with that estimated by our proposed CS enhanced EnKF scheme. A high SSIM value indicates more similarity between the two fields and hence good quality of estimation.

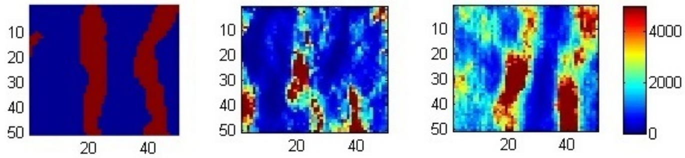

Fig. 2. Reference permeability field (left), Estimated field using the standard EnKF (middle), Estimated field using the proposed scheme (right).

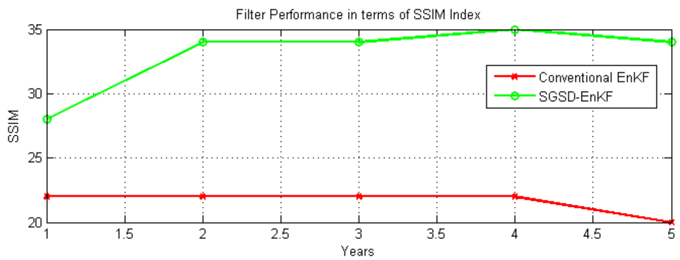

Fig. 3. Structural similarity between the reference and estimated fields over the 5 year simulation period

Figure 2 presents the reference field to be estimated. The well configuration possesses one injector well at the center of the field and four producing wells at the corners. Bottom hole pressure (BHP), water cut ratio (WCR), and production flux were measured at all wells every 30 days and assimilated into the reservoir simulator for a total simulation period of 5 years with a ensemble size of 80 members.

The estimated permeability fields at the end of the 5-years simulation period using the standard EnKF filter and the technique proposed herein are also presented in Figure 2. As evident from the images, there is a clear mismatch in the channel recovery as it results from the standard EnKF where only small traces of the channels are recovered at the bottom of the domain, but at inaccurate locations. The performance of the proposed technique is however much better, progressively improving over the simulation period as indicated by the SSIM values plotted in Figure 3. The recovered field exhibit similar structure to the reference field while utilizing only 5 data points.

5. CONCLUSION

We proposed a new compressed sensing framework for enhanced recovery of subsurface geological structures using the EnKF. The proposed scheme enables problem representation using much less number of coefficients and incorporation of useful prior information into the estimation process. Numerical results demonstrate the effectiveness of the proposed scheme in recovering the flow channels structures and show clear advantage over the conventional EnKF framework.

6. REFERENCES

[1] Michael J Ronayne, Steven M Gorelick, and Jef Caers, “Identifying discrete geologic structures that produce anomalous hydraulic response: An inverse modeling approach,” *Water Resources Research*, vol. 44, no. 8, pp. W08426, 2008.

[2] Behnam Jafarpour, “Wavelet reconstruction of geologic facies from nonlinear dynamic flow measurements,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 5, pp. 1520–1535, 2011.

[3] Geir Evensen, “The Ensemble Kalman Filter: Theoretical formulation and practical implementation,” *Ocean Dynamics*, vol. 53, no. 4, pp. 343–367, Nov. 2003.

[4] Behnam Jafarpour and Dennis B McLaughlin, “History matching with an ensemble Kalman filter and discrete cosine parameterization,” *Computational Geosciences*, vol. 12, no. 2, pp. 227–244, 2008.

[5] Emmanuel J Candès, Justin Romberg, and Terence Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489–509, 2006.

[6] Mohammadreza Mohammad Khaninezhad, Behnam Jafarpour, and Lianlin Li, “Sparse geologic dictionaries for subsurface flow model calibration: Part I. Inversion formulation,” *Advances in Water Resources*, vol. 39, pp. 106–121, 2012.

[7] K Lie, S Krogstad, I Ligaarden, and Others, “Open-source MATLAB implementation of consistent discretisations on complex grids,” *Computational Geosciences*, vol. 16, no. 2, pp. 297–322, 2012.

[8] M. Aharon, M. Elad, and A. Bruckstein, “ K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation,” *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[9] Joel A Tropp and Anna C Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.

[10] A. Elsheikh, M. Wheeler, and I. Hoteit, “Sparse calibration of subsurface flow models using nonlinear orthogonal matching pursuit and an iterative stochastic ensemble method,” *Advances in Water Resources*, vol. 56, no. 0, pp. 14–26, 2013.

[11] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.